

{ ISMP
2024 }

July 21-26

25th International Symposium on
Mathematical Programming

M O N T R É A L

Complete program



<https://ismp2024.gerad.ca/>

25th International Symposium on Mathematical Programming (ISMP)

Montréal, July 21-26, 2024

Table of contents

Welcome to ISMP2024	4
Organizing committees	6
Sponsors	9
General information	11
Special events	13
Schedule	16
Invited speakers	17
Parallel sessions - ordered by streams	25
Parallel sessions - ordered chronologically	62
Abstracts	84
Participants	340
Floor plan and map to the banquet	352

Welcome to ISMP2024



It is with great pleasure that we extend a warm welcome to all 1500 attendees to Montréal for the triennial International Symposium on Mathematical Programming (ISMP), the flagship conference of the Mathematical Optimization Society (MOS).

This will be the 25th edition of the symposium. While ISMP's presence in Montréal in 2024 is a significant occasion, it's not the symposium's first visit to our city. In 1979, the symposium was held here, featuring a remarkable scientific program. ISMP draws together researchers from across the globe alongside industry professionals in the field of mathematical optimization. Here, participants present their latest findings, engage in discussions on emerging challenges, and bridge theory with practical applications.

This edition stands as a testament to the collaborative spirit of the international community. The organization of the symposium was piloted by a small committee from Polytechnique Montréal and supported by many larger ones. The scientific program, led by Michel Gendreau, features an impressive lineup of invited speakers. Under the leadership of Gilles Savard, sponsors and partners have been diligently secured, ensuring the continued success of the symposium. Guy Desaulniers oversees finance matters, ensuring prudent management of resources and budgets. Nadia Lahrichi leads the social events, cultivating an environment that encourages networking among participants. We would like to extend our thanks to the local organizers from the many universities in Montréal, and to the team of 84 researchers from around the world, who worked together to recruit speakers to fill the sessions across the 28 mathematical optimization streams.

We extend our gratitude to all our sponsors whose generous support will undoubtedly make our event a resounding success. Your contributions have played a pivotal role in bringing the symposium to life. Specifically, we would like to thank our Gold Sponsors for financing Thursday's banquet, which hopefully will create a memorable and enjoyable experience. Additionally, we express our thanks to our Silver Sponsors for their sponsorship of the welcome reception.

The contributions of the support services of our institutions, Polytechnique Montréal and HEC Montréal, and of the local organizing team were crucial to the seamless running of the event. We extend our heartfelt gratitude to the GERAD administrative staff for the invaluable support, and to the scheduling team for their efforts in rigorously optimizing the schedule with up to 33 presentations in parallel, and to our colleagues in building the special issue in the Mathematical Programming journal. Finally, we express our deepest appreciation to the dedicated cohort of volunteer students (recognizable by cyan bandanas) and staff who play a crucial role in the logistical organization during the ISMP week.

The success of ISMP lies in your hands. Your scientific contributions enrich the program and pave the way for fruitful collaborations. We thank you for your participation and encourage you to make the most of this congress, seizing the opportunity to learn, connect, initiate collaborations and to enjoy the beautiful attractions that the city has to offer.

On behalf of the organizing committee, I hope that you will enjoy the symposium.

Charles Audet
ISMP2024 Chair





Dear colleagues,

On behalf of the Mathematical Optimization Society (MOS), it is my pleasure to welcome you to the 25th International Symposium on Mathematical Programming.

The MOS is the international society for everyone interested in mathematical optimization. It was launched just over 50 years ago, in 1973, and the triennial International Symposia on Mathematical Programming (ISMP) are our flagship conferences. The Symposia bring the mathematical optimization community together and provide a dedicated forum for the presentation of research results in the theory, algorithms, computation, modelling and applications of mathematical optimization. The Symposia began in 1951; they have been held (almost) every three years and generally, every second meeting has been held in North America. It is fitting to hold this 25th Symposium in Canada, and to be back in Montréal where the 10th Symposium was held in 1979.

Mathematical optimization is a highly active and exciting field. The large number and variety of streams at the Symposium reflect the dynamism of our community, covering areas that have been central to the MOS for decades, such as nonlinear optimization, combinatorial optimization and graph theory, as well as areas that are increasing rapidly in importance, such as quantum computing, multi-stage stochastic programming and reinforcement learning. The Symposium is the opportunity to recognize the best work of our community via the range of triennial prizes that we award at every Symposium. Please join us for the Opening Ceremony during which we will announce the award recipients.

The breadth of the field is reflected in the range of the Society's publications and conferences. Our flagship journal Mathematical Programming began publication in 1970, nearly 55 years ago. The more recently launched Mathematical Programming Computation is celebrating this year its 15th anniversary. Jointly with SIAM, we publish the MOS/SIAM Book Series on Optimization that includes research monographs, textbooks at all levels, books on applications, and tutorials. Our newsletter Optima keeps us abreast of recent technical developments and activities in our field. Beyond the Symposium, the Society supports in different ways several specialized conferences, notably on combinatorial optimization, continuous optimization, mixed integer optimization, and stochastic optimization.

The future of the MOS is bright. I am delighted that more than one third of our attendees are students or postdocs. To them I say: You are the future of mathematical optimization! We value you tremendously and want to support you at every stage of your career. We acknowledge that as a community, we face significant challenges, notably the need to improve the peer-review process for publication and the simplistic use of publication metrics for academic hiring and promotion. Two panel discussions during the Symposium will provide fora to discuss them. I encourage everyone, and especially our early career researchers, to participate in the panels and to engage with our attempt to steer a course for the future success of our members.

Thank you to all the members of the organization of the ISMP for their sustained efforts in ensuring the success of this Symposium. Your contribution to the community and to the Society is important and greatly appreciated. Thank you also to all the attendees; the Symposium would not be a success without your participation.

In closing, I wish you an engaging and fruitful Symposium and a wonderful time in Montréal!

Miguel F. Anjos
Chair, Mathematical Optimization Society



Mathematical
Optimization Society

Organizing committees

Conference chair

Charles Audet, Polytechnique Montréal and GERAD

Steering committee

Guy Desaulniers, Polytechnique Montréal and GERAD
Michel Gendreau, Polytechnique Montréal and CIRRELT
Nadia Lahrichi, Polytechnique Montréal and CIRRELT
Gilles Savard, Polytechnique Montréal and GERAD

Local committee

Daniel Aloïse
Fabian Bastin
Margarida Carvalho
Marilène Cherkesly
Erick Delage
Youssef Diouane
Franklin Djeumou Fomeni
Emma Frejinger
Alain Hertz
Tim Hoheisel
Sébastien Le Digabel
Antoine Legrain
Dominique Orban
Courtney Paquette
Frédéric Quesnel
Walter Rei

Technical staff

Karine Hébert
Marilyne Lavoie
Marie Perreault
Charles Bédard

Scheduling

Mathieu Gervais-Dubé
Alain Hertz
Sébastien Le Digabel
Antoine Legrain

MP special issue

Youssef Diouane
Franklin Djeumou Fomeni
Alain Hertz
Dominique Orban
Courtney Paquette

Scientific committee

Cluster 1 : Continuous optimization

1a: Nonlinear Programming

Frank E. Curtis	Daniel Robinson	Andreas Wächter
-----------------	-----------------	-----------------

1b: Global Optimization

Eligius Hendrix	Sergiy Butenko	Mohit Tawarmalani
-----------------	----------------	-------------------

1c: Nonsmooth Optimization

Michael Friedlander	Marc Teboulle	Dima Drusviatskyi
---------------------	---------------	-------------------

1d: Semi-Definite Programming

Henry Wolkowicz	Christoph Helmberg	Angelika Wiegele
-----------------	--------------------	------------------

1e: Variational Analysis, Variational Inequalities and Complementarity

Walaa Moursi	Radu Bot	Patrick Mehlitz
--------------	----------	-----------------

1f: Random Methods for Continuous Optimization

Lin Xiao	Peter Richtarik	Simon Lacoste-Julien
----------	-----------------	----------------------

1g: Derivative-free and Simulation-based Optimization

Stefan Wild	Ana Custodio	Francesco Rinaldi
-------------	--------------	-------------------

1h: Optimal Control, PDE Constrained Optimization & Multi-level Methods

Roland Herzog	Michael Ulbrich	Scott MacLachlan
---------------	-----------------	------------------

Cluster 2 : Discrete optimization

2a: Mixed Integer Linear Programming

Ivana Ljubic	Ricardo Fukasawa	Yuri Faenza
--------------	------------------	-------------

2b: Mixed Integer Non-linear Programming

Claudia D'Ambrosio	Gonzalo Muñoz	Dan Bienstock
--------------------	---------------	---------------

2c: Combinatorial Optimization and Graph Theory

Bernard Ries	Hadrien Mélot	Celina de Figueiredo
--------------	---------------	----------------------

2d: Machine Learning and Discrete Optimization

Thibaut Vidal	Andrea Lodi	Dolores Romero Morales
---------------	-------------	------------------------

2e: Approximation and Online Algorithms

David P. Williamson	Mohit Singh	Thomas Erlebach
---------------------	-------------	-----------------

2f: Computational Discrete and Integer Optimization

Timo Berthold	Ambros Gleixner	Marco Luebbecke
---------------	-----------------	-----------------

2g: Constraint Programming

Louis-Martin Rousseau	Andre Cire	David Bergman
-----------------------	------------	---------------

Cluster 3: Optimization under uncertainty**3a: Continuous Stochastic Programming**

Guzin Bayraksan

Claudia Sagastizábal

Huifu Xu

3b: Discrete Stochastic Programming

Merve Bodur

Jim Luedtke

Ward Romeijnders

3c: Robust and Distributionally Robust Optimization

Melvyn Sim

Dick den Hertog

Ruiwei Jiang

3d: Multi-stage Stochastic Programming and Reinforcement learning

Raghu Pasupathy

Andy Philpott

Ilya O. Ryzhov

3e: Data-driven optimization

Phebe Vayanos

Zuo-Jun Max Shen

Paul Grigas

Cluster 4: Applications and computation**4a: Networks and Telecom**

Arie Koster

Luis Gouveia

Bernard Fortz

4b: Transportation and logistics

Daniele Vigo

Maria Gracia Speranza

Emma Frejinger

4c: Scheduling and manufacturing

Ruben Ruiz

Tom McCornick

Raf Jans

4d: Energy and Environment

Miguel Anjos

Luce Brotcorne

Golbon Zakeri

4e: Healthcare and life sciences

Stefan Nickel

Paola Cappanera

Nadia Lahrichi

4f: Finance and social sciences

Pavlo Krokhmal

Nisheeth K. Vishnoi

Geneviève Gauthier

4g: High performance implementation and quantum computing

Ted Ralphs

Tobias Achterberg

Giacomo Nannicini

4h: Game Theory

Alexandra Schwartz

Bernhard von Stengel

Adrian Vetta

4i: Industrial applications

Antoine Legrain

Daniel Aloise

{ ISMP 2024 } GOLD SPONSORS

Sponsors of Thursday's Banquet



{ ISMP 2024 } SILVER SPONSORS

Sponsors of Sunday's Welcome Reception



{ ISMP 2024 } BRONZE SPONSORS



General information

Venue

The symposium is hosted at the [Palais des Congrès de Montréal](#). There are many entrances to access the venue.

Main entrance: [1001 Place Jean-Paul-Riopelle](#)
Direct access from Place-d'Armes metro station
201 Viger Avenue West
301 Saint-Antoine Street West
200 Viger Avenue West
Jeanne-Mance and Viger
Jeanne-Mance and Saint-Antoine
1000 Saint-Urbain Street
De la Gauchetière Street (corner of Esplanade)



Registration

Registration and badge pickup is on the second floor of The Palais des Congrès (210 Préfonction Ouest, above the Jean-Paul Riopelle entrance). The desk is open on Sunday July 21th from 14h to 18h, and throughout the symposium from Monday to Friday.

Wifi

Wifi is available at the Palais des Congrès (at 15Mbps per device), in all ISMP conference rooms and corridors.

SSID: ISMP2024
Password: optimization

Projection setup

Each conference room is equipped with a computer connected to a projector. Speakers can either upload their talks on site on the computer using a USB drive, or can use the HDMI connector (one needs to disconnect it from the computer and connect your own). We recommend that the files be transferred before the beginning of the session. A volunteer student is assigned to each session to assist the file transfer, if necessary.

Files in the conference room computers will be deleted after the presentations.

Instructions to chairpersons

Please arrive at your lecture room 10 minutes before the scheduled start time of the session. Collect all presentations, preferably in PDF format, onto the room computer and ensure that video retransmission is functioning properly for each presentation. It is imperative to follow the designated time slots from the conference program. Parallel presentations are 30 minutes long, including 5 minutes for questions and including the setup time. As session chair, it is your responsibility to ensure that speakers respect the schedule.

In case of a cancelled talk in a session, the other talks need to remain at the scheduled time.

In case the chair is absent, please nominate a chair among the present speakers of the session.

Optimizing the ISMP schedule

Approximately 1400 talks and meetings have been scheduled for the 2024 edition of ISMP. The symposium runs over 5 days, and 38 rooms with different sizes were available. The talks are partitioned into four main clusters and 30 streams. Many sessions were pre-formed by the scientific committee, but most of the talks were submitted by the participants without specifying a session. Each registered presentation includes a list of keywords and suggested streams. The scheduled was designed in three steps. i- A mixed integer linear programming model is used to assign each presentation to a stream and a time slot. Constraints include unavailability of speakers, non-ubiquity, preferred ordering of sessions and limiting simultaneous sessions. ii- Multiple re-optimizations were conducted to handle latecomers and accommodate additional requirements, to alter the schedule as little as possible. A Gurobi license was graciously offered. iii- Each session was assigned to a room by taking into account the predicted stream popularity, and by avoiding room change within a session. The talk [TC953](#) describes the methodology.

Toward a greener symposium

The organizing committee is committed to making the symposium as environmentally friendly as possible. To achieve this, we minimized the printing of documents, providing all materials electronically and avoiding any printed handouts for our numerous participants. Additionally, reusable cups, glasses, and plates were used during coffee breaks and social events.

We also surveyed a subset of participants to determine their preference between receiving various gifts or using the corresponding funds to offset the carbon footprint from travel to Montréal. We are proud to announce that 86% favored making a donation to an environmental organization. We will gladly make such a contribution.

We hope future organizers will continue striving to make ISMP as green as possible.

Special events

Welcome reception

From 17h30 to 20h on Sunday the 21th, a welcome reception will be hosted for all participants at the seventh floor of the Palais des Congrès in the Chapiteau (room 720, see map on page [352](#)). A pair of tickets for free beverages will be given at the registration desk on the second floor during the registration period (from 14h to 18h). Plenty of food will be offered.

Banquet

For those that have selected it when registering, the conference banquet will be held at the Belvédère in the Old Port of Montréal [2 Rue de la Commune Ouest, Montréal, QC H2Y 4B2](#). It is a 15-minute walk from the Palais des Congrès. Remember to bring your badge, as it is your ticket to enter the banquet. The event is sold out.

Farewell reception

On Friday at 16h, instead of the usual coffee breaks, a farewell reception will be held in the second floor (exhibit room 210DH). It will be the occasion to say goodbye with some refreshments.

Opening ceremony

On Monday 9h in the room 210A-F, the opening ceremony of ISMP2024 features the Prize Awards:

- The Paul Y. Tseng Memorial Lectureship in Continuous Optimization, for outstanding contributions in continuous optimization.
- The A.W. Tucker Prize, for an outstanding thesis in mathematical programming.
- The Lagrange Prize in Continuous Optimization, for outstanding work in continuous optimization. Sponsored jointly by MOS and SIAM.
- The Beale - Orchard Hays Prize, for outstanding work in computational mathematical programming.
- The Delbert Ray Fulkerson Prize, for outstanding papers in discrete mathematics. Sponsored jointly by MOS and AMS.
- The George B. Dantzig Prize, for original research having a major impact on mathematical optimization. Sponsored jointly by MOS and SIAM.

Exhibits

The Room 210DH adjacent to the Plenary and Keynote auditoriums will host the coffee breaks, as well as the following exhibits.

- AMPL
- CAE
- Cardinal Operations (COPT)
- Gurobi
- Hexaly
- Mathematical Science / Cambridge University Press
- SIAM
- Springer

Board meetings

MOS council and editors of the following journals are invited to participate to their board meetings in Room 516B (except MPC which is held outside of the Palais des Congrès).

- Computational Optimization and Applications (COAP)
Monday 12h20
- Mathematical Optimization Society (MOS) council meeting
Monday 18h30
- Optimization and Engineering (OE)
Tuesday 12h20
- Mathematical Programming Computation (MPC)
Tuesday 18h30(15th anniversary celebration)
- Mathematical Programming (MP)
Wednesday 12h20

Business meeting

Open to all, two business meetings will be hosted to discuss societal issues.

- Mathematical Optimization Society (MOS) business meeting
Tuesday 17h20 Room 210CG
- Stochastic Programming Society (SPS) business meeting
Wednesday 18h30 Room 516D

Panels

Open to all, two panels will be hosted to discuss specific issues.

- Improving journal review process – Wednesday 16h20 Room 516C

The review process for top optimization journals seems to be changing. Papers are getting longer with more technical content while reviewers find themselves lacking time and incentives to read papers in depth. From an editor's standpoint, it seems that securing reliable and thorough reviews has become harder in recent years. While, ideally, reviewing can be intellectually rewarding, more recognition and/or encouragement may be needed to maintain a high-quality review process. In this session, we will review the key issues and hold a panel on the topic of reviewer recognition with editors from SIOPT, MPA, MPC, and MOR. A significant amount of time will be dedicated to Q&A with the audience.

- Publications as a metric in promotion and hiring – Thursday 16h20 Room 516C

We will discuss recent changes in the publication culture in the field of optimization and discuss what some deem as undesirable effects of these changes. Specifically, in the last one or two decades, we have been seeing a substantial increase in the number of papers, both posted online and submitted to conferences and journals. The ease at which work (published or not) can be made visible and the rise of online metrics such as Google Scholar puts pressure on researchers (especially junior ones) to focus on their quantity of papers. More importantly, this pressure is reinforced by the evaluation processes for hiring and promotion in academia and research labs, as well as the evaluation process of grant proposals submitted to funding agencies. We will review some data collected via recent surveys and hold a panel consisting of senior faculty members with relevant experience. A significant amount of time will be dedicated to Q&A with the audience.

A.W. Tucker Prize

The speakers of this Tucker Prize session will be announced during the opening ceremony.
Tuesday 8h30 Room 513F

Paul Y. Tseng Memorial Lectureship in Continuous Optimization

Thursday 16h20 Room 511F

In memory of Jean-Louis Goffin

Wednesday, during all three parallel sessions, in Room 511F

1. Recent advances in large scale optimization
2. Analytic Centre Cutting Plane Methods: The contributions of J-L Goffin's work
3. A tribute to J-L Goffin's academic influence

Remote presentations

A session was added on Friday morning to allow participants who could not attend the meeting to remotely present their talk.

Schedule

	Sun, Jul 21		Mon, Jul 22		Tue, Jul 23		Wed, Jul 24		Thu, Jul 25		Fri, Jul 26			
8:00			Registration 8:00 - 9:00 P2100											
8:30														
9:00			Opening Ceremony 9:00 - 10:30		Parallel Sessions 8:30 - 10:30		Parallel Sessions 8:30 - 10:30		Parallel Sessions 8:30 - 10:30		Parallel Sessions 8:30 - 10:30	Remote Session 8:30 - 10:30		
9:30														
10:00														
10:30			Coffee Break (20 mins) 10:30 - 10:50 - Room 210DH											
10:50			Plenary 10:50 - 12:20		Plenary 10:50 - 12:20		Plenary 10:50 - 12:20		Plenary 10:50 - 12:20		Plenary 10:50 - 12:20			
12:20	Lunch (1h40min) 12:20 - 14:00													
14:00	Registration 14:00 - 17:50 - Room P2100		Keynote 14:00 - 15:00	Parallel Sessions 14:00 - 16:00	Keynote 14:00 - 15:00	Parallel Sessions 14:00 - 16:00	Keynote 14:00 - 15:00	Parallel Sessions 14:00 - 16:00	Keynote 14:00 - 15:00	Parallel Sessions 14:00 - 16:00	Keynote 14:00 - 15:00	Parallel Sessions 14:00 - 16:00		
15:00			Keynote 15:00 - 16:00		Keynote 15:00 - 16:00		Keynote 15:00 - 16:00		Keynote 15:00 - 16:00		Keynote 15:00 - 16:00			
16:00			Coffee Break (20 mins) 16:00 - 16:20 - Room 210DH and 5th floor										Farewell Reception 16:00 - 18:00 - Room 210DH	
16:20			Keynote 16:20 - 17:20	Parallel Sessions 16:20 - 17:50	Keynote 16:20 - 17:20	Parallel Sessions 16:20 - 17:50	Keynote 16:20 - 17:20	Parallel Sessions 16:20 - 17:50	Keynote 16:20 - 17:20	Parallel Sessions 16:20 - 17:50				
17:00														
17:20														
17:30														
17:50	Welcome Reception 17:30 - 20:00 - Room 720													
18:00														
18:30										Banquet 18:30 See on Map				
20:00														
	Sun, Jul 21		Mon, Jul 22		Tue, Jul 23		Wed, Jul 24		Thu, Jul 25		Fri, Jul 26			

Legend: The following syntax is used for the parallel sessions:

<Stream id> - <Time code> - <Room> - <Session name>

For example, the first stream of Monday PM is:

<1a> - <MA11> - <516D> - <Advances in Large-Scale Nonlinear Optimization I>

- <Stream id> is described on pages 7 and 8.
- <Time code> The first symbol is the day of the week in {M,T,W,Th,F}; the second is the time slot A = 8:30-10:30, B = 14:00-16:00 and C = 16:20-17:50; the third is a reference number in N.
- <Room> (see page 353) and <Session name> are self explanatory.

Invited speakers

Plenary speakers

Tamás Terlaky, Lehigh University

Monday 10:50 to 12:20 – 210A-F, Chair: Miguel Anjos

Quantum Computing and Optimization, and Interior Point Methods

This talk briefly reviews the current state of quantum computing (QC) hardware, and the opportunities and challenges quantum computing offers in solving optimization problems. The Quantum Computing (QC) and the Interior Point Methods (IPM) revolutions inspire novel challenges and novel methodologies. Optimization is in the heart of the quest to evidence quantum advantage. However, the inexactness and condition number dependence characteristics of NISQ (Noisy Intermediate-Scale Quantum) devices forced us to think differently about solving optimization problems.

Considering IPMs for linear and semi-definite optimization (LO and SDO) problems, QC inspired to design Inexact Infeasible and Inexact Feasible Primal-Dual, and Inexact Dual IPM variants. These are novel algorithms in both the QC and classic computing environments. Enhancing Quantum Interior Point Methods (QIPMs) with Iterative Refinement (IR) leads to exponential improvements in the worst-case overall running time of QIPMs, compared to previous best-performing QIPMs. We also discuss how the proposed IR scheme can be used in classical inexact IPMs with conjugate gradient methods. Further, the proposed IR scheme exhibits quadratic convergence for LO and SDO towards an optimal solution without any assumption on problem characteristics. On the practical side, IR can be useful to find precise solutions while using inexact LO and SDO solvers.

Coralía Cartis, University of Oxford

Tuesday 10:50 to 12:20 – 210A-F, Chair: Dominique Orban

Tensor methods for nonconvex optimization

We investigate the use of high-order polynomial local models within adaptive regularization frameworks for solving nonconvex optimization problems. Though tensor-based models and methods have preoccupied optimization researchers in the past, this topic is experiencing a recent resurgence, possibly originating with the development of carefully-regularized variants of Newton’s method. It was shown, in both the convex and nonconvex worlds, that regularization algorithms allow higher-than-second derivatives to be naturally incorporated into their local model construction, while controlling convergence by means of a regularization parameter present in these models. Furthermore, the global rate of convergence/worst-case evaluation complexity of the ensuing methods proved to be optimal within their respective algorithmic classes that use corresponding derivative orders or their approximations. We review these developments, as well as establish connections to polynomial optimization and related techniques, in our efforts to efficiently minimise polynomial subproblems, to local or global optimality. A robust adaptive procedure for the choice of the regularization parameter is key to numerical performance, and we detail some of the intricacies involved in its design. Finally, we bring these ingredients together, in our quest to develop practical tensor methods, with exact and inexact derivatives, subproblem solutions and adaptive parameters, and ask the challenging question: are tensor methods preferable to second-order ones in numerical practice?

David Morton, Northwestern University

Co-authors: Nazlican Arslan, Oscar Dowson, Bernardo Pagnoncelli

Wednesday 10:50 to 12:20 – 210A-F, Chair: Michel Gendreau

Markov decision processes and multi-stage stochastic programming: modeling and computation

We incorporate modeling features from Markov decision processes (MDPs) in a class of multi-stage stochastic programs. This class includes structured MDPs with continuous states and actions. We extend Markov-chain based policy graphs to include a form of statistical learning and to include decision-dependent one-step transition probabilities. We illustrate the expressiveness of our modeling approach. To computationally handle these new features, we extend the growing class of stochastic dual dynamic programming algorithms.

Ozlem Ergun, Northeastern University

Thursday 10:50 to 12:20 – 210A-F, Chair: Margarida Carvalho

Optimizing global food aid operations

We describe our efforts in optimizing global food aid operations using mathematical programming in collaboration with UN World Food Program and USAID for over 15 years.

The World Food Programme (WFP) is the largest humanitarian agency fighting hunger worldwide, reaching approximately 90 million people with food assistance across 80 countries each year. We describe a mixed integer linear programming model that simultaneously optimizes the food basket to be delivered, the sourcing plan, the delivery plan, and the transfer modality of a long-term recovery operation for each month in a predefined time horizon. By connecting traditional supply chain elements to nutritional objectives, we are able to make significant breakthroughs in the operational excellence of WFP's most complex operations. We show three examples of how the optimization model is used to support operations: (1) to reduce the operational costs in Iraq by 12% without compromising the nutritional value supplied, (2) to manage the scaling-up of the Yemen operation from three to six million beneficiaries, and (3) to identify sourcing strategies during the El Niño drought of 2016.

Each year, the Bureau for Humanitarian Assistance (BHA) under the U.S. Agency for International Development (USAID) plays a vital role in the global distribution of food aid. However, meeting the increasing demand with finite resources remains a persistent challenge. We examine the USAID global food aid supply chain, focusing on its multi-echelon, multi-commodity, and multi-modal nature. We adopt a multi-stage stochastic programming model to optimize warehouse locations and safety stock levels. This model integrates pre-disaster pre-positioning with post-disaster inventory replenishment, effectively catering to both deterministic (ongoing) and stochastic (sudden-onset) demands. Leveraging Variational Recurrent Autoencoders (VRAE) and Long Short-Term Memory (LSTM) networks, our approach enhances demand pattern analysis and realistic scenario planning. We aim to optimize BHA/FFP's food aid procurement, prepositioning, and distribution strategies, thereby enhancing the resiliency and responsiveness of USAID's global food aid operations.

Nelson Maculan, Federal University of Rio De Janeiro
Co-authors: Marcia Fampa, Hacène Ouzia, Renan Pinto
Tuesday 10:50 to 12:20 – 210A-F, Chair: Daniel Aloise

Mathematical Optimization Models for the Euclidean Steiner Tree Problem in $\mathbb{R}^d$

In this talk, we review the mathematical optimization models for the Euclidean Steiner Tree Problem (ESTP) in dimensions proposed in the literature. The development of such models for the ESTP began in the late 1990s. The ESTP is a mixed integer nonlinear optimization problem with a history dating back to the 17th century. Several properties of its optimal solutions are well known, but it is still a big challenge to encode these properties in its modeling, aiming for its numerical resolution with branch-and-bound algorithms. We identify some of the difficulties and present the modeling techniques used in the literature to overcome them.

Keynote speakers

Frank E. Curtis, Lehigh University
Monday 14:00 to 15:00 – 210CG, Chair: Youssef Diouane

Stochastic Algorithms for Nonconvex Constrained Optimization

I will discuss recent work with my collaborators on the design and analysis of stochastic algorithms for solving constrained optimization problems involving continuous objective and constraint functions that may be nonconvex. The algorithms can be viewed as extensions of the stochastic-gradient method from the unconstrained to the constrained setting. I will motivate this work through applications including physics-informed learning and fair learning, where in particular we have found benefits of using our Newton-based (i.e., sequential quadratic optimization (SQP) and interior-point) methods rather than penalty-based approaches. I will summarize the convergence guarantees offered by our methods, discuss computational challenges and how to overcome them, and present the results of numerical experiments.

Russell Luke, Georg August University of Goettingen
Monday 15:00 to 16:00 – 210CG, Chair: Tim Hoheisel

Random Algorithms for Large-Scale Optimization: from expansive Markov operators to stochastic tomography in X-ray imaging

Randomization of algorithms is an established technique for dealing with large scale optimization problems. The convergence analysis, however, is limited for the most part to convex problems and almost sure convergence of the ergodic sequence or similar, and as such the conventional analytic strategies only apply to a subset of important problems. Viewing these algorithms as Markov chains on a space of probability measures allows one to go beyond almost sure convergence, but here the state of the art for the theory is not much better, being limited to Markov operators that are contractions or, until only recently, nonexpansive. In the convergence theory for deterministic algorithms, we have shown that calmness of set-valued mappings, balanced with metric subregularity at fixed points suffices to guarantee local convergence with rates of any algorithm that can be written as a fixed point iteration. Lifting these properties to their stochastic analogs on probability measure spaces yields a general, but conceptually simple, framework for the analysis of broad class of stochastic algorithms that provides sufficient conditions for local convergence in distribution, with rates, in wide variety of settings. To demonstrate the potential of this approach, we survey two new results enabled by our framework. One of these relates to the computation of Fréchet means of phylogenetic trees in a Hadamard space. The second result, which we spend a little more time with, concerns convergence in distribution of a classical stochastic gradient descent algorithm applied to a nonconvex stochastic tomography problem in X-ray imaging.

Xiaojun Chen, The Hong Kong Polytechnic University
Monday 16:20 to 17:20 – 210CG, Chair: Tim Hoheisel

Nonsmooth Nonconvex-Nonconcave Min-Max Optimization Problems

This talk considers nonsmooth nonconvex-nonconcave min-max optimization problems with convex feasible sets. We discuss the existence of local saddle points, global minimax points and local minimax points, and study the optimality conditions for local minimax points. We show the existence of local saddle points and global minimax points of the convex-concave saddle point problem with cardinality penalties and the relations with its continuous relaxation problems. Moreover, we give an explicit formula for the value function of the inner maximization problem of a class of robust nonlinear least square problems and complexity bound for finding an approximate first order stationary point. A smoothing quasi-Newton subspace trust region algorithm is presented for training generative adversarial networks as nonsmooth nonconvex-nonconcave min-max optimization problems. Examples of retinal vessel segmentation in fundoscopic images are used to illustrate the efficiency of the algorithm.

Pascal Van Hentenryck, Georgia Tech
Tuesday 14:00 to 15:00 – 210CG, Chair: Emma Frejinger

Machine Learning for Power Systems

Machine learning has the potential to transform how power systems are planned and operated. This talk reviews some recent developments in machine learning for power systems, focusing on the concept of trustworthy optimization proxies that provide approximations to optimization models in milliseconds. The talk presents primal proxies that provide near-optimal feasible solutions to Optimal Power Flow (OPF) and Security-Constrained OPF, and dual proxies that deliver feasible dual bounds to the same problems. These optimization proxies can be integrated into simulation platform that performs, for the time, real-time risk assessment of power systems. At the technical level, the talk reviews the concept of differentiable programming, self-supervised learning, compact optimization learning, and input-convex neural networks. The practicality of the proposed techniques is demonstrated on power systems with up to 30,000 buses.

Sven Leyffer, Argonne National Laboratory
Tuesday 15:00 to 16:00 – 210CG, Chair: Sébastien Le Digabel

Unifying nonlinearly constrained optimization

Nonlinearly constrained optimization problems are commonly solved using iterative methods. We present a double loop framework that allows us to express a broad range of state-of-the-art nonlinear optimization solvers within a common framework. The main components of this framework include strategies for computing a descent direction and mechanisms that promote global convergence. We start by introducing an abstract framework with four common ingredients that describes most nonlinear optimization methods and unifies their workflows. We then present Uno, a modern, lightweight and extensible C++ solver that unifies the workflow of most derivative-based iterative optimization methods. Uno is meant to enable researchers to experiment with novel optimization strategies while leveraging established subproblem solvers and interfaces to modeling languages. We demonstrate that Uno is competitive with state-of-the-art solvers, and illustrate its power by combining and extending existing solvers.

Stefan Irnich, Johannes Gutenberg University Mainz
Tuesday 16:20 to 17:20 – 210CG, Chair: Guy Desaulniers

Recent developments in branch-price-and-cut algorithms

The success of optimization algorithms that are based on column generation can be attributed to four aspects: (1) an extensive formulation often has a stronger linear-relaxation bound compared to the original formulation (the extensive formulation is the one with a large number of variables derived by a Dantzig-Wolfe decomposition from an original model); (2) an extensive formulation can eliminate inherent symmetry; (3) column-generation subproblems can cope with non-linearities of the original formulation; and (4) powerful algorithms have been developed to solve subproblems for important real-world and theoretical problems. The talk discusses important developments in branch-price-and-cut algorithms, giving examples from applications in vehicle routing and scheduling, cutting and packing, and problems in networks.

Veronica Piccialli, Sapienza University of Rome
Wednesday 14:00 to 15:00 – 210CG, Chair: Daniel Aloise

Enhancing Machine Learning through Mathematical Programming: A Journey from Unsupervised to Semi-supervised Learning

The integration of mathematical optimization with machine learning represents a significant advancement in addressing complex learning tasks. While many machine learning models rely on optimization for training, the use of tailored optimization techniques has shown particular promise in handling non-convex formulations. This interdisciplinary approach has led to substantial improvements in solving challenging machine learning problems, particularly those involving unsupervised or semi-supervised learning paradigms. The topic of this talk reinforces the synergy between optimization and machine learning, showing how global optimization techniques can enhance model quality and robustness. Specifically, we explore two fundamental problems: the Minimum Sum of Squares clustering problem (MSSC), including its constrained or semi-supervised variant, and Semi-Supervised Support Vector Machines (S3VM). MSSC aims to partition a set of data points into a fixed number of clusters while minimizing the sum of Euclidean distances to cluster centroids. The integration of background knowledge transforms the MSSC into a semi-supervised task. Similarly, S3VM extends traditional SVMs by leveraging both labeled and unlabeled data to learn a more reliable decision boundary. The complexity of these machine learning tasks makes the associated optimization problems NP-hard, necessitating advanced optimization techniques for efficient solutions. By employing exact algorithms and semidefinite programming tools, we achieve significant advancements in scalability and solution quality with respect to state-of-the-art methods. Through comprehensive analysis, we demonstrate the efficacy of global solutions in enhancing model performance and generalization. The outcome of this research underscores the cross-fertilization between mathematical optimization and machine learning, demonstrating the mutual benefits when combining expertise from both domains.

This research has been jointly conducted by Veronica Piccialli, Anna Russo Russo, Jan Schwiddessen, Antonio M. Sudoso, and Angelika Wiegele.

Suvrit Sra, Technische Universität München/MIT
Wednesday 15:00 to 16:00 – 210CG, Chair: Courtney Paquette

AI and Optimization through a geometric lens

Geometry arises in a myriad ways across the sciences, and quite naturally within AI and optimization too. I'd like to share with you examples where geometry helps us understand problems in machine learning and optimization. Time permitting, I'd like to also mention new results in geometric sampling, e.g., when sampling from densities supported on a manifold, understanding geometry and the impact of curvature are crucial; surprisingly, progress on geometric sampling theory helps us understand certain generalization properties of SGD for deep-learning. Another fascinating viewpoint afforded by geometry is in non-convex optimization: geometry can either help us make training algorithms more practical (e.g., in deep learning), it can reveal tractability despite non-convexity (e.g., via geodesically convex optimization), or it can simply help us understand important ideas better (e.g., eigenvectors, LLM training, etc.).

Ultimately, my hope is to offer the audience insights into geometric thinking, and to share with them some new tools that can help us make progress on modeling, algorithms, and applications. To make my discussion concrete, I will recall a few foundational results arising from our research, provide several examples, and note some open problems.

Benjamin Van Roy, Stanford University
Wednesday 16:20 to 17:20 – 210CG, Chair: Fabian Bastin

Scalable Uncertainty: AI Needs to Know What it Doesn't Know

Large models fuel impressive AI capabilities. However, they currently do not know what they do not know, which is critical to important challenges such as exploration and alignment. Knowing what is not known is essential, for example, to gathering informative feedback that can further improve AI capabilities. In this talk, I will discuss the need for scalable uncertainty and open issues in the design of architectures and training algorithms for epistemic neural networks, which could serve that need.

Andrea Lodi, Cornell Tech
Co-authors: Lara Scavuzzo, Karen Aardal, Neil Yorke Smith,
Thursday 14:00 to 15:00 – 210CG, Chair: Margarida Carvalho

Machine Learning Augmented Branch and Bound for Mixed Integer Linear Programming

Mixed Integer Linear Programming (MILP) is a pillar of mathematical optimization that offers a powerful modeling language for a wide range of applications. During the past decades, enormous algorithmic progress has been made in solving MILPs, and many commercial and academic software packages exist. Nevertheless, the availability of data, both from problem instances and from solvers, and the desire to solve new problems and larger (real-life) instances, trigger the need for continuing algorithmic development. MILP solvers use branch and bound as their main component. In recent years, there has been an explosive development in the use of machine learning algorithms for enhancing all main tasks involved in the branch-and-bound algorithm, such as primal heuristics, branching, cutting planes, node selection and solver configuration decisions. This paper presents a survey of such approaches, addressing the vision of integration of machine learning and mathematical optimization as complementary technologies, and how this integration can benefit MILP solving. In particular, we give detailed attention to machine learning algorithms that automatically optimize some metric of branch-and-bound efficiency. We also address how to represent MILPs in the context of applying learning algorithms, MILP benchmarks and software.

Jean François Cordeau, HEC Montréal
Thursday 15:00 to 16:00 – 210CG, Chair: Guy Desaulniers

Some Applications of Benders Decomposition to Location and Network Design Problems

Benders decomposition is a mathematical decomposition technique designed to solve large-scale linear and mixed-integer programs comprising two sets of variables such that a more tractable subproblem is obtained when the values of some "complicating" variables are fixed. The technique works by projecting the original problem onto the space of the complicating variables, and by using a cutting plane method where cuts are generated by solving the subproblem. Since its introduction by Jacques F. Benders in 1962, the approach has been applied successfully to a wide variety of problems arising, among others, in supply chain management, transportation, telecommunications, and energy management. Despite its success, however, it has long remained less well known than dual decomposition methods such as Lagrangian relaxation and Dantzig-Wolfe decomposition. Over the last two decades, one has witnessed a renewed interest in Benders decomposition with the introduction of several novel ideas that improve performance. The purpose of this talk is to give an overview of the main acceleration techniques by focusing on two families of problems where Benders decomposition has proven especially effective: facility location problems and network design problems. After briefly explaining the general methodology and the main practical enhancements, we will present examples of successful applications to set covering, hub location, and fixed-charge network design problems. In each case, we will focus on strategies for the efficient generation of strong cuts.

****CANCELLED ** Zuo Jun Max Shen,** The University of Hong Kong
Thursday 16:20 to 17:20 – 210CG, Chair: Antoine Legrain

Data-Driven Decision-Making: Leveraging Machine Learning for Operational Excellence

In today's ever-evolving business landscape, organizations have unprecedented access to vast amounts of data generated from their daily operations. These datasets encompass a wealth of historical observations, valuable contextual information, and customer survey data. This wealth of data presents unparalleled opportunities to address uncertain behavior and overcome significant challenges in real-world decision-making. To harness the full potential of available data, machine learning tools have emerged as indispensable techniques for driving high-quality, data-driven solutions. In this talk, we delve into different data-driven decision-making frameworks that seamlessly incorporate machine learning elements, forming their foundation theories and offering valuable and practical insights from operational perspectives. Through the integration of advanced machine learning techniques, these frameworks have proven to be effective and applicable across a wide range of practical decision-making scenarios, including inventory management, assortment optimization, and product design. Drawing from real case studies, we showcase how these data-driven frameworks can empower organizations to make well-informed decisions, enhance operational efficiency, and drive overall performance improvements.

Erick Delage, HEC Montréal

Friday 14:00 to 15:00 – 210CG, Chair: Olivier Bahn

Towards Robust and Risk-aware Contextual Optimization

Recently there has been a surge of interest in operations research (OR) and the machine learning (ML) community in combining prediction algorithms and optimization techniques to solve decision-making problems in the face of uncertainty. This gave rise to the field of contextual optimization, under which data-driven procedures are developed to prescribe actions to the decision-maker that make the best use of the most recently updated information. A large variety of models and methods have been presented in both OR and ML literature under a variety of names, including data-driven optimization, prescriptive optimization, predictive stochastic programming, policy optimization, (smart) predict/estimate-then-optimize, decision-focused learning, (task-based) end-to-end learning/forecasting/optimization, etc. The first part of the talk will identify three main frameworks for learning policies from data and sort out the literature that has pioneered this emerging field. The second part of the talk will present an overview of our groups efforts towards making contextual optimization methods more robust and risk-aware in each of these frameworks.

Guzin Bayraksan, The Ohio State University

Friday 15:00 to 16:00 – 210CG, Chair: Walter Rei

Integrating Stochastic Optimization and Machine Learning using Residuals

We present a data-driven approach that integrates machine learning prediction models with stochastic optimization by utilizing residuals from the learning models. Given a new covariate/contextual observation, the goal is to choose a decision that minimizes the expected objective function conditioned on this observation. We first review a Sample Average Approximation (SAA) approach for approximating this problem that is formed using regression residuals. We then present several extensions and discuss real-world applications. First, in the limited-data regime, we discuss Distributionally Robust Optimization (DRO) variants using Wasserstein distance, sample robust, or phi-divergence based ambiguity sets. Then, we investigate extensions under Decision-Dependent Uncertainty. We then discuss applications of this method in energy and transportation systems using real-world data.

Parallel sessions - ordered by streams

Warning: The schedule is subject to changes. Please consult the [online version](#).

OC - 210A-F OPENING CEREMONY			MONDAY 9:00-10:30
CHARLES AUDET – MIGUEL ANJOS			
CLUSTER 1: CONTINUOUS OPTIMIZATION			
Stream 1A: NONLINEAR PROGRAMMING			
1a MB11 - 516D - Advances in Large-Scale Nonlinear Optimization I			Chair: ALBERT BERAHAS
MERT GURBUZBALABAN- Robustly Stable Accelerated Gradient Methods	ALBERT BERAHAS- Advanced, Adaptive and Flexible Algorithms for Decentralized Optimization	QI WANG- Stochastic Interior-Point method for Inequality constrained optimization	JELENA DIAKONIKOLAS- Non-smooth Optimization on a Finer Scale: Bounded Local Subgradient Variation Perspective
1a MB39 - 510C - Advances in numerical methods for large scale nonlinear optimization and applications I			Chair: MARGHERITA PORCELLI
PAULO J. S. SILVA- Exploring the Spectral Proximal Gradient Method and Applications	MARGHERITA PORCELLI- Regularized methods via cubic model subspace minimization for non-convex optimization	FILIPPO ZANETTI- Scalable interior point methods for large scale discrete optimal transport problems	JACEK GONDZIO- Recent advances in preconditioners for interior point methods
1a MB62 - 516A - Development of nonlinear optimization algorithms and applications			Chair: BIN GAO
ERMIN WEI- Improved lower bounds for stochastic optimization under Markovian sampling	NICOLAS LE ROUX- Designing surrogate functions for reinforcement learning	XIN LIU- The Exactness of the L1 Penalty Function for a Class of Mathematical Programs with Generalized Complementarity Constraints	BIN GAO- Low-rank optimization on Tucker tensor varieties
1a MC55 - 510C - From First- to High-order Methodologies: Recent Advances and Trends I			Chair: ARYAN MOKHTARI
ERFAN YAZDANDOOST HAMEDANI- Efficient Methods for Solving Bilevel-structure Optimization Problem	ARYAN MOKHTARI- Non-asymptotic Global Convergence Rates of BFGS with Exact Line Search	MICHAL DEREZINSKI- Second-order Information Promotes Mini-Batch Robustness in Variance-Reduced Gradients	
1a MC902 - 516D - Nonlinear Programming 1			Chair: XINWEI LIU
RACHAEL TAPPENDEN- Connections between Gradient Descent and the Power Method	PRASHANT KHANDURI- A Discretization Approach for Low-Dimensional Non-Convex Bilevel Optimization	XINWEI LIU- D-stationary points for optimization with least constraint violation and the convergence of an SQP method	
1a TA31 - 516D - Advances in Nonconvex Optimization			Chair: SALAR FATTAHI
XUXING CHEN- From Stability to Chaos: Analyzing Gradient Descent Dynamics in Quadratic Regression	RYAN CORY WRIGHT- Optimal Low-Rank Matrix Completion: Semidefinite Relaxations and Eigenvector Disjunctions	WEIJUN XIE- On the Partial Convexification of the Low-Rank Constrained Optimization	SALAR FATTAHI- A Parametric Approach for Solving Convex Quadratic Optimization with Indicators Over Trees
1a TA40 - 510C - Advances in Large-Scale Nonlinear Optimization II			Chair: RAGHU BOLLAPRAGADA
STEPHEN WRIGHT- First-Order Algorithms for Nonconvex Min-Max Problems with Cohypomonotonicity	RAGHU BOLLAPRAGADA- Efficient Gradient Tracking Algorithmic Frameworks for Decentralized Optimization	QUOC TRAN DINH- Randomized and Stochastic Methods for Fixed-Point and Root-Finding Problems	DANIEL ROBINSON- Stochastic Proximal-Gradient Methods with Nonlinear Constraints
1a TA59 - 516A - Advances in numerical methods for large scale nonlinear optimization and applications II			Chair: ERNESTO BIRGIN
FLORENTIN GOYENS- Riemannian trust-region methods for strict saddle functions with complexity guarantees	JEREMY RIEUSSEC- A stochastic Hessian-free trust-region method for statistical learning	ERNESTO BIRGIN- Implementation and experiments of an augmented Lagrangian method with first-order information	ILARIA TROMBINI- Spectral Stochastic Gradient Method with Additional Sampling

1a TB60 - 516D - Advances in numerical methods for large scale nonlinear optimization and applications III				Chair: STEPHANIA BELLAVIA
WANPING DONG- Exploiting Negative Curvature in Conjunction with Adaptive Sampling: Theoretical Results and a Practical Algorithm	STEPHANIA BELLAVIA- Solving large-scale nonlinear least-squares with random Gauss-Newton models	GIORGIA FRANCHINI- A stochastic gradient method with variance control and variable learning rate for Deep Learning	LUKA RUTESIC- Probabilistic Trust Region method for solving Multi-objective Problems	
1a TB67 - 510C - Numerical Optimization for Machine Learning 1				Chair: FREDERIK KUNSTNER
AUDREY BAZERGHII- Relative Monte Carlo for Reinforcement Learning	WU LIN- Structured Inverse-free Matrix Adaptive Methods for Large Neural Networks	JONATHAN LAVINGTON- Target-based Surrogates for Stochastic Optimization	NICOLAS LOIZOU- Stochastic Extragradients Methods: New Convergence Guarantees for Monotone and Non-monotone Variational Inequalities	
1a TB69 - 516A - From First- to High-order Methodologies: Recent Advances and Trends II				Chair: YANG LIU
KATE WENQI ZHU- Tensor Methods for Nonconvex Optimization using Cubic-quartic regularization models	FADI HAMAD- A consistently adaptive trust-region method	SHUYAO LI- A randomized algorithm for nonconvex minimization with inexact evaluations and complexity guarantees	HANFENG ZENG- A MINRES-based Algorithm Framework for Unconstrained Nonconvex Optimization with Non-positive Curvature Detection	
1a TB136 - 511C - First-order Methods and Large-scale Optimization I				Chair: LEXIAO LAI
ZIQI QIN- Activity Identification in Non-smooth Optimization: A Set-Valued Operator Perspective	SIYU KONG- Goldstein modulus in Lipschitz Minimization	BOHAO MA- Convergence of heavy-ball SGD in the nonconvex case: a novel time window-based analysis	FENG YI LIAO- Error bounds, PL condition, and quadratic growth for weakly convex functions, and linear convergences of proximal point methods	
1a WA196 - 516D - Nonlinear stochastic optimization				Chair: MIAOLAN XIE
YAO JI- Distributed Composite Stochastic Mirror Descent for Stochastic Optimization	MIAOLAN XIE- Stochastic Adaptive Regularization Method with Cubics - A High Probability Complexity Bound	YUCHEN LOU- Design Guidelines of Noise-Tolerant Optimization Methods	PUQIAN WANG- Robustly Learning Single-Index Models via Alignment Sharpness	
1a WA251 - 510C - Algebraic Methods in Nonlinear Programming				Chair: ALI MOHAMMAD NEZHAD
THEO DIAMANDIS- Convex Network Flows	MIN TAO- A full splitting algorithm for fractional programs with structured numerators and denominators	ALI MOHAMMAD NEZHAD- Minimal structures and central paths in nonlinear optimization	SHENG DING SUN- On describing convex hull of quadratic sets via aggregations	
1a WA284 - 516A - Theory of constrained nonlinear optimization				Chair: YANGYANG XU
FARZAD YOUSEFIAN- Complexity Guarantees for Optimal Nash Equilibrium Seeking and Bilevel Variational Inequalities	UDAY SHANBHAG- Global Resolution of Chance-Constrained Optimization Problems: Minkowski Functionals and Monotone Inclusions	ZEV WOODSTOCK- Breaking the cycle: Deterministic block-iterative analysis for the Frank-Wolfe algorithm	YANGYANG XU- Damped Proximal Augmented Lagrangian Method for weakly-Convex Problems with Convex Constraints	
1a WB41 - 516D - Advances in Large-Scale Nonlinear Optimization III				Chair: BAOYU ZHOU
FABIAN BASTIN- Dynamic optimal combination of forecasting models	ANDREAS WAECHTER- On the global convergence of a decomposition algorithm for nonconvex two-level problems	JIAHAO SHI- Modified Line Search Sequential Quadratic Methods for Equality-Constrained Optimization with Unified Global and Local Convergence Guarantees	KATYA SCHEINBERG- Comparing stochastic oracles arising in robust gradient estimation	
1a WB68 - 510C - Numerical Optimization for Machine Learning 2				Chair: MARK SCHMIDT
BETTY SHEA- Why Line-Search When You Can Plane-Search?	FREDERIK KUNSTNER- Searching for optimal per-coordinate step-sizes with multidimensional backtracking	VICTOR SANCHES PORTELLA- When Online Learning Meets Stochastic Calculus	LEONARDO GALLI- Don't be so Monotone: Relaxing Stochastic Line Search in Over-Parameterized Models	

1a WB120 - 516A - Non-Convex, Non-Smooth, Stochastic, and High-Performance Computing Approaches				Chair: DONGDONG GE
JOACHIM DAHL- Recent advances in the COPT conic optimizer	QI DENG- Stochastic Structured Nonconvex Optimization in the Absence of Smoothness	TIANYI LIN- New complexity results on nonconvex nonsmooth optimization	DONGDONG GE- GPU-Centric Algorithm for Non-Linear Programming Problems	
1a WC903 - 516D - Nonlinear Programming 2				Chair: HUYEN TRANG TRAN
YASSINE KAMRI- Numerical design of optimized first-order methods	NIZAR BOUSSELMI- Analysis of Optimization Methods via non-convex Performance Estimation	HUYEN TRANG TRAN- Stochastic ISTA/FISTA Adaptive Step Search Algorithms for Convex Composite Optimization		
1a ThA19 - 516D - Theoretical advances in nonconvex optimization				Chair: HIDEFUMI KAWASAKI
CARLYLE DELIGNY- Practical session : Model and solve a non-linear optimization problem with Artelys Knitro	MOURAD BAIYOU- The k-cut problem in hypergraphs, a computational study	MARWA EL HALABI- Fairness in Submodular Maximization over a Matroid Constraint	HIDEFUMI KAWASAKI- An antipodal theorem for nonlinear optimization problems	
1a ThA42 - 510C - Advances in Large-Scale Nonlinear Optimization IV				Chair: MIKE O'NEILL
ION NECOARA- A stochastic moving ball approximation method for smooth convex constrained minimization	ANGELIA NEDICH- Randomized Algorithms for Variational Inequality Problems	SHAGUN GUPTA- Retrospective Approximation for Stochastic Constrained Problems Using Sequential Quadratic Programming	MIKE O'NEILL- A Double Step-size Stochastic SQP Method with Complexity Guarantees	
1a ThA74 - 516A - The interplay between optimization, statistics and machine learning I				Chair: LIJUN DING
LIJUN DING- Sharpness and well-conditioning of nonsmooth convex formulations in statistical signal recovery	JIALUN ZHANG- Fast Non-convex Matrix Estimation with Global Optimality Certification	JINWEN YANG- cuPDLP: A GPU implementation of restarted primal-dual hybrid gradient for linear programming	CONG MA- Top-K Ranking with a Monotone Adversary	
1a ThA135 - 511C - From First- to High-order Methodologies: Recent Advances and Trends III				Chair: ANDRE MILZAREK
CHUWEN ZHANG- HSODF: A Homogeneous Framework for Second-Order Methods	ANDRE MILZAREK- Descent properties of an Anderson accelerated gradient method with restarting	RUICHEN JIANG- Online Learning Guided Quasi-Newton Methods: Improved Global Non-asymptotic Guarantees	NACHUAN XIAO- Adam-family Methods for Nonsmooth Optimization with Convergence Guarantees	
1a ThB35 - 516D - Nonlinear Programming for Data Science I				Chair: VIVAK PATEL
VIVAK PATEL- Reliable Optimization Algorithms when Objective Function Evaluations are Expensive	PHILIPPE TOINT- Yet another fast variant of Newton's method for nonconvex optimization	KARL WELZEL- Approximate Derivatives for Tensor Methods	RAGHU PASUPATHY- Deterministic and Stochastic Frank Wolfe Recursion on Probability Spaces	
1a ThB37 - 510C - Sum-of-squares methods for continuous optimization and beyond				Chair: FRANCIS BACH
GASPARD BEUGNOT- GloptiNets: Scalable Non-Convex Optimization with Certificates	HAMZA FAWZI- Convex relaxations for Gibbs states	GIOVANNI FANTUZZI- Moment-SOS relaxations for variational problems	LUCAS SLOT- Nonconvergence of some sum-of-squares bounds for global polynomial optimization	
1a ThB93 - 516A - Nonconvex optimization: landscapes and dynamics I				Chair: CHRISTOPHER CRISCITIELLO
CHRISTOPHER CRISCITIELLO- The sensor network localization problem has benign landscape under mild rank relaxation	MATEO DIAZ- Stochastic approximation with decision-dependent distributions: asymptotic normality and optimality	CIFUENTES DIEGO- HALLaR: A new Large-Scale SDP Solver and application for Maximum Stable Set	YOSSI ARJEVANI- Symmetry and Critical Points	
1a ThB137 - 511C - First-order Methods and Large-scale Optimization II				Chair: JUNWEN QIU
PUYA LATAFAT- Adaptivity in convex optimization beyond minimization	YASONG FENG- The Anytime Convergence of Stochastic Gradient Descent with Momentum: From a Continuous-Time Perspective	JIANTING PAN- Efficient sparse probability measures recovery via Bregman gradient	NAOKI MARUMO- Universal heavy-ball method for nonconvex optimization under Hölder continuous Hessians	

1a ThC904 - 516D - Nonlinear Programming 3			Chair: IYAD WALWIL
MICHEL LAHMANN- A parallelizable ADMM approach for block-structured quadratic programs	ADEMIR ALVES RIBEIRO- Fuzzy clustering with capacity constraints: theory, algorithm, convergence analysis and numerical experiments	IYAD WALWIL- The Smoothed Duality Gap as a Stopping Criterion	
1a FA83 - 516D - Gradient-based methods: theory and advances			Chair: SHOTA TAKAHASHI
JINGYANG LI- Nonconvex Stochastic Bregman Proximal Gradient Method with Application to Deep Learning	SHOTA TAKAHASHI- Approximate Bregman Proximal Gradient Algorithm for Relatively Smooth Nonconvex Optimization	KERRY HE- Bregman Proximal Methods for Quantum Information Theoretic Problems	
1a FA94 - 510C - Nonconvex optimization: landscapes and dynamics II			Chair: ANDREEA MUSAT
JU SUN- Deep Learning with Nontrivial Constraints	DAMIEN SCIEUR- Adaptive Quasi-Newton and Anderson Acceleration with Global (Accelerated) Convergence Rates	AKSHAY RAMACHANDRAN- Non-Convexity Two Ways: new bounds for the Paulsen Problem via Geodesic PL-condition	ANDREEA MUSAT- Strict saddle avoidance for Riemannian Gradient Descent
1a FA260 - 516A - Open-source software for nonlinear programming			Chair: CHARLIE VANARET
DAVID COLE- Graph-Based Modeling and Optimization: Algorithms, Properties, and Software	CHARLIE VANARET- Uno, a next-gen solver for unifying nonlinearly constrained nonconvex optimization	MIHAI ANITESCU- Accelerating Optimal Power Flow with GPUs: SIMD Abstraction of Nonlinear Programs and Condensed-Space Interior-Point Methods	DAVID KIESSLING- A Funnel SQP and Feasibility Solver for Direct Optimal Control within the acados Software Framework
1a FB70 - 516D - Nonlinear Programming for Data Science II			Chair: VIVAK PATEL
SERGE GRATTON- High order stochastic optimization without evaluating the function.	ROBERT WEBBER- Randomized matrix decompositions for faster scientific computing	MLADEN KOLAR- Adaptive Stochastic Optimization with Constraints	EDWARD TANSLEY- Random subspace second order methods for nonconvex optimization
1a FB75 - 510C - The interplay between optimization, statistics and machine learning II			Chair: ALEX WANG
SHIQIAN MA- AdaBB: A Parameter-Free Gradient Method for Convex Optimization	FENG RUAN- Kernel Learning in Ridge Regression “Automatically” Yields Exact Low Rank Solution	YUETIAN LUO- Overparameterized Tensor Regression via Riemannian Optimization	VAIDEHI SRINIVAS- The Burer-Monteiro SDP method can fail even above the Barvinok-Pataki bound
1a FB114 - 516A - Computational advancement in nonconvex programming			Chair: ALEXIS MONTISON
JUN YA GOTOH- EM Algorithms for Optimization Problems with Polynomial Objectives	DIGVIJAY BOOB- Level constrained first-order methods for nonconvex function constrained optimization	MIJU AHN- Methods of Difference-of-Convex Programming for Statistical Learning with Structured Sparsity	ALEXIS MONTISON- GPU Implementation of Algorithm NCL
1a FB154 - 511C - First-order Methods and Large-scale Optimization III			Chair: LEXIAO LAI
L FELIPE BUENO- Inexact restoration for minimization with inexact evaluation both of the objective function and the constraints	LI JIANG- A Generalized Version of Chung’s Lemma and its Applications	PIERFRANCESCO BENEVENTANO- On the Trajectories of SGD Without Replacement	ZIJIAN LIU- Revisiting the Last-Iterate Convergence of Stochastic Gradient Methods
Stream 1B: GLOBAL OPTIMIZATION			
1b MC14 - 513B - Global Optimization			Chair: YU HONG DAI
ZHOU WEI- Stability of Error Bounds for Convex Inequalities Systems	ZHILONG DONG- International trading network optimization under extreme shocks	JOHN WARWICKER- Efficient continuous piecewise linear regression for linearising univariate non-linear functions	
1b TA324 - 512H - Global Optimization at work			Chair: SERGIY BUTENKO
ROBERT GOTTLIEB- GPU-Accelerated Deterministic Global Optimization	DMYTRO MATSYURA- Centrality of shortest paths: algorithms and complexity results	JOURDAIN LAMPERSKI- Simple Randomized Rounding for Max-Min Eigenvalue Augmentation	NIKOLAOS V. SAHINIDIS- Convexification of optimization problems involving the Euclidean norm

1b TB5 - 512H - Polynomial and Conic Optimization			Chair: IMMANUEL BOMZE
MARKUS GABL- Strengthening convex relaxations of the radial optimal power flow problem via copositive optimization	IMMANUEL BOMZE- Uncertain standard quadratic optimization under distributional assumptions: a chance-constrained epigraphic approach	QUEENIE HUANG- SDPs for Moment Optimisation with Piecewise SOS-Convex Functions	
1b WA905 - 512H - Global Optimization 1			Chair: FABIO TARDELLA
LEANDRO MAIA- An Adaptive Block Proximal ADMM for Weakly Convex, Linearly-Constrained Composite Functions	HERMANN SCHICHL- Manifolds of solutions in deterministic global optimization	XIANG LI- Difference-of-Convex Piecewise Linear Delta-Approximation and Its Application in Global Optimization	FABIO TARDELLA- Location of optima and vertex optimality
1b WB350 - 512H - Optimization on manifolds			Chair: DAVID GUTMAN
MIHAI I. FLOREA- An optimized composite gradient method for minimizing the gradient mapping	YE HE- Regularized Stein Variational Gradient Flow	DAVID GUTMAN- Tangent Subspace Descent via Discontinuous Subspace Selections on Fixed-Rank Manifolds	MAN CHO SO- Randomized Riemannian Submanifold Sub-gradient Method for Optimization over Stiefel Manifold
1b ThA24 - 512H - Recent Advances in Global Optimization			Chair: ANA ROCHA
ANA ROCHA- On a Penalty Weighted Tchebycheff Scalarizing Function for Multiobjective Optimization	ELIGIUS MT HENDRIX- Investigating polytopal global optimization branch and bound	TUSHAR RATHI- Column generation for multistage stochastic mixed-integer nonlinear programs with discrete state variables	FARUQUE HASAN- Advances in edge-concave relaxation based global optimization
1b ThB26 - 512H - Large-Scale and Structured Optimization			Chair: NAM HO NGUYEN
SYUUJI YAMADA- A procedure for listing KKT points of a quadratic reverse convex programming problem	MILTIADES PARDALOS- Continuous Approaches to Discrete Optimization Problems	QIAN ZHANG- Viral Quasispecies Analysis via Tensor Factorization	NAM HO NGUYEN- Projection-Free Methods for Solving Convex Bilevel Optimization Problems
1b FA27 - 512H - Global Optimal Search: Unleashing Maximum Potential			Chair: FREDERIC MESSINE
TANMAYA KARMAKAR- Computing the conjugate of nonconvex bivariate piecewise linear-quadratic functions.	ADRIAN GÖSS- Norm-induced Non-Convex Relaxations for Lipschitz Constraints	FREDERIC MESSINE- On the impact of reliable boundary calculations by convex relaxation in global optimization	KAMIL KAMIL KHAN- Advances in convex relaxation generation for dynamic global optimization
1b FB1 - 512H - Recent Advances in Global Optimization of Nonconvex Programs			Chair: HARSHA NAGARAJAN
ROBERT BURLACU- Discretization Approaches for Non-Convex Mixed-Integer Nonlinear Programs	ARVIND RAGHUNATHAN- Constructing Tight Quadratic Underestimators for Global Optimization	ERIN GEORGE- A Generalizable End-to-End Learning Approach for Global Optimization of Quadratically-Constrained Quadratic Programs	MOHIT TAWARMALANI- Piecewise Polyhedral Relaxations of Multilinear Optimization
Stream 1C: NONSMOOTH OPTIMIZATION			
1c MB71 - 510B - Inertial and stochastic algorithms			Chair: SEBASTIAN KASSING
SEBASTIAN KASSING- Convergence of Momentum Stochastic Gradient Descent	JUNWEN QIU- Convergence of a Normal Map-Based Prox-SGD Method for Stochastic Composite Optimization	SAM DAVANLOO TAJBAKHS- Stochastic composition optimization in the absence of Lipschitz continuous gradient	KUANGYU DING- Adam-family Methods with Decoupled Weight Decay in Deep Learning
1c MB92 - 511A - Douglas–Rachford Awesome Cones, Great Job! 1			Chair: LUIZ RAFAEL SANTOS
SCOTT BOIVIN LINDSTROM- On tight error bounds for conic optimisation	VERA ROSCHINA- On particularly awesome cones	FELIPE ATENAS- Proximal splitting methods avoid strict saddle points of weakly convex problems	LUIZ RAFAEL SANTOS- Solving large-scale basis pursuit by infeasible alternating projections
1c MC44 - 510B - Design of New Nonsmooth Optimization Algorithms			Chair: BENJAMIN GRIMMER
QIHANG LIN- Primal-Dual First-Order Methods for Convex-Concave Semi-Infinite Programs	ZHICHAO JIA- First-Order Methods for Nonsmooth Nonconvex Functional Constrained Optimization with or without Slater Points	AHMET ALACAOGLU- Variance Reduced Halpern Iteration for Optimal Finite-Sum Min-Max Optimization	

1c TA65 - 510B - Nonlinear optimization problems				Chair: WEI BIAN
WEI BIAN- Convergence and error estimation of smoothing consensus-based optimization algorithm for nonsmooth nonconvex optimization	WENJING LI- On solving a rank regularized minimization problem via equivalent factorized column-sparse regularized models	LEI YANG- A Corrected Inexact Proximal Augmented Lagrangian Method with a Relative Error Criterion for a Class of Group-quadratic Regularized Optimal Transport Problems	HAO WANG- An iteratively reweighted second-order method for nonconvex regularization	
1c TA112 - 511A - Advances in Variational Analysis and Nonsmooth Optimization I				Chair: JOHANNES ROYSET
YING CUI- Variational theory and algorithms for a class of asymptotically approachable nonconvex problems	JOHANNES ROYSET- Approximations of Rockafellians, Lagrangians, and Dual Functions	JONG SHI PANG- Heaviside Composite Optimization: Continuous meets Discrete	DANIEL MIMOUNI- Computing Wasserstein Barycenter via operator splitting: the method of averaged marginals	
1c TB63 - 510B - Nonsmooth Stochastic Optimization and Applications				Chair: HAILIN SUN
INBAR SEROUSSI- High-dimensional Limit of Stochastic Optimization Methods with Structured Data	JIE JIANG- Distributionally Robust Mean-Lower Partial Moment (LPM) Model and its Application in Portfolio Selection	HAILIN SUN- Data-driven Distributionally Robust Multiproduct Pricing Problems under Pure Characteristics Demand		
1c TB207 - 511A - Large scale continuous optimization and first-order methods and related topics				Chair: HAIHAO LU
ZIKAI XIONG- Level-set geometry and the complexity of restarted PDHG for conic LP	HAIHAO LU- A Practical and Optimal First-Order Method for Large-Scale Convex Quadratic Programming	TIANYUN TANG- A feasible method for solving general convex low-rank SDP problems	DAVID APPELEGATE- Practical Large-Scale Linear Programming using Primal-Dual Hybrid Gradient	
1c WA285 - 510B - Efficient algorithms for large-scale sparse optimization problems and applications				Chair: DEFENG SUN
YING LIN- Convex facial reduction algorithm and strong extended dual	GUOJUN ZHANG- An Accelerated Preconditioned ADMM for Solving the Optimal Transport Problem	KAIHUANG CHEN- Collective Certified Robustness against Graph Injection Attacks	DEFENG SUN- An efficient sieving based secant method for sparse optimization problems with least-squares constraints	
1c WA333 - 511A - Recent Advances in Regularized and Composite Optimization				Chair: ALEXANDER HSU
AKATSUKI NISHIOKA- Epi-convergent approximations of discontinuous generalized eigenvalue functions with application to topology optimization	MOHAMED LAGHDAF HABIBOULLAH- A Proximal Modified Quasi-Newton Method for Nonsmooth Regularized Optimization	JAMES BURKE- Feature Selection for Linear Fixed Effects Models	ALEXANDER HSU- An efficient continuation algorithm for regularized optimization through the Moreau envelope	
1c WB331 - 511A - Advances in First Order Methods				Chair: MARC TEBoulLE
JEROME BOLTE- A Bestiary of Counterexamples in Smooth Convex Optimization	SHOHAM SABACH- First Order Methods for Convex Bi-Level Optimization Problems	X.Y. HAN- Survey Descent and Its Initialization	MARC TEBoulLE- A Unified Framework for Convergence Analysis of Proximal Gradient Methods in Nonsmooth Nonconvex Minimax Problems	
1c WB906 - 510B - Nonsmooth Optimization 1				Chair: ISABELLA KÄMING
JUYOUNG JEONG- Commutation principles for nonsmooth variational problems on Euclidean Jordan algebras	POUGKAKIOTIS SPYRIDON- An efficient active-set method with applications to sparse approximations and risk minimization	HONGZHI LIAO- Douglas-Rachford is the best projection method	ISABELLA KÄMING- A new qualification condition for Lipschitzian optimization problems	
1c WC45 - 510B - Large-scale Iterative Optimization				Chair: MATEO DIAZ
ROBERT FREUND- Improving the Geometry of (Conic) Linear Optimization Problems for the Primal-Dual Hybrid Gradient Method (PDHG)	CRISTÓBAL GUZMÁN- The role of sparsity in differentially-private optimization	NAOMI GRAHAM- Density Separation with Tensor Factorization		

1c ThA30 - 510B - Nonsmooth nonconvex optimization with geometric structure				Chair: XIN LIU
DOUGLAS SOARES- GONCALVES- Away-step Frank-Wolfe algorithm for multiobjective optimization on polytopes	WEI LIU- A Near-optimal Method for Quadratic Con- strained Composite Non-convex Non-smooth Problems	QUENTIN REBJOCK- Global and local convergence guaran- tees for bounded-rank optimiza- tion with a desingularization ge- ometry		
1c ThA56 - 511A - Optimization with Projections				Chair: XIAOQI YANG
XIAOQI YANG- Projectional coderivatives with applications	JINGLAI SHEN- Sorting Func- tions for Sparse Projection with Applications to Sparse Optimiza- tion	XIAOXIA LIU- A projection al- gorithm for nonlocal low-rank tensor models with orthogonal constraints		
1c ThB54 - 511A - Projection-free first-order methods for convex optimization				Chair: JAVIER PENA
STEPHEN VAVASIS- Saddle- point Frank-Wolfe methods for Linear Programming	RENBO ZHAO- The Generalized Multiplicative Gradient Method for a Class of Convex Symmetric Cone Optimization Problem	ELIAS WIRTH- Adaptive Open- Loop Step-Sizes for Accelerated Frank-Wolfe Convergence	JAVIER PENA- Affine-invariant convergence of Frank-Wolfe algo- rithms on polytopes	
1c ThB907 - 510B - Nonsmooth Optimization 2				Chair: ZEHUI LIU
SIMONE REBEGOLDI- An ab- stract convergence framework with application to proximal Heavy-ball line-search algorithms	SAMIR ADLY- On the limit form of the Su-Boyd-Candès dynamic version of Nesterov’s accelerated gradient method when the vis- cous parameter becomes large	ZEHUI LIU- Inertial Acceler- ated Stochastic Mirror Descent for Large-Scale Generalized Ten- sor CP Decomposition	LEXIAO LAI- Global stability of first-order methods for coercive tame functions	
1c ThC38 - 510B - Nonsmooth calculus and optimization in machine learning				Chair: JALAL FADILI
ANTONIO SILVETI FALLS- Dif- ferentiating Nonsmooth Solu- tions to Parametric Monotone In- clusion Problems	SHOLOM SCHECHTMAN- The gradient’s limit of a definable family of functions is a conserva- tive set-valued field	VYACHESLAV KUNGURTSEV- Riemannian Stochastic Approx- imation for Minimizing Tame Nonsmooth Objective Functions		
1c FA25 - 510B - Some recent advances in zeroth and first-order optimization methods				Chair: TING KEI PONG
CHING PEI LEE- Structure Utilization for Reducing Dimen- sional Dependency for Nons- mooth Derivative-Free Optimiza- tion	ANDREAS THEMELIS- Adapt- able proximal algorithms: improved stepsize bounds and weaker assumptions	ANDERSEN (MAN SHUN) ANG- MGProx: A nonsmooth multigrid proximal gradient method with adaptive restriction for strongly convex optimization	YAOLIANG YU- Understand- ing Neural Network Binariza- tion with Forward and Backward Proximal Quantizers	
1c FA113 - 511A - Advances in Variational Analysis and Nonsmooth Optimization II				Chair: JOHANNES ROYSET
FRANK FISCHER- An asyn- chronous parallel bundle method	NGHIA TRAN- Geometric char- acterizations of Lipschitz stabil- ity for convex optimization prob- lems	MEISAM RAZAVIYAYN- First- Order Methods for Constrained Optimization Problems Under Distributional Uncertainties: Empirical Evaluations and Theory	KRISHNAKUMAR BALASUBRAMANIAN- Proximal Algorithms for Gaussian varia- tional inference via JKO in the Bures-Wasserstein Space	
1c FB257 - 510B - Complexity of Stationarity Concepts in Non-Smooth Optimization				Chair: MAN CHO SO
LAI TIAN- Testing Approximate Stationarity Concepts for Piece- wise Affine Functions	WENQING OUYANG- Strong Second-order Sufficient Condi- tion of Decomposable Functions	LINGLINGZHI ZHU- Nons- mooth Nonconvex-Nonconcave Minimax Optimization: Primal- Dual Balancing and Iteration Complexity Analysis		

Stream 1D: SEMI-DEFINITE PROGRAMMING			
1d MB6 - 510A - Convex Approaches for Quadratically Constrained Quadratic Programs Chair: SAMUEL BURER			
JIANZHE ZHEN- Reformulation-Perspectification Technique for Global Optimization	SAMUEL BURER- A Slightly Lifted Convex Relaxation for Nonconvex Quadratic Programming with Ball Constraints	BOSHI YANG- Semidefinite representable reformulations for two variants of the trust-region subproblem	ALEX WANG- Hidden convexity, optimization and algorithms on rotation matrices
1d MB29 - 510D - Recent advances in conic optimization - Part 1			Chair: MAKOTO YAMASHITA
KEIYU NOSAKA- Procrustean Data Collaboration for Cross-Silo Privacy-Preserving Machine Learning	AKIKO YOSHISE- Riemannian Interior Point Methods for Constrained Optimization on Manifolds	SENA SAFARINA- Projection Onto Convex Sets Approach in Solving Homogeneous Self-Dual Model of Quadratically Linear Programming Problems (Case Study: Portfolio Optimization)	CHARLES NAMCHASIRI- Dual Spectral Projected Gradient Method for Generalized Log-det Semidefinite Programming
1d MC908 - 510A - Semi-Definite Programming 1			Chair: ELLEN HIDEKI FUKUDA
NATHAN KELLER- On randomly signed sums of vectors	BENOÎT LEGAT- Reducing Sum-of-Squares constraints over a domain	ELLEN HIDEKI FUKUDA- Approximate Karush-Kuhn-Tucker 2 conditions for nonlinear second-order cone optimization problems	
1d TA138 - 510A - Moment, Polynomial and Tensor Optimization - Part 1			Chair: SUHAN ZHONG
CEDRIC JOSZ- Convergence of the momentum method for semi-algebraic functions with locally Lipschitz gradients	KE YE- Certifying nonnegative functions on finite sets via Fourier SOS	JIYOUNG CHOI- Generalized Nash equilibrium problems with Quasi-linear constraints	XINDONG TANG- Solving robotic control problems with safety constraints using moment and SOS relaxations
1d TB66 - 510A - SDP and Combinatorial Optimization			Chair: NATHAN KRISLOCK
MICHAEL SMITH- Eigenvalue cuts for the Traveling Salesman Problem	TIMOTEJ HRGA- Computational study of rounded psd inequalities for the Max-Cut and k-cluster problem	ANGELIKA WIEGELE- Edge Expansion: Exploring SDP-Based Computational Strategies	IAN SUGRUE- Facial Reduction of Semidefinite Relaxations of Binary Quadratic Problems
1d TB87 - 510D - Semidefinite Programming and Convex Algebraic Geometry			Chair: REKHA THOMAS
TAKASHI TSUCHIYA- Closing nonzero duality gaps by perturbation in SDP and its extension	EITAN LEVIN- Any-dimensional learning	SUHAN ZHONG- Polynomial Lower Approximation for Two-Stage Stochastic Optimization	MITCHELL HARRIS- Cubic interpolation to certify nonnegativity
1d TC23 - 510A - Semidefinite and Polynomial Optimization			Chair: FRANK VALLENTIN
WILLEM DE MUINCK KEIZER- Semidefinite Programming Bounds for the Kissing Number	DAVID GROSS- Non-commutative polynomial optimization with polynomial constraints for quantum causality	JEAN B. LASSERRE- On some properties and extensions of the Christoffel function	
1d TC299 - 510D - Semidefinite Programming for Geometric Processing			Chair: AMIT SINGER
ANDREW MCRAE- Benign nonconvexity of Burer-Monteiro SDP factorization for group synchronization	AMIT SINGER- Non-unique games over compact groups and orientation estimation in cryo-EM	DAVID PALMER- SDP Relaxation and Topological Defects in Meshing	
1d WA97 - 510A - Semidefinite programming, sums of squares, and quantum information Chair: HAMZA FAWZI			
SÉBASTIEN DESIGNOLLE- Bell nonlocality and Grothendieck constants via Frank-Wolfe algorithms	OJAS PAREKH- A simple approximation algorithm for Quantum Max Cut via Maximum Matching	SUJIT RAO- Analysis of ncSoS relaxations for the quantum rotor model	ADAM BENE WATTS- Relaxations and Exact Solutions to Quantum Max Cut via the Algebraic Structure of Swap Operators

1d WA208 - 510D - Low-rank semidefinite optimization, algorithms, software and applications				Chair: MICHAL KOCVARA
TIANYUN TANG- Accelerating nuclear-norm regularized low-rank matrix optimization through Burer-Monteiro decomposition	HONG MING CHIU- Well-conditioned primal-dual interior-point method for accurate low-rank semidefinite programming	MICHAL KOCVARA- Loraine news: From a low-rank to a high-precision SDP solver	ZHANWANG DENG- An Augmented Lagrangian Primal-Dual Semismooth Newton Method for Multi-Block Composite Optimization	
1d WB18 - 510A - Semidefinite Programming Performance Estimation				Chair: ETIENNE DE KLERK
BENJAMIN GRIMMER- Accelerating Gradient Descent	ETIENNE DE KLERK- Performance estimation for SDP hierarchies for polynomial optimization on the hypercube	HADI ABBASZADEHPEIVASTI- The exact worst-case convergence rate of the alternating direction method of multipliers	JASON ALTSCHULER- Acceleration by Stepsize Hedging	
1d WB222 - 510D - Algorithms and Applications of Semidefinite Programming				Chair: CIFUENTES DIEGO
ZHUORUI LI- Solving Exact and Noisy Rank-one Tensor Completion with Semidefinite Programming	MONSE GUEDES AYALA- Random Projections for Semidefinite Programming and Polynomial Optimization	ARNESH SUJANANI- HALLaR: A New Solver for Solving Huge SDPs	GEORGINA HALL- Sum of squares submodularity	
1d WC909 - 510A - Semi-Definite Programming 2				Chair: DAVID PAPP
MANUEL RUIZ- Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem	DAVID PAPP- Rational Dual Certificates for Weighted Sums-of-Squares Polynomials: Algorithms and Complexity Bounds			
1d ThA139 - 510A - Moment, Polynomial and Tensor Optimization - Part 2				Chair: XINDONG TANG
ZI YANG- Symmetric Tensor Decompositions and Gaussian Mixture Models	JOE KILEEL- Moment Estimation of Nonparametric Mixtures Through Implicit Tensor Decomposition	ELVIRA MORENO- Spectral Methods for Polynomial Optimization		
1d ThA214 - 510D - On SDP relaxations for combinatorial optimization problems				Chair: MANISH KRISHAN LAL
JAN SCHWIDDESSEN- Exploiting low-rank SDP methods for solving Max-Cut	NATHAN KRISLOCK- Using the ALPS framework to run the BiqCrunch binary quadratic solver on a computer cluster	DANIEL BROSCHE- Combinatorial Derivations and Sidorenko's Conjecture	ANNIE RAYMOND- Sums of squares and Sidorenko's conjecture	
1d ThB123 - 510D - Online optimization and learning in quantum games				Chair: ANTONIOS VARVITSIOTIS
CHRISTOF BRANDSTETTER- A Semi-Definite Programming Approach for the Single-Row Facility Layout Problem with Multiple Objectives	CONSTANTIN ICKSTADT- Semidefinite Network Games	WAYNE LIN- Learning in Quantum Potential Games	EMMANOUIL VASILEIOS VLATAKIS GKARAGKOUNIS- A Quadratic Speedup in Finding Nash Equilibria of Quantum Zero-Sum Games	
1d ThB910 - 510A - Semi-Definite Programming 3				Chair: TYLER WEAMES
CHEE KHIAN SIM- Superlinear convergence of an interior point algorithm on linear semi-definite feasibility problems	JEFERSON ZAPATA- From SDPs to polynomial equations: validating numerical solutions for semidefinite programming.	MAROUAN HANDA- Term-sparse polynomial optimization for the design of frame structures	TYLER WEAMES- Nonsmooth Newton Methods for Solving the Best Approximation Problem; with Applications to Linear Programming	
1d FA17 - 510A - Recent advances in conic optimization - Part 2				Chair: BRUNO LOURENÇO
CASEY GARNER- General Spectral Constrained Matrix Optimization	MITSUHIRO NISHIJIMA- Non-facial Exposedness of Copositive Cones over Symmetric Cones	MEHDI KARIMI- Efficient Quantum Relative Entropy Programming with SDP Preprocessing	XIAOZHOU WANG- Convergence rate analysis of a Dykstra-type projection algorithm	
1d FA89 - 510D - Recent advances in quadratic and semidefinite optimization				Chair: FRANK PERMENTER
CHENYANG YUAN- Rounding semidefinite relaxations of concave functions of quadratic forms	RICHARD ZHANG- Complexity of Chordal Conversion for Sparse Semidefinite Programs with Small Treewidth	KEVIN SHU- A Semidefinite Hierarchy for Expected Independence Numbers of Random Graphs		

1d FB36 - 510A - SDP and Exact Relaxations of Hard Problems			Chair: YONG XIA
LOUIS LORTIE- Forward-Backward Extended DMD with an Asymptotic Stability Constraint	SUNYOUNG KIM- Convex conic reformulation of geometric non-convex conic optimization problems and exact solutions of QC-QPs by SDP relaxations	SHUYANG LING- On the tightness of SDP relaxation for quadratic assignment problem	GODAI AZUMA- High-rank Solution of Sum-of-Squares Relaxations for Exact Matrix Completion
1d FB275 - 510D - Recent Advancements in the Theory and Applications of Semidefinite Programming			
Chair: HAO HU			
MANISH KRISHAN LAL- Navigating hidden convexity in non-convex projection problems	HAO HU- The Maximum Singularity Degree for Linear and Semidefinite Programming	GABOR PATAKI- Ramana's dual redux	HENRY WOLKOWICZ- A Peaceman-Rachford Splitting Method for the Protein Side-Chain Positioning Problem and other Hard Problems
Stream 1E: VARIATIONAL ANALYSIS, VARIATIONAL INEQUALITIES AND COMPLEMENTARITY			
1e MB312 - 513D - Convex analysis and beyond			Chair: ADRIAN LEWIS
SHUBHANKAR BHATT- Polar convexity in finite dimensional Euclidean spaces	MICHAEL ORLITZKY- The uniqueness of Lyapunov rank among symmetric cones	SEDI BARTZ- Constructions of c-splitting potentials for multi-marginal monotone sets	ADRIAN LEWIS- Exploring averages of finite sets in nonpositively curved metric spaces
1e TA313 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 1			
Chair: RADU IOAN BOT			
HEINZ BAUSCHKE- On a result by Baillon, Bruck, and Reich	SHAMBHAVI SINGH- On the Bredies-Chenchene-Lorenz-Naldi algorithm	RUSSELL LUKE- The proximal point algorithm without monotonicity	RADU IOAN BOT- Fast continuous time methods for monotone equations
1e TB311 - 514A - New insights into first- and second-order variational analysis			Chair: TIM HOHEISEL
STEPHEN SIMONS- Quasidensity and certain classes of multifunctions	TING KEI PONG- Kurdyka-Lojasiewicz exponent for a class of Hadamard-difference-parameterized models	HELMUT GFRERER- On second-order variational analysis of prox-regular functions	TIM HOHEISEL- Stability of nonsmooth optimization problems
1e TB314 - 514B - Algorithms for optimization, monotone inclusions and variational inequalities 2			
Chair: ZI XU			
JIAYUAN WU- A projected semismooth Newton method for a class of nonconvex composite programs with strong prox-regularity	VIKTOR PAVLOVIC- Accelerated gradient descent: A guaranteed bound for a heuristic restart strategy	EWA BEDNARCZUK- Forward-Backward algorithms for weakly convex problems	ZI XU- Zeroth order and first order primal dual alternating proximal gradient algorithms for nonsmooth nonconvex minimax problems with coupled linear constraints
1e TC911 - 514A - Variational Analysis, Variational Inequalities and Complementarity 1			Chair: JEIN SHAN CHEN
MONIKA SYGA- Zero duality gap for optimization problems in spaces without linear structure	EDOUARD OYALLON- Accelerating Communications in Decentralized Learning via Randomization	JEIN SHAN CHEN- On some new NCP functions	
1e WA338 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 3			
Chair: WALAA MOURSI			
BRUNO LOURENÇO- Exotic error bounds, Karamata theory and convergence rate analysis	JONATHAN ECKSTEIN- DEFBAL – a connection between the ADMM and forward-backward methods	BETHANY CALDWELL- The Douglas-Rachford algorithm for inconsistent problems	WALAA MOURSI- The Chambolle-Pock algorithm revisited: splitting operator and its range with applications
1e WB227 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 4			
Chair: JUAN PABLO CONTRERAS			
AFROOZ JALILZADEH- Single-loop stochastic variance reduced method for variational inequalities	SIQI QU- Tikhonov regularized exterior penalty dynamics for constrained variational inequalities	ZHUOYU XIAO- Synchronous and asynchronous gradient-response schemes for computing quasi-Nash equilibria under uncertainty	JUAN PABLO CONTRERAS- Stochastic Halpern iteration in non-Euclidean spaces and applications to synchronous Q-learning

1e WB329 - 515C - Conic optimization				Chair: JAKUB MARECEK
HECTOR RAMIREZ- Extensions of constant rank qualification constraint condition to nonlinear conic programming	ALES WODECKI- Learning quantum Hamiltonians at any temperature in polynomial time with Chebyshev and bit complexity	ANTONIO BELLON- Time-varying semidefinite programming: path-following a Burer-Monteiro factorization	JAKUB MARECEK- Time-varying semidefinite programming: geometry of the trajectory of solutions	
1e ThA354 - 514C - Recent advances in Variational Inequalities and Complementarity				Chair: JEAN PIERRE DUSSAULT
MATEO DIAZ- The radius of statistical efficiency	EBRAHIM SARABI- The Role of Subgradients in Second-Order Variational Analysis	JEAN PIERRE DUSSAULT- Polyhedral Newton-min algorithms for complementarity problems	BAPTISTE PLAQUEVENT- Last-Iterate Levenberg-Marquardt globalization of the polyhedral Newton-min algorithm for complementarity problems	
1e ThA912 - 514A - Variational Analysis, Variational Inequalities and Complementarity 2				Chair: TATJANA CHAVDAROVA
KRZYSZTOF RUTKOWSKI- On projections onto Parametric Polyhedral Sets	LINGLING XU- Douglas-Rachford splitting algorithm with line search	FELIX SCHNEPPE- Primal-Dual Methods with Adjoint Mismatch - Existence of Stationary Points, Convergence and Error Bounds	TATJANA CHAVDAROVA- Last-Iterate Convergence of Saddle-Point Optimizers via High-Resolution Differential Equations	
1e ThB226 - 512E - Nonconvexity, stochasticity and hierarchy in optimization problems				Chair: MATHIAS STAUDIGL
ALEXANDER BODARD- The inexact power augmented Lagrangian method for constrained nonconvex optimization	LUKE MARRINAN- A quasi-Newton method for nonsmooth, nonconvex, stochastic optimization	YUYANG QIU- Zeroth-order federated methods for stochastic MPECs and nondifferentiable nonconvex hierarchical optimization	MATHIAS STAUDIGL- Hessian barrier algorithms for non convex conic optimization	
1e ThB913 - 514A - Variational Analysis, Variational Inequalities and Complementarity 3				Chair: ABDERRAHIM JOURANI
XIAOXIAO MA- Calm local optimality for nonconvex-nonconcave minimax problems	ABDERRAHIM JOURANI- On desingularizing functions in convex programming			
1e FA103 - 514A - Recent advances in bilevel optimization				Chair: JAN HAROLD ALCANTARA
SHANGZHI ZENG- Moreau envelope based reformulation and algorithm for bilevel programs	JEONGYEOL KWON- On first-order methods for stochastic bilevel optimization	ALAIN ZEMKOHO- Relaxation methods for pessimistic bilevel optimization	JAN HAROLD ALCANTARA- Theoretical smoothing frameworks for general nonsmooth bilevel problems	
1e FB336 - 514A - Nonsmooth and hierarchical optimization in machine learning				Chair: RAHUL PARHI
ZHAOSONG LU- First-order methods for bilevel optimization	DAVID BENFIELD- Classification under strategic adversary manipulation using pessimistic bilevel optimisation	RAHUL PARHI- A Banach-space view of neural network training		
Stream 1F: RANDOM METHODS FOR CONTINUOUS OPTIMIZATION				
1f TA130 - 512E - Fast randomized algorithms for distributed optimization				Chair: LAURENT CONDAT
CORRADO COPPOLA- CMA Light: a minibatch algorithm for large-scale non-convex finite sum optimization	JIAXIN XIE- On adaptive stochastic heavy ball momentum for solving linear systems	MING YAN- Optimal gradient tracking for decentralized optimization	LAURENT CONDAT- Communication-efficient distributed optimization algorithms	

1f TB131 - 512C - Beyond Standard Assumptions in Optimization: Structured Non-Convexity, Non-Monotonicity, and Non-Smoothness				Chair: EDUARD GORBUNOV
YI ZHOU- Generalized-Smooth Nonconvex Optimization is As Efficient As Smooth Nonconvex Optimization	ILYAS FATKHULLIN- Stochastic Optimization under Hidden Convexity	BRECHT EVENS- Convergence of Douglas–Rachford Splitting and Primal-Dual Hybrid Gradient for Non-Monotone, Non-Smooth Problems	CONG FANG- On the Lower Bound of Minimizing Polyak-Lojasiewicz functions	
1f WA334 - 512C - Random subspace methods for continuous optimization				Chair: PIERRE LOUIS POIRION
MICHAEL ULBRICH- A Semismooth Newton Stochastic Proximal Point Algorithm with Variance Reduction	AKIKO TAKEDA- Applying Random projection techniques to large-scale nonconvex optimization problems	BENEDETTO MANCA- Random projections for quadratically constrained quadratic programming	PIERRE LOUIS POIRION- Random subspace Newton method for unconstrained non-convex optimization	
1f WB49 - 512C - Splitting and Randomized Methods 1				Chair: DEREN HAN
MEHIDDIN AL BAALI- Improved L-BFGS Methods for Large-Scale Unconstrained Optimization	LIPING WANG- A Decentralized Primal-Dual Method with Quasi-Newton Tracking	ZIYAN LUO- Group SLOPE Penalized Low-Rank Tensor Regression	DEREN HAN- Revisit Iterative Complexity of Some Projection Methods for Monotone Linear Variational Inequalities	
1f ThA128 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization I				Chair: NICOLAS LOIZOU
MARK SCHMIDT- Global-Local Smoothness: Line Search can really help! Really!	RYAN D’ORAZIO- What is the right extension of Polyak’s step-size to mirror descent?	COURTNEY PAQUETTE- Scaling Law: Finding Compute Optimal Curves on a Simple Model with SGD	XIAOYU WANG- Generalized Polyak Step Size for First Order Optimization with Momentum	
1f ThB50 - 512C - Splitting and Randomized Methods 2				Chair: YANNAN CHEN
ALOIS PICHLER- Uniform Function Estimators in Reproducing Kernel Hilbert Spaces, with applications in stochastic optimization	XUDONG LI- Data-Driven Minimax Optimization with Expectation Constraints	YANNAN CHEN- A Scalable Optimization Approach for the Multilinear System Arising from Scattered Data Interpolation		
1f ThC129 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization II				Chair: FRANCESCO ORABONA
OLIVER HINDER- The Price of Adaptivity in Stochastic Convex Optimization	LOUIS SHARROCK- Coin Sampling: Learning-Rate-Free Optimisation on the Space of Probability Measures	EVGENII CHZHEN- Parameter-free projected gradient descent with application to contextual bandits with knapsacks		
1f FA51 - 512C - Splitting and Randomized Methods 3				Chair: CHUNFENG CUI
CAIRONG CHEN- Optimal and approximate optimal parameters of the SOR-like iteration method for solving absolute value equations	XINGJU CAI- A regularized splitting method for sums of multi-block maximal monotone operators and its relationship with symmetric ADMM	CHUNFENG CUI- Extended alternating structure-adapted proximal gradient algorithm for multiblock nonconvex optimization		
1f FB52 - 512C - Splitting and Randomized Methods 4				Chair: JINGYA CHANG
ISABEL JACOB- A Recursive Multilevel Algorithm for Deep Learning	TING LI- Improved variance reduction extragradient method with line search for stochastic variational inequalities	JINGYA CHANG- Differentially Private Federated Learning via Inexact ADMM with Multiple Local Updates		
Stream 1G: DERIVATIVE-FREE AND SIMULATION-BASED OPTIMIZATION				
1g MB302 - 515A - Challenging DFO and SBO Applications 1				Chair: YVES LUCET
MORGANE MENZ- Optimization under hidden constraints and its application to the robust design of wind turbines	FILIPPO MARINI- Black-box optimization for the design of a jet plate for impingement cooling	LUCA BOISNEAULT- Blackbox optimization for origami-inspired multistable structures	YVES LUCET- Convex optimization of vertical alignment of roads	

1g TA306 - 515C - Large-scale and Parallel DFO				Chair: YOUSSEF DIOUANE
PIERRE YVES BOUCHET- A derivative-free approach to partitioned optimization	HAITIAN LI- Blockwise direct search method	JEFFREY LARSON- Using concurrent objective evaluations in derivative-free optimization	YOUSSEF DIOUANE- A direct-search method for decentralized derivative-free optimization	
1g TB322 - 515C - Recent Progresses in DFO 1				Chair: GIAMPAOLO LIUZZI
GABRIEL JARRY BOLDUC- The cosine measure relative to a subspace	ZAIKUN ZHANG- A (forgotten) subspace framework for large-scale derivative-free optimization	SAMUEL MENDOZA- Parallel versions of the mesh adaptive direct search algorithm	GIAMPAOLO LIUZZI- Worst case complexity bounds for linesearch-type derivative-free algorithms	
1g TC305 - 515C - Multiobjective DFO				Chair: LUDOVIC SALOMON
EVERTON SILVA- An inexact restoration direct multisearch filter approach to constrained optimization	LUDOVIC SALOMON- A comparison of search strategies for multiobjective direct search algorithm	DIMO BROCKHOFF- Gradient-based and derivative-free multi-objective optimization via iterative single-objective optimization: MO-BFGS and MO-SLSQP		
1g WA309 - 515C - Multifidelity and Bilevel DFO				Chair: FRANCESCO RINALDI
YUNSOO HA- Multifidelity Monte Carlo estimates for accelerating stochastic optimization with trust regions	XAVIER LEBEUF- Hierarchically constrained multi-fidelity blackbox optimization	SAEED GHADIMI- Zeroth-order bilevel optimization	FRANCESCO RINALDI- Inexact direct-search methods for bilevel optimization problems	
1g WC323 - 515C - Recent Progresses in DFO 2				Chair: ANA LUISA CUSTODIO
CHRISTOPHE TRIBES- Improving MADS search step with a quadratic programming solver	ANDREA BRILLI- Advancing derivative-free optimization: exploring logarithmic barrier and exterior penalty in direct search	ANA LUISA CUSTODIO- A hybrid direct search and projected simplex gradient method for convex constrained optimization		
1g ThA325 - 515C - DFO based on Approximations to Derivatives 1				Chair: STEFAN M. WILD
SARA SHASHAANI- Is building first-order simulation oracles really worth it?	KILLIAN WOOD- Adaptive stochastic subspace descent	JIAQI WANG- Self-Correcting DFO Algorithm: An Attempt at a Convergence Rate	STEFAN M. WILD- Derivative-free stochastic optimization via adaptive sampling strategies	
1g ThB300 - 515C - Stochastic DFO 1				Chair: MATT MENICKELLY
ARMAND GISSLER- Convergence analysis of CMA-ES	YIWEN CHEN- Q-fully quadratic modeling and its application in a random subspace derivative-free method	PENGCHENG XIE- A new two-dimensional model-based subspace method for large-scale unconstrained derivative-free optimization: 2D-MoSub	MATT MENICKELLY- Noise-aware derivative-free optimization	
1g ThC303 - 515C - Challenging DFO and SBO Applications 2				Chair: EDWARD HALLÉ HANNAN
SYLVERIO POOL MARQUEZ- Management of categorical variables in a mixed-variable (surrogate-based) evolutionary algorithm (Part I)	CHARLOTTE BEAUTHIER- Management of categorical variables in a mixed-variable (surrogate-based) evolutionary algorithm (Part II)	EDWARD HALLÉ HANNAN- A graph-structured distance for heterogeneous datasets with meta variables		
1g FA326 - 515C - DFO based on Approximations to Derivatives 2				Chair: ANNE AUGER
YANJUN LIU- Black-Box Optimization Algorithms for Problems with Convex Regularizers	TIANYU WANG- Zeroth-order low-rank Hessian estimation via matrix recovery	GEOVANI GRAPIGLIA- Derivative-free implementation of the regularized Newton method with lazy approximated Hessians	ANNE AUGER- On the connection between affine-invariance, convergence and learning second order information	
1g FB327 - 515C - Stochastic DFO 2				Chair: WARREN HARE
KWASSI JOSEPH DZAHINI- Stochastic derivative-free optimization algorithms using random subspace strategies	CUNXIN HUANG- Non-convergence analysis of probabilistic direct search	CLÉMENT ROYER- Full-low evaluation methods for bound and linearly constrained derivative-free optimization	WARREN HARE- Expected decrease for derivative-free algorithms using random subspaces	

Stream 1H: OPTIMAL CONTROL, PDE CONSTRAINED OPTIMIZATION, AND MULTI-LEVEL METHODS

1h MB88 - 515C - Nonsmooth PDE Constrained Optimization 1 Chair: HARBIR ANTIL

HARBIR ANTIL- PDE Constrained Optimization and Digital Twins	ANTON SCHIELA- Semismoothness in Infinite Dimensional Non-Smooth Optimization	DANIEL WACHSMUTH- A numerical solution approach for non-smooth optimal control problems based on the Pontryagin maximum principle	TUOMO VALKONEN- Interweaved first-order methods for PDE-constrained and bilevel optimisation
--	---	---	--

1h TA90 - 512B - Nonsmooth PDE Constrained Optimization 2 Chair: ANNIKA SCHIEMANN

MICHAEL HINTERMÜLLER- A hybrid physics-informed neural network based multiscale solver as a PDE constrained optimization problem	PAUL MANNS- Decomposition in Integer Optimal Control with Total Variation Regularization	MARVIN SEVERITT- Integer linear programming for two-dimensional trust-region subproblems arising in integer optimal control	ANNIKA SCHIEMANN- Integer Optimal Control Problems with TV regularization on Multi-dimensional Domains
--	--	---	--

1h TB91 - 512B - Nonsmooth PDE Constrained Optimization 3 Chair: HENRIK WYSCHKA

DREW KOURI- An Inexact Trust-Region Algorithm for Nonsmooth Risk-Averse Optimization	NICOLAI SIMON- Control in the Coefficients of an Obstacle Problem	ROHIT KHANDELWAL- A Discontinuous Galerkin Method for Optimal Control of the Obstacle Problem	HENRIK WYSCHKA- Limiting Descent Directions in p-Harmonic Shape Optimization
--	---	---	--

1h TC320 - 512B - Analysis of Algorithms for PDE Constrained Optimization 1 Chair: STEFAN ULBRICH

STEFAN ULBRICH- Snapshot generation for POD models of semilinear parabolic optimal control problems based on a simplified Newton method	JESPER SCHRÖDER- On the well-posedness of the Mortensen observer for a defocusing cubic wave equation	ROBERT J BARALDI- Adaptive Randomized Sketching for Nonsmooth Dynamic Optimization	
---	---	--	--

1h WA914 - 512B - Optimal Control, PDE Constrained Optimization, and Multi-level Methods 1 Chair: SEPEHR MOALEMI

JOHANNES HAUBNER- Image Registration Using Optimal Control of a Linear Hyperbolic Transport Equation	JAD WEHBEH- Robust Optimal Control Through Scenario-Based Solutions to Semi-Infinite Programs With Existence Constraints	DOMINIK H. CEBULLA- Towards robust optimization of chromatographic separation processes with flow reversal	SEPEHR MOALEMI- Input-Output Stability of First-Order Optimization Algorithms: A Passivity Approach
--	--	--	---

1h WB321 - 512B - Challenging PDE Constrained Applications Chair: ANTON BONDAREV

AKWUM ONWUNTA- Low-rank tensor approximation for risk averse optimization	JANNIK BREITKOPF- Optimal Control of the Generalized Riemann Problem for Hyperbolic Systems of Conservation Laws	MARCEL MAXIMILIAN STEINHARDT- A Variational Calculus for Optimal Control of Networks of Scalar Conservation or Balance Laws	LUKAS KANNENGIESSER- Parameter Identification in Boundary Value Problems for Fiber Spinning
---	--	---	---

1h ThA361 - 512B - Analysis of Algorithms for PDE Constrained Optimization 2 Chair: ARSEN HNATIUK

THOMAS LYNN- Sensor path planning for optimal experimental design	ANDREI DRAGANESCU- Low-rank Gradient Flow – a First-Order Algorithm for Non-convex Optimization	CHRISTOPH PLATE- Optimal Experimental Design for Universal Differential Equations	ARSEN HNATIUK- Lazyfied Generalized Conditional Gradient
---	---	---	--

1h ThB318 - 512B - Shape Optimization Chair: THEODOR KOMANN

ZAKARIA HOUTA- Level Set-Based Topology Optimization Of 3D Magnetic Circuits With Mechanical Constraints	KATHRIN WELKER- PDE constrained shape optimization in the space of piecewise-smooth shapes	ROLAND HERZOG- Mesh Denoising and Inpainting using the Total Variation of the Normal and a Shape Newton Approach	THEODOR KOMANN- Robust Optimization of Electric Machines: A Gradient-Based Approach for Parameter and Shape with Isogeometric Analysis
--	--	--	--

1h ThC319 - 512B - Multi-Level and Parallel Methods for PDE Constrained Optimization Chair: WING HONG FELIX KWOK

RADOSLAV VUCHKOV- Multigrid-in-time methods for nonlinear optimization of dynamical systems	ED BUELER- A full approximation scheme multilevel method for solving nonlinear variational inequalities	WING HONG FELIX KWOK- Pipeline waveform relaxation for optimal control	
---	---	--	--

CLUSTER 2: DISCRETE OPTIMIZATION

Stream 2A: MIXED INTEGER LINEAR PROGRAMMING

2a MB157 - 511F - Vehicle Routing Problems under Uncertainty Chair: PEDRO MUNARI

MATHEUS JUN OTA- Approximating the recourse cost function of the vehicle routing problem with stochastic demands.	ALBERT SCHROTENBOER- Distributionally Robust Cyclic Inventory Routing	RAFAEL AJUDARTE DE CAMPOS- Polynomial-time separation algorithms based on closed-form equations for the RVRP under travel time uncertainty	PEDRO MUNARI- Pickup and Delivery Vehicle Routing Problems under Uncertainty via Robust Optimization
---	---	--	--

2a TA188 - 513B - Network-flow based MIP formulations Chair: FRANÇOIS CLAUTIAUX

MIKE HEWITT- A Branch-and-Price Approach for a Consolidation-based Formulation of the Scheduled Service Network Design Problem	SEAN LO- Exact Methods For Electric Routing Operations: A Subpath-Based Pricing Problem Decomposition	FRANÇOIS CLAUTIAUX- A network-flow formulation for a scheduling problem in the car industry	LUIS MARQUES- Dynamic programming based arc-flow formulations for cutting and packing problems
--	---	---	--

2a TB915 - 513B - Mixed Integer Linear Programming 1 Chair: QUOC TUNG LE

NILS HASSAN QUTTINEH- A column generation approach for a challenging real-world routing problem	BÁRBARA RODRIGUES- Complexity of Determining Unboundness in Linear Bilevel Optimization	QUOC TUNG LE- On the geometric and computational complexity of polynomial bilevel optimization	
---	---	--	--

2a WA231 - 511E - Bilevel Optimization with Discrete Decisions (I) Chair: MARTINA CERULLI

MARTIN LABBE- Playing Stackelberg Security Games in perfect and imperfect formulations	YASMINE BECK- A Toll-Setting Problem with Gamma-Robust Wardrop Equilibrium Conditions	MARTIN SCHMIDT- BOBILib - A Bilevel Optimization Benchmark Instance Library	FEDERICO BATTISTA- A Branch-and-Cut Algorithm for Mixed-Integer Bilevel Linear Optimization Based on Improving Directions
--	---	---	---

2a WA916 - 513B - Mixed Integer Linear Programming 2 Chair: VANESSA ANN BEDDOE SAWKMIE

YOUNGJOO ROH- Lifting of the cover inequalities for the robust knapsack problem	FLORIAN ROESEL- A novel Pareto-optimal cut selection strategy for Benders Decomposition	ENRICO MALAGUTI- Bulk Optimization approaches to Graph Problems	VANESSA ANN BEDDOE SAWKMIE- Route 'em and Count 'em : A Two-Stage Stochastic Programming Model in Undersea Warfare
---	---	---	--

2a WB236 - 513D - Bilevel Optimization with Discrete Decisions (II) Chair: JOHANNES THÜRAUF

LEONARDO LOZANO- Network Relaxations for Discrete Bilevel Optimization under Linear Interactions	ANDREAS HORLÄNDER- Using Disjunctive Cuts in a Branch-and-Cut Method to Solve Convex Integer Nonlinear Bilevel Problems	NOAH WENINGER- Interdiction of minimum spanning trees and other weighted matroids	MARGARIDA CARVALHO- Combinatorial pricing problems: methods using embedded dynamic programming models
--	---	---	---

2a WB917 - 513B - Mixed Integer Linear Programming 3 Chair: DAVID WARME

JUAN FRANCISCO CORRECHER- A block decomposition algorithm for an extension of the vertex cover problem	JUANJO PEIRÓ- Preliminary results on an extension of the vertex cover problem	ERIK JANSEN- Analyzing the Sensitivity of Integer Linear Programs via Optimization Oracles	DAVID WARME- 1,000,000 Point Steiner Tree Solved to Optimality
--	---	--	--

2a WC160 - 513B - Network Design Chair: NAVNEET VIDYARTHI

PABLO FELIPE OTEÍZA- ML-based Algorithm for Solving the Transmission Network Expansion Planning Problem	JAEHEE JEONG- Optimizing Electric Vehicle Fleet Operations for Sustainable Last-mile Delivery Under Uncertainty	NAVNEET VIDYARTHI- Discrete Location Problem with Probabilistic Service Level Constraints	
---	---	---	--

2a ThA237 - 513B - Recent Advances in Combinatorial Optimization Chair: AHMAD ABDI

SAUMYA SINHA- Lagrangian relaxation and duality for multiobjective integer programming	KIRILL KUKHARENKO- On the Number of Degenerate Simplex Pivots	MEIKE NEUWOHNER- A better-than-k/2-approximation for weighted k-Set Packing	
--	---	---	--

2a ThB161 - 513D - Advances in ILP: Branch and cut				Chair: ALINSON SANTOS XAVIER
MARC PFETSCH- Super-Polynomial Lower Bounds for Branch-and-Bound with General Disjunctions via Interpolation	JOSEPH PAAT- Compressing Branch-and-Bound Trees	ANTONIA CHMIELA- Polyhedrality made easy: Using general cut operators to determine when cut closures are polyhedral	ALINSON SANTOS XAVIER- Learning to Generate Cutting Planes with MIPLearn	
2a ThB918 - 513B - Mixed Integer Linear Programming 4				Chair: VENKAT AKHIL ANKEM
RENAUD CHICOISNE- Strong Cutting Planes for the Feature Selection Problem	LOUISA SCHROEDER- The undirected circular facility layout problem	FERNANDO DIAS- Inexact Flow Decomposition in Quasispecies Network via Integer Linear Programming	VENKAT AKHIL ANKEM- Production scheduling of open pit mines using new cutting planes	
2a ThC180 - 513B - Advances in Mixed Integer Programming Theory				Chair: DANIEL DADUSH
THOMAS ROTHVOSS- The Subspace Flatness Conjecture and Faster Integer Programming	STEFAN WELTGE- Integer programs with nearly totally unimodular matrices: proximity	DANIEL DADUSH- Integrality Gaps for Random Integer Programs via Discrepancy		
2a FA228 - 513B - Integer programming: structure and algorithms				Chair: MANUEL APRILE
ALEXANDRA ANNA LASSOTA- Parameterized Algorithms for Block-Structured Integer Programs with Large Entries	HYUNWOO LEE- Cooperative Integer Programming Games	LUCA FERRARINI- Total Matching and Subdeterminants	STEFAN KOBER- Integer programs with nearly totally unimodular matrices: the cographic case	
2a FA262 - 513D - Aspects of Symmetry and Fairness in Integer Programming and Combinatorial Optimization				Chair: CHRISTOPHER HOJNY
ADRIAN VETTA- FPT Algorithms using Minimal Parameters for a Generalized Version of Maximin Shares	ANNIKA JAEGER- Handling symmetries in stable set problems and perfect mixed graphs	STEN WESSEL- Fair Schedules for Single Round Robin Tournaments with Ranked Participants	CHRISTOPHER HOJNY- Detecting and Handling Reflection Symmetries in MINLP	
2a FB151 - 513D - Recent Developments in MIP Theory and Applications				Chair: BURAK KOCUK
BURAK KOCUK- Proximity results in conic mixed-integer programming	GUSTAVO ANGULO- On formulations of the close-enough TSP	DAVID BERNAL NEIRA- A Convexification-Based Outer-Approximation Method for Convex and Nonconvex MINLP	DIEGO ALEJANDRO MORAN RAMIREZ- M-Decomposable Sets in Integer Programming	
2a FB919 - 513B - Mixed Integer Linear Programming 5				Chair: JAKOB NORDSTRÖM
NATHAN BENEDETTO- Generalized Cuts and Grothendieck Covers: Algorithmic Performance	JAKOB NORDSTRÖM- Combinatorial solving with provably correct results			
Stream 2B: MIXED INTEGER NONLINEAR PROGRAMMING				
2b MB12 - 514C - Binary Quadratic Programming				Chair: SVEN MALLACH
LUCAS LÉTOCART- Matrix generation for binary quadratic programming	BORZOU ROSTAMI- Exact and Approximate Solution Methods for Convex Binary Quadratic Optimization	LUCAS WADDELL- Binary Quadratic Optimization: A Polyhedral Characterization of Linearizable Instances	SVEN MALLACH- Inductive Linearization for Binary Quadratic Programs	
2b MC286 - 513F - MINLP Applications				Chair: AKSHAY GUPTA
OLHA SILINA- The lattice of di-joins	JUSTO PUERTO- Mixed integer linear programming formulations and column generation algorithms for the Minimum Normalized Cuts problem	DAVID REY- Scaling-up exact methods for solving the discrete bilevel network design problem under user equilibrium		
2b TA152 - 513F - Solving ACOF problems				Chair: DANIEL BIENSTOCK
HASSAN HIJAZI- AC-Unit Commitment: Capturing Losses	ANDY SUN- New Advances in Solving Large-Scale Unit-Commitment-AC-OPF Problem	PASCAL VAN HENTENRYCK- Fusing AI and Optimization for Security-Constrained Economic Dispatch	MATIAS VILLAGRA- Accurate and Warm-Startable Linear Cutting-Plane Relaxations for ACOF	

2b TB195 - 513F - Optimization over integers, cones and polynomials				Chair: LUZE XU
SHIXUAN ZHANG- Spurious local minima in nonconvex sum-of-square optimization	KURT ANSTREICHER- Solving Nonconvex Optimization Problems using Outer Approximations of the Set-Copositive Cone	LUZE XU- Integer Points in Arbitrary Convex Cones: The Case of the PSD and SOC Cones	SHAONING HAN- On Polynomial-Time Solvability of Combinatorial Markov Random Fields	
2b TC920 - 513F - Mixed Integer Nonlinear Programming 1				Chair: ANDREA GHEZZI
EHSAN YADEGARI- Periodic Capacity Development in Nonlinear Closed-loop Logistics Planning: an advanced Continuous Genetic Algorithm	HENRI LEFEBVRE- A First-Relax-then-Reformulate Approach for Bilevel Problems with Nonconvex Mixed-Integer Nonlinear Lower Levels	ANDREA GHEZZI- A Sequential Benders-based Mixed-Integer Quadratic Programming Algorithm		
2b WA8 - 513F - Mixed Integer Quadratic/Conic Optimization				Chair: ATAMTURK ALPER
LINCHUAN WEO- An Outer Approximation Method for Solving Mixed-Integer Convex Quadratic Programs with Indicators	FATMA KILINC KARZAN- New Perspectives on Deriving Compact Extended Formulations for Optimization Problems with Indicator Variables	ANDRES GOMEZ- Convexification and solution of mixed-integer quadratic programs using decision diagrams	SANTANU DEY- Sensitivity Analysis for Mixed Binary Quadratic Programs	
2b WB7 - 513F - Algorithms for Global Mixed Integer Nonlinear Optimization				Chair: PIETRO BELOTTI
IMRE POLIK- Solving MINLPs to Global optimality with the FICO Xpress Solver	JULIO GONZÁLEZ DÍAZ- R-POSa: A free global solver for mixed integer polynomial optimization problems	ROBERT LUCE- Techniques and advances for solving MINLPs with Gurobi	PIETRO BELOTTI- Spatial branching for convex MIQO problems	
2b WC346 - 513F - MINLP meets Machine Learning				Chair: EMILIANO TRAVERSI
JOSÉ NETO- Outlier Detection in Regression: Conic Quadratic Formulations	SHIQIANG ZHANG- Optimization over trained graph neural networks: mixed-integer formulations, symmetry breaking, and optimal molecular design	EMILIANO TRAVERSI- ML-based Clique Merging Algorithms to Solve Semidefinite Relaxations of Optimal Power Flow Problems		
2b ThA77 - 513F - Polynomial Optimization Algorithms and Applications				Chair: BISSAN GHADDAR
ÁNGEL MANUEL GONZÁLEZ RUEDA- Parameter estimation in ODEs: assessing the potential of local and global solvers	IGNACIO GÓMEZ CASARES- Domain reduction techniques in polynomial optimization	ROHIT KANNAN- Learning to Accelerate the Global Optimization of Quadratically-Constrained Quadratic Programs	BISSAN GHADDAR- Strong and fast SDP bounds for water network problems	
2b ThB34 - 513F - Advances on Multi-Objective and Separable Mixed Integer Nonlinear Programming				Chair: MARIANNA DE SANTIS
PHILIP DE CASTRO- Pareto LEAP: An Algorithm for Biobjective Mixed Integer Nonlinear Programming	MORITZ LINK- Pseudo enclosures for multi-objective mixed-integer nonconvex optimization	DANIELE PATRIA- Detecting the integer efficient assignments for multiobjective mixed-integer convex quadratic problems	LUCA MENCARELLI- MWU 2.0 with approximation guarantee for non-convex separable problems	
2b ThC159 - 513F - Discrete structures in non-convex problems				Chair: SONIA CAFIERI
KIM YU- On constrained mixed-integer DR-submodular minimization	GONZALO MUÑOZ- Exploiting the polyhedral geometry of stochastic linear bilevel programming	SONIA CAFIERI- On computing upper bounds in nonlinear problems involving disjunctive constraints		
2b FA348 - 513F - Convexification Techniques for MINLP				Chair: AMÉLIE LAMBERT
ABHISHEK PATHAPATI- Exact Augmented Lagrangian Duality for Mixed Integer Convex Optimization	AMÉLIE LAMBERT- Using quadratic cuts to iteratively strengthen convexifications of box quadratic programs	TOBIAS KUEN- The if-then Polytope: Conditional Relations over Multiple Sets of Binary Variables	MATTHIAS WALTER- Cutting planes for multilinear optimization	
2b FB9 - 513F - Theory of Mixed Integer Non-linear Programming				Chair: AIDA KHAJAVIRAD
ROBERT HILDEBRAND- Complexity of Mixed Integer Nonlinear Programming with Nonconvexities	SILVIA DI GREGORIO- A Knowledge Compilation Take on Binary Polynomial Optimization	AIDA KHAJAVIRAD- On the power of linear programming for data clustering	BARBARA ANNA FIEDOROWICZ- Sample Complexity and Computational Results for Branch and Cut Using Neural Networks	

Stream 2C: COMBINATORIAL OPTIMIZATION AND GRAPH THEORY			
2c MB223 - 511C - Complexity and algorithmic aspects of structured families of graphs 1 Chair: KATHIE CAMERON			
AURÉLIE LAGOUTTE- Local Certification of Geometric Graph Classes	MORITZ MÜHLENTHALER- Topological tools for H-recoloring and H-mixing	CÉSAR HERNÁNDEZ CRUZ- Hereditary properties defined by 2-edge-colourings avoiding a set of 2-edge-coloured graphs	KARTHEKEYAN CHANDRASEKARAN- Splitting-off in Hypergraphs
2c MB921 - 511B - Combinatorial Optimization and Graph Theory 1 Chair: TASUKU SOMA			
STEPHAN MARNACH- Weighted graph burning: Complexity and MIP formulations	VLADIMIR KOLMOGOROV- Computing cuts in graphs: Homory-Hu tree and the Sparsest Cut problem	JENNY SEGSCHEIDER- On the Polynomial Solvability of the Two-Stage Robust Perfect b-Matching Problem	TASUKU SOMA- Spectral sparsification of hypergraphs
2c MC250 - 511B - Graph Theory Chair: FRÉDÉRIC MEUNIER			
HADRIEN MÉLOT- Obtaining bounds on topological indices of chemical graphs	ALAIN HERTZ- Extremal Chemical Graphs for the Arithmetic-Geometric Index	GRÉGOIRE BEAUDOIRE- Antimagic labelling of (k,2)-bipartite biregular graphs	
2c TA263 - 511C - Graph Coloring Chair: ECKHARD STEFFEN			
JONAS ALKER- Column Generation for Generalized Min-Cost Flows	ISAAK H. WOLF- Pairwise disjoint perfect matchings in r -graphs	ALEXANDER CLOW- Injective Edge Colouring Some Classes of Sparse Graphs	ECKHARD STEFFEN- Sets of r -graphs that color all r -graphs
2c TA922 - 511B - Combinatorial Optimization and Graph Theory 2 Chair: ADOLFO ESCOBEDO			
MALEK MOUHOUB- Preference-based Constrained Optimization	JOSE WALTEROS- Exact approaches for solving interdependent multi-stage interdiction games: Applications in Communication Problems	EDWARD DUC HIEN NGUYEN- Sparse factorization of the square all-ones matrix of arbitrary order	ADOLFO ESCOBEDO- The Edge-based Contiguous p-median Problem with Connections to Territorial Districting
2c TB86 - 511B - Spectral Graph Theory and Optimization Chair: RENATA DEL VECCHIO			
HERMIE MONTERDE- Graphs with well-structured eigenbases	LEONARDO LIMA- On the max k -cut and the eigenvalue of a graph	CYBELE VINAGRE- Main Signless Laplacian Eigenvalues of Quasi-Threshold Graphs	SHAHLA NASSERASR- Zero forcing and eigenvalue multiplicities
2c TC359 - 511B - Advances in Combinatorial Optimization Chair: TAMON STEPHEN			
STEFFEN BORGWARDT- An Introduction to Circuit Walks	AMY WIEBE- A Yannakakis-type theorem for affine semigroups	TAMON STEPHEN- Hypergraph Transversal Pairs Near the Fredman-Khachiyan Bound	
2c WA169 - 511C - Complexity and algorithmic aspects of structured families of graphs 2 Chair: CÉSAR HERNÁNDEZ CRUZ			
KATHIE CAMERON- Reconfiguration of Vertex Colourings in Hereditary Classes of Graphs	JING HUANG- Gallai-like characterization of strong cocomparability graphs	M. GABRIELA ARAUJO PARDO- On Mixed Cages	JAN BOK- Complexity dichotomies for list homomorphism problems of signed graphs
2c WA923 - 511B - Combinatorial Optimization and Graph Theory 3 Chair: SATORU IWATA			
ANJA FISCHER- The betweenness model for the single-row facility layout problem and beyond	FREIJA VAN LENT- Envy-free Pricing of Seats in a Planetarium	FRANCISCO BARAHONA- Packing Hypertrees and the k -cut problem in hypergraphs	SATORU IWATA- Finding Maximum Edge-Disjoint Paths Between Multiple Terminals
2c WB15 - 511B - Discrete Optimization Chair: MARCIA FAMPA			
GABRIEL PONTE- Good and Fast Row-Sparse ah-Symmetric Reflexive Generalized Inverses	DANIELA LUBKE- A column-generation based heuristic for integrating machine scheduling and personnel allocation	YUAN ZHOU- Semialgebraic characterization of triangulation	YUSHAN QU- On disjunction convex hulls by lifting
2c WB185 - 511C - Complexity and algorithmic aspects of structured families of graphs 3 Chair: CHÍNH HOÀNG			
ANGELE FOLEY- H-Chromatic Symmetric Functions	CHÍNH HOÀNG- Coloring graphs by forbidden induced subgraphs	R SRITHARAN- Intersection of chordal graphs and some related partition problems	BEN CAMERON- Efficient certifying algorithms for k -colouring subfamilies of P_5 -free graphs

2c WC353 - 511B - Combinatorial Optimisation			Chair: FABIO FURINI
FRANKLIN DJEUMOU FOMENI- Recent heuristic development for the quadratic knapsack problem	FABIO FURINI- Critical infrastructures location via submodular maximization	ASVIN GOEL- Processual programming - Using business process models to formulate mathematical programs	
2c ThA924 - 511B - Combinatorial Optimization and Graph Theory 4			Chair: LUCAS MOSCHEN
ROLAND GRAPPE- Totally equimodular matrices	VIET HUNG NGUYEN- A 2/3-approximation algorithm for the maximum weight spanning star forest problem	SIYUE LIU- On the Congruency-Constrained Matroid Base	LUCAS MOSCHEN- Combinatorial algorithms for large household allocation problems
2c ThB352 - 511B - Integer/Linear Optimisation			Chair: DANIELE CATANZARO
YASUKO MATSUI- Enumerating full binary trees in polynomial delay	DANIELE CATANZARO- Optimizing over Path-Length Matrices of Unrooted Binary Trees	BENTO NATURA- Breaking the quadratic gap for strongly polynomial solvers to combinatorial linear programs	ZHUAN KHYE KOH- A strongly polynomial algorithm for the minimum cost generalized flow problem
2c ThC355 - 511B - Kernels in Graphs			Chair: FRÉDÉRIC MEUNIER
HÉLÈNE LANGLOIS- Kernels and quasi-kernels	FRÉDÉRIC MEUNIER- Extensions of classical results on kernels	PÉTER ERDOS- Quasi-kernels in finite and countable infinite (?) directed graphs	
2c FA351 - 511C - Combinatorial Optimisation and Graph Theory			Chair: MARTIN DREES
RUI GONG- Rounding the Lovasz Theta Function with a Value Function Approximation	MARTIN DREES- Intersecting and Dense Restrictions of Clusters in Polynomial Time	VÍCTOR BLANCO- Optimizing Mediated Graphs: Applications in Conic Optimization and Representations of Circuits as Sums-Of-Squares	NIKLAS SCHLOMBERG- Packing cycles in planar and bounded-genus graphs
2c FA925 - 511B - Combinatorial Optimization and Graph Theory 5			Chair: AKIYOSHI SHIOURA
RICARDO EULER- Labeling Methods for Partially Ordered Paths	YUICHIRO MIYAMOTO- Measuring Constituency Based On Minimum Campaign Distance And Application To Evaluate The Merger Of Electoral Districts	MARTINA CERULLI- The Colorful Components Problems	AKIYOSHI SHIOURA- On the monotonicity in steepest descent algorithm for M-convex functions
2c FB926 - 511B - Combinatorial Optimization and Graph Theory 6			Chair: JOSEPH POREMBA
HIROTAKE ONO- Bounding the Proper Orientation Number by Tree-depth	CLAUDIA O LÓPEZ SOTO- Beam Search to Minimise Cutting Patterns with Maximum Utilisation	EL MEHDI OUDAUD- 3D Phase Unwrapping using 3D Surface minimization.	JOSEPH POREMBA- Uncrossable Multicommodity Flows
Stream 2D: MACHINE LEARNING AND DISCRETE OPTIMIZATION			
2d MB248 - 513F - Distance Optimization in Data Science			Chair: EMILIO CARRIZOSA
JASONE RAMIREZ AYERBE- A Novel Column Generation Framework for One-for-many Counterfactual Explanations	NURIA GÓMEZ VARGAS- Solving the Contextual Multiobjective Inverse Ideal Point Problem via Mathematical Optimization	THOMAS HALSKOV- Distance-Based Fairness in Classification and Regression	EMILIO CARRIZOSA- Sequential Counterfactual Decisions
2d MC298 - 513A - Classification and regression via mathematical optimization			Chair: DOLORES ROMERO MORALES
ANTONIO CONSOLO- Soft regression trees: a model variant and a decomposition training algorithm	MARIA EDUARDA PINHEIRO- Mixed-Integer Linear Optimization for Cardinality-Constrained Random Forests	DOLORES ROMERO MORALES- On Enhancing the Explainability and Fairness of Tree Ensembles	
2d TA297 - 513A - Optimization for Safe and Private Machine Learning			Chair: THIBAUT VIDAL
FRAUKE LIERS- A Framework for Data-driven Explainability in Mathematical Optimization	ROBERT BASSETT- Get Off The Hype Train: Why Using Neural Networks Can Be Very Risky	THIBAUT VIDAL- OR with a White Hat : Evidencing Privacy Vulnerabilities in ML Models	MATTEO FRANCOBALDI- SMLE: Safe Machine Learning via Embedded Overapproximation

2d TB199 - 513A - Learning to Select				Chair: VASSILIS DIGALAKIS
GEORGE ORFANIDES- Minimizing difference of submodular functions via DC programming	VASSILIS DIGALAKIS- Slowly varying regression under sparsity	MICHAEL LI- Learning to Cover: Online Learning and Optimization with Irreversible Decisions	STEFANOS POULIDIS- Optimal Experimentation for Learning Personalized Policies Across Locations	
2d TC289 - 513A - On Machine Learning for Mathematical Optimization				Chair: ANDREA LODI
HONGYU CHENG- A Data-Driven Approach to Cut Generating Function Selection	RICK WILLEMSSEN- Variable aggregation in a Benders decomposition for the p-median problem	ROSARIO MESSANA- Interactive learning of a knapsack constraint using a membership oracle		
2d WA927 - 513A - Machine Learning and Discrete Optimization 1				Chair: YASSINE YAAKOUBI
ROLAND WUNDERLING- Gurobi Machine Learning	JULIEN PALLAGE- Wasserstein Distributionally Robust Shallow Convex Neural Networks	DEFENG LIU- A machine learning approach for neighbor generation in metaheuristic search	YASSINE YAAKOUBI- Machine Learning for Distributionally Robust Warm-Starting in Mineral Supply/Value Chains	
2d WB230 - 513A - Combinatorial Optimization Augmented Machine Learning and Decision Aware Learning				Chair: AXEL PARMENTIER
CHEIKH AHMED- Learning to solve large-scale districting problems	AXEL PARMENTIER- Surrogate policies with generalization guarantees for combinatorial optimization problems	ESTHER JULIEN- Neur2BiLO: Neural Bilevel Optimization	ELLEN VITERCIK- From Large to Small Datasets: Size Generalization for Clustering Algorithm Selection	
2d WC203 - 513A - Machine Learning and Combinatorial Optimization				Chair: BARTOLOMEO STELLATO
JIM LUEDTKE- Valid Inequalities for Verification of Binarized Neural Networks	SIMONE MILANESI- Slimming down and beefing up: MILP Multi-Objective Optimization and Training for Few-Bit Neural Networks	BARTOLOMEO STELLATO- Differentiable Cutting-plane Layers for Mixed-integer Linear Optimization		
2d ThA211 - 513A - Explainable Artificial Intelligence and Mathematical Programming				Chair: JANNIS KURTZ
PATRICK VOSSLER- ODTlearn: A Python Package for Learning Optimal Decision Trees for Prediction and Prescription	ILKER BIRBIL- Finding Regions of Counterfactual Explanations via Robust Optimization	NATHANAEL JO- Learning Optimal Prescriptive Trees from Observational Data	JANNIS KURTZ- Counterfactual Explanations for Linear Optimization	
2d ThB200 - 513A - Scaling Sparse Machine Learning				Chair: VASSILIS DIGALAKIS
ISHY ZAGDOUN- Projection onto a Capped Rotated Second-Order Cone with Applications to Sparse Regression Relaxations	GUANYI WANG- Inexact Stochastic Proximal Methods for Sparse Group Structure Identification	ILIAS ZADIK- On the power of greedy local search for sparse statistical problems	ANNA DEZA- Logic Rules and Chordal Graphs for Sparse Learning	
2d ThC928 - 513A - Machine Learning and Discrete Optimization 2				Chair: RAHUL MAZUMDER
JARED TANNER- Deep neural network initialisation: Nonlinear activations impact on the Gaussian process	ALEXANDRE FOREL- Functionally-Identical Pruning of Tree Ensembles	RAHUL MAZUMDER- Discrete Optimization Methods for Pruning and Compressing Large Neural Networks		
2d FA173 - 513A - Progress on Machine Learning Approaches for Solving Combinatorial Optimization				Chair: ZAIWEN WEN
ZAIWEN WEN- A Monte Carlo Policy Gradient Method with Local Search for Binary Optimization	AKANG WANG- Mixed-Integer Linear Optimization via Learning-Based Two-Layer Large Neighborhood Search	JUNCHI YAN- Understanding and Recomposing the Learning and Searching Techniques in ML4CO: An Deep Dive into the TSP Case	JIANYONG SUN- Learning to learn for discrete optimization: some recent developments	
2d FB929 - 513A - Machine Learning and Discrete Optimization 3				Chair: FEDERICO D'ONOFRIO
BRIAN LIU- FASTopt: An Optimization Framework for Fast Additive Segmentation in Transparent ML	PIERRE HOUDOUIN- Reliability evaluation of gaussian processes applied to power systems	STEVEN DAHDAH- Closed-loop Koopman operator approximation	FEDERICO D'ONOFRIO- Novel Mixed-Integer Optimization Approaches for Interpretable SVMs	

Stream 2E: APPROXIMATION AND ONLINE ALGORITHMS			
2e MB153 - 513E - Stochastic Combinatorial Optimization I			Chair: NICOLE MEGOW
MARC UETZ- Stochastic Scheduling of Bernoulli Jobs	SHARAT IBRAHIMPUR- Approximation Algorithms for Stochastic Minimum Norm Combinatorial Optimization	THOMAS KESSELHEIM- Learning Prophet Inequality and Pandora's Box with Limited Feedback	NICOLE MEGOW- Set Selection under Explorable Stochastic Uncertainty
2e MB239 - 514A - Network design			Chair: NEIL OLVER
LAURA VARGAS KOCH- Single-source unsplittable flows in planar graphs	ELLIS HERSHKOWITZ- New Structures and Algorithms for Length-Constrained Expander Decompositions	NATHAN KLEIN- Ghost Value Augmentation for k-Edge-Connectivity	RHEA JAIN- LP-Based Algorithms for Two-Cost Network Design
2e TA229 - 512A - Algorithms for Pricing and Assortment Optimization			Chair: OMAR EL HOUSNI
RAJAN UDWANI- Submodular Order Functions and Their Applications	MIKA SUMIDA- Optimizing and Learning Assortment Decisions in the Presence of Platform Disengagement	SANJAY MEHROTRA- Learning to Price Under Competition for Multinomial Logit Demand	OMAR EL HOUSNI- Maximum Load Assortment Optimization
2e TA930 - 514B - Approximation and Online Algorithms 1			Chair: ALENA OTTO
DANISH KASHAEV- Online Matching on 3-Uniform Hypergraphs	SANDER BORST- Nearly optimal node selection in the explorable heap model	ANDREAS KARRENBauer- TSP-T3CO: A Classification Scheme for Variants of the Traveling Salesman Problem	ALENA OTTO- Reliable Adjustable Planning in Case of Dynamically Arriving Information: Example of E-Commerce Warehouses
2e TC358 - 514B - Approximation in Algorithmic Game Theory			Chair: CHAITANYA SWAMY
POOJA KULKARNI- Fair Allocation of Indivisibles Beyond Additive Valuations	RIAN NEOGI- Budget-Feasible Mechanism Design: Simpler, Better Mechanisms and General Payment Constraints	CHAITANYA SWAMY- Multidimensional Budget-Feasible Mechanism Design	
2e WA931 - 514B - Approximation and Online Algorithms 2			Chair: JAROSLAV FOWKES
THOMAS ERLEBACH- Scheduling with Obligatory Tests	ALBERTO LOCATELLI- An Exact Method for the Job Sequencing and Tool Switching Problem	ALANTHA NEWMAN- Recent progress for correlation clustering	JAROSLAV FOWKES- Approximating large-scale Hessian matrices using secant equations
2e WB146 - 514B - Stochastic Combinatorial Optimization II			Chair: VISWANATH NAGARAJAN
YIFAN SUN- Fast Online Node Labeling for Very Large Graphs	ROHAN GHUGE- Semi-Bandit Learning for Monotone Stochastic Optimization	VISWANATH NAGARAJAN- Power of Adaptivity for Stochastic Submodular Cover	KEVIN SCHEWIOR- Recent Advances in Stochastic Function Evaluation: Score Classification and Voting
2e ThA293 - 514B - Approximation Algorithms for the Traveling Salesman Problem			Chair: DAVID WILLIAMSON
BILLY JIN- A Lower Bound for the Max Entropy Algorithm for the Traveling Salesman Problem	JENS VYGEN- Approximation Algorithms for Traveling Salesman Problems	RICHARD SHAPLEY- Small additive error for unsplittable multicommodity flow in outerplanar graphs	DAVID WILLIAMSON- A 4/3-Approximation Algorithm for Half-Integral Cycle Cut Instances of the TSP
2e ThB244 - 514B - Advanced algorithms for chip design			Chair: STEPHAN HELD
SUSANNE ARMBRUSTER- Depth Optimization in Binary Addition	STEPHAN HELD- Tighter Approximation for the Uniform Cost-Distance Steiner Tree Problem	DANIEL BLANKENBURG- Resource Sharing revisited: Local weak duality and optimal convergence	JAN MALTE SCHÜRKES- Decision-guided SAT for hard routing problems
2e FA150 - 514B - Submodularity and Beyond in Approximation Algorithms			Chair: KARTHEKEYAN CHANDRASEKARAN
ISHAN BANSAL- Improved Approximation Algorithms Beyond Uncrossable Functions	SHUBHANG KULKARNI- Polyhedral Aspects of Feedback Vertex Set and Pseudoforest Deletion Set	KENT QUANRUD- Quotient sparsification for submodular functions	SABINE MÜNCH- Maximizing a Monotone Submodular Function Under an Unknown Knapsack Capacity

2e FA245 - 512A - TSP and vehicle routing			Chair: STEFAN HOUGARDY
LUISE PUHLMANN- Improved guarantees for the a priori TSP	STEFAN HOUGARDY- The k-Opt algorithm for the Traveling Salesman Problem has exponential running time for $k \geq 5$	ELAHE AMIRI- Anytime Optimization Approaches for Online dial-a-ride problem	DIRK MÜLLER- Vehicle Routing with Time-Dependent Travel Times: Theory, Practice, and Benchmarks
2e FB225 - 514B - Fairness via Optimization Modeling and Approximation Algorithms			Chair: DAVID SHMOYS
VARUN SURIYANARAYANA- Improved Approximation Algorithms for the Joint Replenishment Problem with Outliers, and with Fairness Constraints	MAROUANE IBN BRAHIM- Assortment Optimization with Visibility Constraints	ARAVIND SRINIVASAN- Fairness, Randomization, and Approximation Algorithms	DAVID SHMOYS- Algorithmic Tools for Redistricting: Fairness via Analytics
2e FB295 - 512A - Algorithms with Predictions			Chair: CLIFFORD STEIN
ERIC BALKANSKI- Fair Secretaries with Unfair Predictions	BENJAMIN MOSELEY- Incremental Topological Ordering and Cycle Detection with Predictions	THOMAS LAVASTIDA- Algorithms with Prediction Portfolios	CLIFFORD STEIN- Energy-Efficient Scheduling with Predictions
Stream 2F: COMPUTATIONAL DISCRETE AND INTEGER OPTIMIZATION			
2f MB344 - 513C - Modeling Software			Chair: DAN STEFFY
GREGOR HENDEL- FICO® Xpress: New and Augmented Interfaces for Modeling Optimization Problems	DAN STEFFY- Practical guidelines for model improvement and reformulation	MILES LUBIN- What's new in JuMP	CAN LI- OptiChat: An AI Assistant for Explaining Optimization Models Powered by LLM
2f MB932 - 513B - Computational Discrete and Integer Optimization 1			Chair: FRED GARDI
FREDERIC KRONER- Optimizing Airport Selection for Electric Aircraft Operations Across Europe	SABRINA AMMANN- A speed-up for Helsgaun's TSP heuristic by relaxing the positive gain criterion	LISA MARIE MANKE- A New Labeling Approach for a Special Case of the Multicommodity Flow Problem	FRED GARDI- Hexaly, a new kind of global optimization solver
2f MC339 - 513E - Extensions of Integer Programming			Chair: MARIANNA DE SANTIS
MARIANNA DE SANTIS- On the accurate detection of the Pareto frontier for bi-objective mixed integer linear problems	MARK TURNER- PySCIPOpt-ML: Embedding Trained Machine Learning Models into Mixed-Integer Programs	PEDRO MACIEL XAVIER-QUBO.jl: A tale of implementation and benchmarking of a Quantum Optimization Ecosystem in Julia	
2f TA269 - 513E - Solver Software I			Chair: TIMO BERTHOLD
TIMO BERTHOLD- What's New in FICO Xpress Solver?	MICHAEL WINKLER- What's new in Gurobi 11	JULIAN HALL- Enhancing the HiGHS MIP solver	NILS CHRISTIAN KEMPE- Recent Progress in the Cardinal Optimizer
2f TB340 - 513E - Solver Software II			Chair: PHILIPP CHRISTOPHEL
SOFIANE TANJI- A modeling language-based approach to automatically recommend first-order optimization methods.	SVEN WIESE- Symmetry detection in Mixed-Integer Conic Programming	JULIEN DARLAY- Hybridizing combinatorial heuristics and continuous optimization methods for Mixed-Integer Programming	PHILIPP CHRISTOPHEL- Recent Advances in the SAS Linear and Mixed-integer Optimization Solvers
2f WA281 - 513E - New Algorithms for Large-Scale Optimization			Chair: DANIEL ALOISE
ANTONIO M. SUDOSO- Dynamic constraint aggregation for solving the minimum sum-of-squares clustering problem	ALPHA SALIOU BARRY- Generalized integral simplex	FRANCOIS SOUMIS- Integral Simplex for the Set Partitioning Problem	YOUSSEF EMINE- New complementary problem formulation for the improved primal simplex
2f WB341 - 513E - Computational Mixed Integer Programming			Chair: PIERRE BONAMI
GIONI MEXI- Generalized Resolution Conflict Analysis in MIP Solvers	PIERRE BONAMI- Recent additions to the collection of cutting planes in Gurobi	KRUNAL KISHOR PATEL- Leveraging Discounted Pseudocosts in MILP for Enhanced Branching Strategies	LUKAS SCHÜRMANN- Variable Fixing for Integer Linear Programming via Objective Rotation
2f WC266 - 513E - Certified Optimization			Chair: JONAD PULAJ
JONAD PULAJ- An SMT-based checker for VIPR certificates of MILP results	ALEXANDER HOEN- Certificates of Correctness for Presolving Binary Programs	ALDO KIEM- The Four-Color Ramsey Multiplicity of Triangles	

2f ThA343 - 513E - Decomposition Methods				Chair: MARCO LÜBBECKE
KIBA EK KIM- A scalable Master Problem in Dual Decomposition of Stochastic MIP	MARCO LÜBBECKE- A Row-and-Column Generation Approach to Min-Sum Job-Shop Scheduling	ANTOINE LEGRAIN- Decomposition method for a capacitated multi-vehicle covering tour problem with intermediate facilities	ELINA RÖNNBERG- Integer programming column generation: Accelerating branch-and-price for set covering, packing, and partitioning problems	
2f ThB342 - 513E - Combinatorial Optimization and Applications				Chair: JAAP PEDERSEN
MOHAMMED GHANNAM- Branch and Price for the Length-Constrained Cycle Partition Problem	SHENG JIE CHEN- Cutting planes for the maximum k-vertex cover problem	JAAP PEDERSEN- Quota Steiner Tree Problem and its application on Wind Farm Planning	JULIAN YARKONY- A New Class of Compact Formulations for Vehicle Routing Problems	
2f ThC270 - 513E - Augmentation Methods and Dynamic Programming				Chair: DOMINIC YANG
TOLGA BEKTAS- Augmentation search for integer programmes defined on polyhedra	DOMINIC YANG- Algebra-Inspired Randomized Augmentation Methods in Combinatorial Optimization	IBRAHIM DAN DIJE- A lifted-space dynamic programming algorithm for the Cubic Knapsack Problem (CKP)		
2f FA224 - 513C - Tightening MILPs by Reformulation				Chair: DOMINIK KAMP
JÖRG RAMBAU- Augmented subproblem grinding for the lot-type design problem	JOÃO DIONÍSIO- An acceleration scheme for column generation in production-maintenance scheduling problems	JOHANNES SCHMIDT- A new approach for inter-sample avoidance using an integer polytope formulation	DOMINIK KAMP- Dynamic Demand Propagation in Guaranteed Service Models	
2f FA933 - 513E - Computational Discrete and Integer Optimization 2				Chair: QUENTIN LOUVEAUX
ANTHONY KARAHALIOS- Column Elimination	ANDREAS TILLMANN- Sequential Mixed-Integer Programming for Matrix Sparsification	SHUNJI UMETANI- An improved weighting local search for large-scale binary integer programs	QUENTIN LOUVEAUX- Sensitivity analysis for linear change of the constraint matrix of an LP	
2f FB268 - 513E - Integer Programming Applications in Mobility and Traffic				Chair: NIELS LINDNER
NIELS LINDNER- On the Split Closure of the Periodic Timetabling Polytope	ENRICO BORTOLETTO- Integrated Periodic Event Scheduling and Infrastructure Assignment: approaches via Cyclic Orders and Graph Matchings	BERENIKE MASING- A Flexible Model for Integrated Line Planning and Periodic Timetabling with Track Choice in Practice	FABIAN LÖBEL- The Approximation Error Caused by Linearizing Charging Behavior for Electric Vehicle Scheduling Problems	
Stream 2G: CONSTRAINT PROGRAMMING				
2g WA243 - 512F - Constraints and Machine Learning				Chair: CHENBO SHI
RICARDO FUKASAWA- Optimal counterfactual explanations for k-Nearest Neighbors using Mathematical Optimization and Constraint Programming	RAHUL PATEL- MORBDD: Multiobjective Restricted Binary Decision Diagrams by Learning to Sparsify	LOUIS MARTIN ROUSSEAU- Learning Lagrangian Multipliers for the Travelling Salesman Problem	CHENBO SHI- Constraint Learning to Define Trust Regions in Optimization Over Pre-Trained Predictive Models	
2g WC241 - 514B - CPAIOR at ISMP: Mathematical Programming, Constraint Programming, and Hybrid Methods				Chair: LAURENT MICHEL
JOHN HOOKER- A reinterpretation of cutting planes	SEYEDE SAEDE HOSSEINI- A Constraint-Programming-Based Branch-and-Check Approach for Stochastic Consistent Home Health Care Routing and Scheduling Problem	LAURENT MICHEL- C++DDOpt : A Modeling and Solving Interface for Decision Diagram-based Optimization.		
2g FB242 - 512G - Learning and Explaining Constraints				Chair: NADJIB LAZAAR
J. CHRISTOPHER BECK- Domain-Independent Dynamic Programming	ALEXANDER TESCH- Lazy MIP Solving with Methods from SAT	IAN MIGUEL- Automated Streamlining for Constraint Satisfaction Problems.	NADJIB LAZAAR- Query-Based Constraint Acquisition	

CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY

Stream 3A: CONTINUOUS STOCHASTIC PROGRAMMING

3a MB107 - 515B - Multistage Stochastic Programs Chair: ZHIPING CHEN

ZHIPING CHEN- A Bayesian approach to data-driven multi-stage stochastic optimization	RAIAN NOUFEL LEFGOUM- Multistage stochastic optimization of an elementary hydrogen infrastructure	LEONARD MACLEAN- The Kelly Strategy - Modifications to Address Market Changes and Investor Preferences	JIA LIU- Multi-stage distributionally robust optimization with Bayesian learning moments uncertainty
--	---	--	--

3a MC934 - 515A - Continuous Stochastic Programming 1 Chair: JUNYU ZHANG

DI ZHANG- A Sampling-based Progressive Hedging Algorithm for Stochastic Programming	VINCENT LECLÈRE- Generalized adaptive partition-based method for two-stage stochastic linear programs: geometric oracle and analysis	JUNYU ZHANG- Stochastic Bregman Proximal Gradient Method Revisited: Kernel Conditioning and Painless Variance Reduction	
---	--	---	--

3a TA115 - 515A - Advances in First-Order Methods for Continuous and Stochastic Optimization Chair: TIANJIAO LI

LIWEI JIANG- Asymptotic Normality and Optimality in Nonsmooth Stochastic Optimization	TIANJIAO LI- A simple uniformly optimal method without line search for convex optimization	JIAMING LIANG- Universal First-Order Methods for Convex Optimization Problems	JIMMY ZHANG- Optimal and parameter-free gradient minimization methods for convex and nonconvex optimization
---	--	---	---

3a TB111 - 515A - PDE-constrained optimization under uncertainty Chair: DANIEL WALTER

CAROLINE GEIERSBACH- Optimality Conditions with Probabilistic State Constraints	JOHANNES MILZ- Convergence rates for ensemble-based solutions to optimal control of uncertain dynamical systems	MATTHIAS HEINKENSCHLOSS- Adaptive Surrogate Modeling for Trajectory Optimization with Model Inexactness	DANIEL WALTER- Empirical estimators for risk-neutral composite optimal control with applications to bang-bang control
---	---	---	---

3a WA102 - 515A - Multistage Optimization Programs under Uncertainty Chair: FRANCESCA MAGGIONI

RUI GAO- Multistage distributional robustness, time consistency and optimal policy	ABDEL LISSER- Distributionally robust chance-constrained Markov decision processes	DARINKA DENTCHEVA- Risk Evaluation and Control for Distributed Multi-Agent Systems	
--	--	--	--

3a WB232 - 515A - Decision-Focused Learning and Bilevel Models for Stochastic Optimization Chair: TITO HOMEM DE MELLO

RUBEN VAN BEESTEN- Assessing solution quality in risk-averse stochastic programs	FRANCISCO JARA MORONI- An implementation of the Adaptive Partition Method to Stochastic Bilevel Problems	VÍCTOR BUCAREY- Decision-focused predictions via pessimistic bilevel optimization: a computational study	TITO HOMEM DE MELLO- Application-Driven Optimal Pointwise Forecasts for Stochastic Optimization
--	--	--	---

3a WC258 - 515A - Theory and Algorithms for Chance Constrained Optimization Chair: ANIRUDH SUBRAMANYAM

DANIELA BERNHARD- Joint chance-constrained optimization with continuous distributions	SAI RAKSHITH- Self-Structured Importance Sampling for Chance-Constrained Optimization	YUKI SHIRAI- Chance-Constrained Optimization for Uncertain Contact-Rich Systems	
---	---	---	--

3a ThA101 - 515A - Decision-making under Preference Uncertainty Chair: WENJIE HUANG

SILVANA PESENTI- Risk Budgeting Allocation for Dynamic Risk Measures	REZA VALIMORADI- Nonparametric Inverse Optimization: Theory and Applications	HUIFU XU- Preference Ambiguity and Robustness in Multistage Decision Making	WENJIE HUANG- Multi-attribute Preference Robust Optimization with Quasi-Conave Choice Functions
--	--	---	---

3a ThB935 - 515A - Continuous Stochastic Programming 2 Chair: DIONYSIS KALOGERIAS

NICOLAS GAGNÉ- Duality in Optimal Stopping	MARIUS ROLAND- Adaptive Partitioning for Chance-Constrained Problems with Finite Support	CARLOS CARDONHA- Consistent Assortment Optimization under the MNL with Exogenous Data	DIONYSIS KALOGERIAS- Strong Duality in Risk-Constrained Nonconvex Functional Programming
--	--	---	--

3a ThC252 - 515A - Advances in Risk Optimization and Design			Chair: QUANYAN ZHU
YINGDONG LU- Federated Learning for Discrete Optimal Transport with Large Population under Incomplete Information	UTSAV SADANA- Data driven decision making under uncertainty with entropic risk measure	SHUTIAN LIU- Stackelberg risk preference design	
3a FA247 - 515A - Latest Developments in Stochastic Optimization			Chair: JONATHAN LI
DAVID SAUNDERS- Reinforcement Learning for Portfolio Optimization in Model with Regime-Switching	ANDREW LIM- Optimal trade execution with learning	QINYU WU- On Generalization and Regularization via Wasserstein Distributionally Robust Optimization	TIANTIAN MAO- Model Aggregation for Risk Evaluation and Robust Optimization
3a FB936 - 515A - Continuous Stochastic Programming 3			Chair: SHIBSHANKAR DEY
FEDERICO BOBBIO BOBBIO- Stable Matching with Contingent Priorities	SHIBSHANKAR DEY- On the Chance-Constrained Optimization with Gaussian Mixture Models		
Stream 3B: DISCRETE STOCHASTIC PROGRAMMING			
3b MB287 - 512H - Recent modeling language extensions: Stochastics and beyond			Chair: MICHAEL FERRIS
STEVEN DIRKSE- GAMSPy: Algebraic Modeling and Python	OLIVIER HUBER- DAG-based modeling framework for extended mathematical programming	ROBERT FOURER- Advances in Automated Conversion of Optimization Problems	MICHAEL FERRIS- Stochastic Equilibrium Problems
3b MC57 - 512A - Modeling and Optimization under Uncertainty			Chair: SUVRAJEET SEN
DAVID MORTON- Distributionally robust two-stage convex quadratic programming	SUVRAJEET SEN- Multi-scale and Multi-stage Stochastic Discrete Optimization and its Applications in Energy and Sustainability	SIMGE KUCUKYAVUZ- Integer Programming for Learning Directed Acyclic Graphs from Non-identifiable Gaussian Models	
3b TB187 - 512A - Modeling and Algorithms for Stochastic Discrete Optimization			Chair: HAOXIANG YANG
ZHUO ZHANG- Optimization via Simulation with Discontinuous Recourse Function	XINGYU LIN- A Distributed Optimization Method for Large-Scale Production Planning Problem	CHENG GUO- Risk-Aware Security-Constrained Unit Commitment	LUHAO ZHANG- Beyond Absolute Continuity: a New Class of Dynamic Risk Measures
3b TC96 - 512A - Theory and Applications of Robust and Stochastic Integer Programming			Chair: WEIJUN XIE
YONGJIA SONG- Stochastic Programming for Disaster Housing Assistance Planning	JUAN ALBERTO ESTRADA GARCIA- Dynamic Transmission Line Switching Amidst Wildfire-Prone Conditions	BRADLEY STURT- Improving the Security of United States Elections with Robust Optimization	
3b WA58 - 512A - Stochastic Integer Programming: Models and Algorithms			Chair: JIM LUEDTKE
ERFANEH NIKZAD- Master Surgery Scheduling Problem Considering Uncertainty in Time of Stay in Downstream Units; A Matheuristic Solution Approach	JEFF LINDEROTH- Probing-Enhanced Stochastic Programming	WOOJIN KIM- The Chance-constrained Stochastic Diversion Path Problem with Sample Average Approximation	EDUARDO MORENO- Stochastic facility location problem with outsourcing costs
3b WB178 - 512A - Recent Advances to Solve Discrete Stochastic Problems			Chair: AYSENUK KARAGOZ
HARSHIT KOTHARI- Accelerating Benders decomposition for solving a sequence of sample average approximation problems	WARD ROMEIJNDERS- Scaled cuts for multistage stochastic mixed-integer programs	AYSENUK KARAGOZ- Single-Scenario Facet Preservation for Stochastic Mixed-Integer Programs	JANOSCH ORTMANN- Test sets to compute opportunity cost matrices
3b ThA937 - 512A - Discrete Stochastic Programming 1			Chair: MARYAM DARYALAL
EMANUELE CONCAS- A Decomposition Method for Network Revenue Management with Dynamic Pricing and Bundles	PIETER KLEER- Robust Optimality of Bundling I.I.D. Goods Beyond Finite Variance	MARIA CAROLINA BAZOTTE CORGOZINHO- The SAA Method for Two-Stage Stochastic Programs with Endogenous Uncertainty	MARYAM DARYALAL- A Risk-aware Location-Allocation-Pricing Problem with Stochastic Price-Sensitive Demands

3b ThB124 - 512A - Sequential Decision Making Under Uncertainty				Chair: RUI CHEN
RUI CHEN- Online Decision Making with Nonconvex Local and Convex Global Constraints	HAOMING SHEN- Value of Stochastic Solution with Right-Hand Side Uncertainty	DAHYE HAN- Regularized MIP Model for Optimal Power Flow with Energy Storage Systems and its Applications	ALFREDO TORRICO- Two-sided assortment optimization: Adaptivity gaps and approximation algorithms	
3b ThC116 - 512A - Advanced algorithms in stochastic mixed-integer programming				Chair: KIBA EK KIM
KAIWEN FANG- Strengthening Benders cuts via disjunction	HANBIN YANG- Stochastic Cutting-plane Dynamic integer Programming	GEUNYEONG BYEON- Decomposition Algorithm for Two-Stage Distributionally Robust Optimization over 1-Wasserstein Balls: Enhancing Performance and Scalability		
Stream 3C: ROBUST AND DISTRIBUTIONALLY ROBUST OPTIMIZATION				
3c MB108 - 511D - Applications of Robust and Distributionally Robust Optimization				Chair: ZHEN YANG
ZHI CHEN- Budget-Driven Multi-Period Hub Location: A Robust Time Series Approach	IAN YIHANG ZHU- Network Flow Models for Robust Binary Optimization with Selective Adaptability	LI CHEN- Robust Optimization with Moment-Dispersion Ambiguity	ZHEN YANG- Wasserstein Regularization for 0-1 Loss	
3c TA99 - 510D - Robust Optimization and Machine Learning 1				Chair: ARAS SELVI
SERGIO CAMELO- Platform Design for the First Mile of Commodity Supply Chains	MAN CHUNG YUE- An MILP-Based Solution Scheme for Factored and Robust Factored Markov Decision Processes	ARAS SELVI- Wasserstein Adversarial Logistic Regression with Synthetic Data	ABHILASH CHENREDDY- End-to-end Conditional Robust Optimization	
3c TA170 - 511E - Recent Advances in the Theory of Stochastic and Robust Optimization				Chair: HAOMING SHEN
ANGELOS GEORGHIOU- Robust Data-driven Prescriptiveness Optimization	LOUIS CHEN- On the Adversarial Robustness of Benjamini-Hochberg	LI XIAOBO- Fourth-order Marginal Moment Model: Reformulations and Applications		
3c TA938 - 511D - Robust and Distributionally Robust Optimization 1				Chair: MATTEO PETRIS
VITOR LUIZ PINTO DE PINA FERREIRA- Cut-Generation Algorithms for Uncertain Unit Commitment under Mixed-Information Models	ZUKUI LI- Distributionally robust optimization using optimal transport for Gaussian mixture models	MATTEO PETRIS- Robust approaches for the Kidney Exchange Problem		
3c TB100 - 511D - Robust Optimization and Machine Learning 2				Chair: MENG MENG LI
YUHAO WANG- Bayesian Risk-Averse Reinforcement Learning	CHIN PANG HO- Robust Regret Markov Decision Processes	DUO ZHOU- Distributionally Robust Path Integral Control	MENG MENG LI- Towards Optimal Offline Reinforcement Learning	
3c TB168 - 511E - Robust Optimization in practice				Chair: AHMADREZA MARANDI
HOURA MAHMOUDZADEH- Superior and Light Pareto Robust Optimization for IMRT Treatment Planning	JUSTIN STARREVELD- Dealing With Uncertainty Over Time When Optimizing Industrial Decarbonization Pathways	PERIKLIS PETRIDIS- A Holistic Framework for Decarbonization at OCP	MINAKSHI PUNAM MANDAL- Robust Facility Location in Disaster Preparation for Earthquakes with Aftershocks	
3c WA106 - 511D - Robust Optimization and Machine Learning 3				Chair: ARAS SELVI
JOSE BLANCHET- On Foundations of Distributionally Robust Reinforcement Learning	JIE WANG- Regularization for Adversarial Robust Learning	KARTHYEK MURTHY- Distributionally Robust Affine Markov Decision Processes	CAIHUA CHEN- Enhancing Data-Driven Distributionally Robust Optimization: Jointly Incorporating First-Order Moment and Wasserstein Distance	
3c WB125 - 511E - Robust Optimization and Machine Learning 4				Chair: MENG MENG LI
SHIMRIT SHTERN- First-order data-driven distributionally robust optimization	AHARON BEN TAL- New algorithm for the difference of convex functions optimization problem	VAMSI POTLURU- Towards Watermarking Tabular Datasets	SHUO SUN- Online MDP with Transition Prototypes: A Robust Adaptive Approach	

3c WB939 - 511D - Robust and Distributionally Robust Optimization 2			Chair: FABRICIO OLIVEIRA
EVA LEY- Robust Optimization Under Controllable Uncertainty	KEVIN MARTIN AIGNER- Clustering-based scenario reduction for distributionally robust optimization with approximation guarantees	TIANXIAO WANG- Distributionally Robust Optimization Approaches for Resilience and Low Carbon Enhancement of Distribution Power Systems	FABRICIO OLIVEIRA- Streamlining Emergency Response: A K-Adaptable model and a Column-and-Constraint-Generation Algorithm
3c ThA174 - 511D - Innovative RO Techniques in Transportation and Location Science			Chair: IHSAN YANIKOGLU
OMAR BENNOUNA- Tractable smart predict and optimize in the non-realizable case	IHSAN YANIKOGLU- Robust Lane Covering Problem	BESTE BASCIFTCI- Distributionally Robust Optimization with Multimodal Decision-Dependent Ambiguity Sets	FARZAD AVISHAN- Electric Bus Fleet Scheduling under Travel Time and Energy Consumption Uncertainty
3c ThA277 - 511E - Optimization - Machine Learning Interfaces			Chair: CHAITHANYA BANDI
MIRAI TANAKA- Equivalent reformulation of multistage robust optimization under time-varying exogenous uncertainty with an application to optimal bidding in energy markets	TIM NIEMANN- General Polyhedral Approximation of Two-Stage Robust Linear Programming	DEEKSHA SINHA- Multipurchase behavior: Modeling, estimation, and optimization	CHAITHANYA BANDI- A Robust Approach to Classification and Truth-Validation: Incorporating Human and Large Language Model Decision Making
3c ThB296 - 511D - Distributionally Robust Optimization: Theory Meets Applications			Chair: RUIWEI JIANG
ALIAA ALNAGGAR- Two-Stage Distributionally Robust Optimization for Service Region Design in Crowdsourced Delivery	KANGLIN LIU- Reliable Facility Location with Wasserstein Ambiguity	HOANG NAM NGUYEN- Distributionally Robust Optimization of Probability of Exceedance	NOAM GOLDBERG- Smooth Uncertainty Sets: Dependence of Uncertain Parameters via a Simple Polyhedral Set
3c FA133 - 511D - Robust Optimization meets other domains			Chair: AHMADREZA MARANDI
JUDITH BRUGMAN- Robust Appointment Scheduling with General Convex Uncertainty Sets	JOHANNES THÜRAUF- Connections and Reformulations between Robust and Bilevel Optimization	AYOUB FOUSSOUL- LP-based approximations for disjoint bilinear and two-stage adjustable robust optimization.	AHMADREZA MARANDI- Quadratic Optimization Through the Lens of Adjustable Robust Optimization
3c FA140 - 511E - Robust Optimization for Operations Problems			Chair: ZHENYU HU
MIGUEL CARRASCO- Robust topology optimization using the Bernstein approximation	KAI PAN- Optimized Dimensionality Reduction for Moment-based Distributionally Robust Optimization	CHUN CHENG- Robust Workforce Management with Crowdsourced Delivery	ZHENYU HU- Robust Capacity Planning with General Upgrading
3c FB104 - 511D - Data-driven Robust Optimization			Chair: JUSTIN DUMOCHELLE
BO LIN- Conformal Inverse Optimization	IRINA WANG- Learning Decision-Focused Uncertainty Sets in Robust Optimization	JUSTIN DUMOCHELLE- Neur2RO: Neural Two-Stage Robust Optimization	
3c FB122 - 511E - Optimization with Marginals and Moments			Chair: KARTHIK NATARAJAN
KARTHIK NATARAJAN- Distributionally robust optimization through the lens of submodularity	ARJUN RAMACHANDRA- When submodularity meets pairwise independence	HAILONG SUN- Convex Optimization for Bundle Size Pricing Problem	HE SIMAI- Progresses in Modeling and solving distributional robust optimization problems
3c FB184 - 511A - Exact methods for robust nonlinear optimization			Chair: DICK DEN HERTOOG
DANIQUE DE MOOR- RPT for robust nonlinear optimization problems	DICK DEN HERTOOG- Reformulation-Perspectification Technique for problems with conic uncertainty sets	THORORIS KOUKOUVINOS- An exact method for a class of robust nonlinear optimization problems	
Stream 3D: MULTI-STAGE STOCHASTIC PROGRAMMING AND REINFORCEMENT LEARNING			
3d MB167 - 513A - Reinforcement Learning in Industry			Chair: DANIEL JIANG
ETHAN CHE- Adaptive Experimentation at Scale: A Computational Framework	ALEC KOPPEL- Exploration Incentives in Model-Based Reinforcement Learning	ZHIWEI (TONY) QIN- RL and ADP for Rideshare Operations	ZHEQING ZHU- Deep Reinforcement Learning for Recommender Systems and Beyond

3d MC783 - 515B - Multi-stage Stochastic Programming and Reinforcement learning 1 Chair: ELIZABETH GLISTA			
JEAN PAUL WATSON- Recent Developments in the MPI-SPPY Solver Library for Scalable Stochastic Programming	ZITONG WANG- Resampling Stochastic Gradient Descent Cheaply for Efficient Uncertainty Quantification	ELIZABETH GLISTA- Stochastic Capacity Expansion Planning considering Power Flow and Unit Commitment	
3d TA53 - 515B - Approximation of Large Scale Structured Markov Decision Processes in Operations Applications Chair: SELVAPRABU NADARAJAH			
ANDRE AUGUSTO CIRE- Combinatorial Fluid Models for Routing	SELVAPRABU NADARAJAH- Structure Aware Lagrangian Relaxations for Weakly Coupled Markov Decision Processes	DAVID BROWN- Unit Commitment without Commitment: A Dynamic Framework for Managing an Integrated Energy System Under Uncertainty	DANIEL JIANG- Faster Approximate Dynamic Programming by Freezing Slow States
3d TB22 - 515B - Stochastic Dual Dynamic Programming : theory and applications Chair: JOAQUIM DIAS GARCIA			
BERNARDO FREITAS PAULO DA COSTA- Duality, upper bounds, and policy quality in Multistage Stochastic Programming	GRANI A HANASUSANTO- Data-Driven Stochastic Dual Dynamic Programming: Performance Guarantees and Regularization Schemes	ANDY PHILPOTT- Multistage stochastic inventory optimization with contracting	JOAQUIM DIAS GARCIA- A Multicut Approach to Compute Upper Bounds for Risk-Averse SDDP
3d TC940 - 515B - Multi-stage Stochastic Programming and Reinforcement learning 1 Chair: GIULIA PEDRIELLI			
QIUSHI HAN- A Low-Rank ADMM Splitting Approach for Semidefinite Programming	GIULIA PEDRIELLI- Scaling Bayesian optimization with Multi-agent Rollout		
3d WA148 - 515B - Advances in Approximate Dynamic Programming with Energy Applications Chair: STEPHAN MEISEL			
ALESSIO TRIVELLA- Network dual reoptimization policies and bounds for managing energy real options	WOUTER VAN HEESWIJK- Reinforcement Learning for Energy-Efficient Job Shop Scheduling with Large Structured Action Spaces	NEGAR SOHEILI- Stochastic Approximation for Upper Bound Optimization in Markov Decision Processes	STEPHAN MEISEL- Reinforcement Learning or Policy Sampling: An Online Coupling Policy with an Energy Storage Application
3d WB176 - 515B - Advances in Stochastic Programming Chair: GUANGHUI LAN			
GUANGHUI LAN- Optimal L-p Risk Minimization	DOMINIC KEEHAN- Epi-Consistency of Multistage Stochastic Optimization Problems	YIFAN HU- Three-stage Stochastic Programming is As Easy As Classical Stochastic Optimization	ALEXANDER SHAPIRO- Risk averse and distributionally robust modeling of multistage stochastic programming
3d WC328 - 515B - Advances in Multi-stage Stochastic Programming and Reinforcement learning Chair: YIZHE HUANG			
YAN LI- Policy Mirror Descent Inherently Explores Action Space	XIN CHEN- Nonconvex Landscape of Policy Optimization for A Class of Finite Horizon Markov Decision Processes and Applications in Operations Models	YIZHE HUANG- Adversarially Robust Optimal Control with Causal Transport Distance	
3d ThA246 - 515B - Statistical Learning, Optimization and Stochastic Programming for Reinforcement Learning Chair: MING YIN			
YINGRU LI- HyperAgent: Advancing Scalable Exploration through Fast Uncertainty Estimation in RL	DONGRUO ZHOU- Variance-Dependent Regret Bounds for Non-stationary Linear Bandits	HAQUE ISHFAQ- Provable and Practical: Efficient Exploration in Reinforcement Learning via Langevin Monte Carlo	MING YIN- Rethinking of Deep RL: Learning the Target Network in Function Space
3d ThB186 - 515B - Efficient RL: Adaptive Experimentation and Estimation Chair: IMON BANERJEE			
DAVID WOODRUFF- Stochastic Programming Without Assuming a Probability Distribution	SOMAYEH MOAZENI- Bayesian Reinforcement Learning with Limited Evaluations and Nonlinear Objectives	ETHAN CHE- Planning Adaptive Experiments with Model-Predictive Control	IMON BANERJEE- Sample efficient estimation of the transition kernels of controlled Markov chains

3d ThC255 - 515B - Adaptive stochastic optimization with applications in energy system operations and transition			
Chair: ANDY SUN			
PETER BARKLEY- Design of Optimized Proximal Splitting Algorithms	PANAGIOTIS ANDRIANESIS- A New Risk Assessment and Management Paradigm in Electricity Markets	THOMAS LEE- Bundle Method Approaches for Capacity Expansion Planning with Large Network and Temporal Scales	
3d FA182 - 515B - Advances at the intersection of reinforcement learning and stochastic optimization			
Chair: SAEED GHADIMI			
DANIEL RANGEL MARTINEZ- A Recurrent Reinforcement Learning Strategy for Online Scheduling of a State Task Network Under Uncertainty	WARREN POWELL- A Unified Framework for Sequential Decision Problems: From Dynamic Programming/RL to Stochastic Programming	SWATI GUPTA- Inducing Group Fairness in Real-World Pipelines through Uncertainty Protections	ANDRZEJ RUSZCZYNSKI- A Functional Model Method for Nonconvex Nonsmooth Conditional Stochastic Optimization
3d FB109 - 515B - Advances in multi-stage stochastic programming			
Chair: BERNARDO PAGNONCELLI			
YONGJIA SONG- Multi-Stage Stochastic Programming for Integrated Hurricane Evacuation and Logistics Planning	HAMED RAHIMIAN- Multistage Chance-Constrained Programming	MERVE BODUR- Interpretable Vaccine Administration and Inventory Replenishment Policies via Smooth-in-Expectation Decision Rules	BERNARDO PAGNONCELLI- Solving multistage equilibrium problems: the Krusell-Smith model
Stream 3E: DATA-DRIVEN OPTIMIZATION			
3e MB254 - 514B - Data-driven optimization in urban systems			
Chair: MANXI WU			
HARIPRIYA PULYASSARY- Network Flow Problems with Electric Vehicles	JULIA YAN- Pricing shared rides	ARON BRENNER- A Generative Learning Approach for Data-Driven Robust Optimization	MANXI WU- Information design for spatial resource allocation
3e MC177 - 514C - Learning for Inventory			
Chair: YAQI XIE			
SEAN SINCLAIR- Hindsight Learning for MDPs with Exogenous Inputs	YAQI XIE- VC Theory for Inventory Policies	ZIJIN ZHANG- Costly Data Collection in the Newsvendor Problem	JIASHUO JIANG- Learning to Order for Inventory Systems with Lost Sales and Uncertain Supplies
3e TA134 - 514C - Robust or Go Bust: Tales in Distributionally Robust and Data-Driven Optimization			
Chair: WOLFRAM WIESEMANN			
DAN IANCU- Distributionally Robust Linear Quadratic Control	WOLFRAM WIESEMANN- Randomized Assortment Optimization	DANIEL KUHN- Wasserstein Distributionally Robust Optimization with Heterogeneous Data Sources	CAGIL KOCYIGIT- Learning Optimal and Fair Policies for Online Allocation of Scarce Societal Resources from Data Collected in Deployment
3e TA158 - 512C - Statistical and Computational Aspects of Robust Estimation			
Chair: SOROOSH SHAFIEE			
SOROOSH SHAFIEE- Outlier-Robust Wasserstein DRO	JONATHAN LI- Coherent Wasserstein Distributionally Robust Optimization	TAOLI ZHENG- Universal Gradient Descent Ascent Method for Smooth Minimax Optimization	CYNTHIA RUSH- The out-of-sample prediction error of the square-root lasso and related estimators
3e TB73 - 514C - Data-driven Methods for Contextual Stochastic Optimization			
Chair: MENG QI			
RARES CRISTIAN- Robust End-to-End Learning under Endogenous Uncertainty	MINGLONG ZHOU- Robust Actionable Prescriptive Analytics	MENG QI- Integrated Conditional Estimation-Optimization	PAUL GRIGAS- Conformal Contextual Optimization with a Smart Predict-then-Optimize Method
3e TC198 - 514C - Learning and Verification of Continuous Optimizers			
Chair: BRANDON AMOS			
RAJIV SAMBHARYA- Data-Driven Performance Guarantees for Classical and Learned Optimizers	VINIT RANJAN- Verification of First-Order Methods for Parametric Quadratic Optimization	BRANDON AMOS- Amortized optimization for optimal transport and LLM attacks	

3e WA941 - 514C - Data-driven optimization 1			Chair: SAEID ABBASIPARIZI
ADRIAN RIEKERT- Stochastic gradient methods with adaptive learning rates	PRIYANK AGRAWAL- A Tractable Online Learning Algorithm for the Multinomial Logit Contextual Bandit	HANSHENG JIANG- Smoothness-Adaptive Dynamic Pricing with Nonparametric Demand Learning	SAEID ABBASIPARIZI- A Benders Decomposition Algorithm for the Overland Search and Rescue Problem Considering Search Area Selection

3e [WB82](#) - 514C - **Data-driven Optimization Under Uncertainty: Theory and Applications** Chair: KARMEL SHEHADEH

HAO HAO- Robust Design of Critical Supply Chains	YILING ZHANG- Integer Programming Approaches for Distributionally Robust Chance Constraints with Adjustable Risks	KARMEL SHEHADEH- On the Trade-off Between Distributional Belief and Ambiguity: Conservatism, Bias, and Asymptotic Properties	XIAN YU- Residuals-Based Contextual Distributionally Robust Optimization with Decision-Dependent Uncertainty
--	---	--	--

3e [WC942](#) - 514C - **Data-driven optimization 2** Chair: MOHAMMAD REZA BELBASI

NURSEN AYDIN- Differentially Private Linear Optimization for Multi-Party Resource Sharing	JAMES KOTARY- Metric Learning to Accelerate Convergence of Operator Splitting Methods	MOHAMMAD REZA BELBASI- It's All in the Mix: Wasserstein Machine Learning with Mixed Features	
---	---	--	--

3e [ThB194](#) - 514C - **Learning and Calibration in Data-Driven Optimization** Chair: TIANYU WANG

MELVYN SIM- Service Oriented Considerate Routing: Data, Predictions and Robust Decisions	VISHAL GUPTA- Directional Gradients and Surrogate Losses in the Predict-then-Optimize Framework	CHUNLIN SUN- Decoupling Learning and Decision-Making: Breaking the $O(\sqrt{T})$ Barrier in Online Resource Allocation with First-Order Methods	TIANYU WANG- Optimizer's Information Criterion: Dissecting and Removing Bias in Data-Driven Optimization
--	---	---	--

3e [FA234](#) - 514C - **Data-driven decision making in high-stakes domains: notes on fairness and interpretability** Chair: SINA AGHAEI

LESIA SEMENOVA- Is There an Accuracy-Simplicity Tradeoff in High-Stakes Decision Domains?	SINA AGHAEI- Learning Interpretable and Fair Housing Allocation Policy for Individuals Experiencing Homelessness in Los Angeles	JOHANN GAEBLER- The Measure and Mismeasure of Fairness	SINA BAHARLOUEI- f-FERM: Robust and Fair Empirical Risk Minimization via f-divergences
---	---	--	--

3e [FA259](#) - 512B - **New models and applications of data-driven optimization** Chair: HODA BIDKHORI

MATEUS WAGA- Robust Portfolio Optimization Meets Arbitrage Pricing Theory	CAGIN URU- Data-Driven Robust Sequential Search	HODA BIDKHORI- Missing data in data-driven optimization	TRISTAN POLLNER- Pivotal Sampling for Online Bayesian Matching
---	---	---	--

3e [FB141](#) - 514C - **Advances in data-driven decision making** Chair: VELIBOR MISIC

NATHAN JUSTIN- Learning Optimal Classification Trees Robust to Distribution Shifts	YA CHI CHU- Randomized Nyström Preconditioned Interior Point-Proximal Method of Multipliers (IP-PMM)	LONG HE- On the Equivalence and Performance of Distributionally Robust Optimization and Robust Satisficing Models in OM Applications	VELIBOR MISIC- Randomized policy optimization for high-dimensional optimal stopping
--	--	--	---

CLUSTER 4: APPLICATIONS AND COMPUTATION

Stream 4A: NETWORKS AND TELECOM

4a [TB943](#) - 512G - **Networks and Telecom 1** Chair: CESAR A URIBE

CHARLES NOURRY- Budgeted spanning trees and polyhedra	CESAR A URIBE- On Graphs with Finite-Time Consensus and Their Use in Gradient Tracking		
---	--	--	--

4a [WB944](#) - 512G - **Networks and Telecom 2** Chair: SACHIN JAYASWAL

OLIVIER BÉLANGER- Online Convex Optimization for On-Board Routing in High-Throughput Satellites	ANTOINE BERNARD- Application placement simulations over HEAVEN platform	SACHIN JAYASWAL- Fiber-to-the-Home Passive Optical Access Network Design: A New Formulation and Valid Inequalities	
---	---	--	--

4a WC212 - 514A - Telecommunication Network Optimization			Chair: DIMITRI PAPADIMITRIOU
BERNARD FORTZ- Robust optimization for the Segment Routing Traffic Engineering Problem	PEGAH ALIZADEH- Joint Optimization of inter- and intra-frequency events in Multi-Carrier Scenarios with UE Mobility in Radio Networks	DIMITRI PAPADIMITRIOU- Stochastic Lagrangian-based method and application to multi-constrained network design	
Stream 4B: TRANSPORTATION AND LOGISTICS			
4b MB190 - 511E - Network Design for Transportation Planning II			Chair: TEODOR GABRIEL CRAINIC
JULIE KIENZLE- The Inter-modal Railroad Blocking and Railcar Fleet-Management Planning Problem	TEODOR GABRIEL CRAINIC- Multi-layer network design	AMIRMOHAMMAD FATHOLLAHIFARD- The Integrated Multi-Tier Hierarchical Hub Location and Scheduled Service Network Design Problem	EHSAN MIRZAEI- Profit maximizing network design problem with location decisions considering incompatible commodities and binary demand selection
4b MC202 - 516E - Dynamic Facility and Network Design			Chair: SANJAY DOMINIK JENA
PETER SCHÜTZ- Solving a stochastic multi-horizon facility location problem with capacity expansion	GABRIEL HOMSI- A slope scaling heuristic for the multiperiod strategic planning of carbon capture and sequestration value chains	SANJAY DOMINIK JENA- Cumulative Customer Demand in Facility Location and Network Design	
4b TA192 - 516E - Network Design for Transportation Planning III			Chair: WALTER REI
TOM VAN WOENSEL- Design for integrated Freight and Passenger Transportation: Some recent advances	JULIA LANGE- A Construction Matheuristic for Two-Tier Synchronized City Logistics	FRANCESCO CONTU- A multi-commodity location-network design problem with vehicle selection in City Logistics	WALTER REI- The Two-Echelon Multicommodity Location-Routing Problem with Stochastic and Correlated Demands
4b TA274 - 512G - Large Scale Optimisation in Transportation			Chair: FRANCOIS SOUMIS
NADIA RASOULI- A Branch-and-Price Algorithm for Multiple Depot Bus scheduling Problem in Public Transit Systems	FRÉDÉRIC QUESNEL- Considering regularity in the aircrew pairing problem	PHILIPPE RACETTE- Accelerated insight into the crew rostering problem with machine learning methods	MOHAND RYAD CHELBANI- Solving the Crew Split Problem: Innovative Team Formation Strategies for the Crew Pairing Problem
4b TB166 - 513D - Network Design for Transportation Planning I			Chair: TEODOR GABRIEL CRAINIC
ALI ROUHANI- Benders Decomposition for Two-layer Network Design with Capacity Decisions	SAEED DEHGHANI- From Tactical to Operational Planning: Implementation Policies for Consolidation-based Freight Transportation Systems	BITA PAYAMI SHABESTARI- Service Network Design with Uncertainty on Water Levels for Intermodal River Transport	NADIA TAHMASBI- Tactical Planning in Multi-Stakeholder Freight Transportation Systems under Uncertainty
4b TB945 - 511F - Transportation and logistics 1			Chair: JEAN FRANÇOIS CÔTÉ
ZAHRA ANSARILARI- A Comprehensive Stochastic Programming Model for Transfer Synchronization in Transit Networks	LARS HELLEMO- Flexible and modular energy systems modelling with JuMP; A case study from the Arctic	NIMA MORADI- Electric Vehicle Routing Problem with Robots, Parcel Lockers, and Time Windows	JEAN FRANÇOIS CÔTÉ- Service level requirements in real-life-sized bike sharing systems
4b TB946 - 516E - Transportation and logistics 2			Chair: SHEN PENG
TAHA VAROL- Neural Network Estimators for Optimal Tour Lengths of Traveling Salesperson Problem Instances with Arbitrary Node Distributions	LAURA KOLCHEVA- On-line Stochastic Optimization for Real-Time Transfer Synchronization in Public Transportation Networks	SHEN PENG- Distributionally robust facility location problem under facility disruption, uncertain demand and cost	
4b WA947 - 516E - Transportation and logistics 3			Chair: DAVID SCHINDL
MICHEL GENDREAU- A two-stage stochastic programming model with recourse for a Production Routing Problem with uncertain availability of vehicles	BAHMAN BORNAY- Generalized Dial-A-Ride Problem on Road Networks with Time-dependent Travel Times	ROSEMARIE SANTA- A Stochastic Approach for Mobile Clinic Deployment Planning	DAVID SCHINDL- A column generation approach for the routing of electricity technicians

4b WB147 - 516E - Large-scale optimization in action: methods and applications				Chair: ALEXANDRE JACQUILLAT
SEBASTIEN MARTIN- A Better Match for Everyone, Reinforcement Learning for Ridesharing	ALEXANDRE JACQUILLAT- Branch-and-price for prescriptive contagion analytics	ARTHUR DELARUE- Heterogeneous Treatment Effects In Matching Marketplaces Under Interference: From Estimation to Decision	SIMON THOMÄ- Modeling and Optimizing Large-Scale Inventory Routing in East Africa: A Data-Driven Optimization Approach	
4b ThA948 - 511F - Transportation and logistics 4				Chair: ALI KERMANI
FAIZ HAMID- Valid Inequalities for the Vehicle Routing Problem with Drones	ALI KERMANI- A Matheuristic for the Stochastic Production Routing Problem with Adaptive Routing and Service Levels			
4b ThA949 - 516E - Transportation and logistics 5				Chair: CAGLA DURSUNOGLU
MOHAMMAD KARIMI- An exact branch-and-price-and-cut algorithm for a practical dial-a-ride problem	MALTE BILLEN- A Column-Generation Algorithm for the Inter-city Electric Coach Scheduling Problem	CAGLA DURSUNOGLU- A unifying framework for selective routing problems		
4b ThB155 - 511F - Models and Algorithms in Last-mile Delivery				Chair: OKAN ARSLAN
HANG DONG- EV delivery, charging and parking planning method under uncertain arrival times at logistic depots	XIN WANG- Optimizing Ultra-Fast Delivery Networks and Service Guarantees Under Uncertainty	MILAD ELYASI- Technician Routing and Scheduling under Uncertainty	OKAN ARSLAN- A contextual framework for learning routing experiences in last-mile delivery	
4b ThB950 - 516E - Transportation and logistics 6				Chair: ALEJANDRO TORIELLO
SEYOUNG OH- An integer optimization approach for the graph-free multi-agent pathfinding problem	ASEFEH HASSANI GOODARZI- Optimizing Containerized Multimodal Transport toward Sustainability and resilience	MICHAEL LEVIN- Charging Station Location and Fleet Sizing for Shared Autonomous Electric Vehicles using Benders' Decomposition	ALEJANDRO TORIELLO- Optimizing the Trade-Off Between Batching and Waiting: Subadditive Dispatching	
4b ThC951 - 516E - Transportation and logistics 7				Chair: ALI AKBAR SADAT ASL
STEFAN ROPKE- Dive-and-cut-and-price heuristics for vehicle routing problems	BIENVENU BAMBI- Automatic model decomposition in Hexaly Optimizer	ALI AKBAR SADAT ASL- An Efficient Real-time Multi-modal Single-leg Transportation System: Balancing Perspectives on User Needs, Operator Efficiency, and Sustainability		
4b FA193 - 516E - Network Design for Transportation Planning IV				Chair: MARIA ELENA BRUNI
GUIDO PERBOLI- Stochastic Variable Cost and Size Bin Packing with Capacity Loss and Supplier Selection problem	GIACOMO LANZA- Continuous-Time Service Network Design with Stochastic Travel Times	MAHDI MOSTAJABDAVEH- Door-to-Door Multimodal Long-Distance Freight Transportation Planning	MARIA ELENA BRUNI- Scheduled Service Network Design with Packing Considerations	
4b FB164 - 516E - Humanitarian Logistics				Chair: VITTORIO NICOLETTA
HANI ZBIB- A Mutual Catastrophe Insurance Framework for Horizontal Collaboration in Prepositioning Strategic Reserves	VITTORIO NICOLETTA- Mitigating fire risk towards critical and residential structures near a high ignition area using Critical Node Detection	ALBORZ HASSANZADEH- The Value of Demand Prediction for Improved Food Security	MAEDEH SHARBAF- A bilevel optimization approach for shelter network design and evacuation planning problem: An application to flood preparedness in Haiti	
4b FB172 - 511F - Demand-based and stochastic urban transportation planning				Chair: LÉA RICARD
GIANMARCO ANDREANA- Competing on Emissions Charges	ÖYKÜ NAZ ATTILA- Hub Transportation Problem with Chance Constrained Due Dates	JIA TENG YIN- Collaborative Optimization of Rolling Stock Allocation and Timetable Coordination in a Multi-Modal Rail Network: MILP Formulation and Decomposition-Based Algorithm	JIAQI LIANG- A Reinforcement Learning Approach for Dynamic Rebalancing in Bike-Sharing Systems	

Stream 4C: SCHEDULING AND MANUFACTURING			
4c TB952 - 512E - Scheduling and manufacturing 1			Chair: MAURÍCIO GONÇALVES
ZOE FORNIER- Growing a stochastic subtree for mixed integer linear programming under uncertainty	LÉA RICARD- Stochastic optimization of bus schedules	MAURÍCIO GONÇALVES- Integrated lot sizing and blending problems under demand uncertainty	
4c TC953 - 513D - Scheduling and manufacturing 2			Chair: KENZA ALIOUI
ASHLEY PEPER- Scheduling Cybersecurity Mitigations to Delay Attacker Projects	MATHIEU GERVAIS DUBÉ- Scheduling ISMP 2024	KENZA ALIOUI- Open shop scheduling problem with reentrance, delays, and no-wait constraints	
4c ThA954 - 512G - Scheduling and manufacturing 3			Chair: PAUL PATIENCE
ALIREZA ALAVI- Fair Assortment Optimization for an Online Fashion Retailer in the Presence of Size Substitution Effect	XINYI GUO- A Nested Logic-Based Benders Decomposition for the Multi-Shop Car Resequencing Problem with A Painted Body Storage	PAUL PATIENCE- Optimizing a cellular solid compression problem via the NOMAD blackbox optimizer	
4c FA955 - 512E - Scheduling and manufacturing 4			Chair: FARIN RASTGAR AMINI
TAZU NOMOTO- Daily production scheduling method for mixed-model assembly line and different workload models	DEBORA PRETTI RONCONI- Flexible job shop scheduling problem with sequencing flexibility: constructive heuristics and local searches considering position-based learning effect	HARCÈNAGE DANSOU- Solving an integrated lot sizing and preventive maintenance planning for a single machine	FARIN RASTGAR AMINI- Data Mining-Driven Shift Enumeration for Accelerating the Solution of Large-Scale Personnel Scheduling Problems
Stream 4D: ENERGY AND ENVIRONMENT			
4d MB119 - 516E - Optimization models for electric vehicle charging problems 1			Chair: MIGUEL ANJOS
NAGISA SUGISHITA- Strengths of the Formulations of the Fuel Refuelling Location Problem	ISMAIL SEVIM- A Constrained Traffic Assignment Problem Model for Congestion Analysis of Electric Vehicle Fast Chargers	GAËL GUILLOT- Optimal Electric Vehicle Charging with Dynamic Pricing, Customer Preferences and Power Peak Reduction	TOMMASO SCHETTINI- Optimizing Electric Vehicle Charger Locations for Ride-hailing Services through Discrete Simulation-based Optimization
4d MC337 - 513C - Optimal transition of energy systems/power grids 2			Chair: MICHEL DENAULT
OLIVIER BAHN- Modeling of a Hydrogen Economy Using Integrated Assessment Approaches	SARA GHABOULIAN ZARE- Robust Pathways to H2 Adoption: Leveraging ETEM-YUL and ETEM-CAN Frameworks		
4d TA117 - 513D - Advances in generation, transmission and distribution modelling			Chair: ALEXANDRE BESNER
ALEXANDRE BESNER- RALPH - Solving the unit commitment problem under hydraulic and transient stability constraints	SALVADOR PINEDA- Tight Big-Ms for Optimal Transmission Switching	SANJA BAJIC- Application of operational research in Hydro Quebec distribution system planning	YVES ALAIN MBEUTCHA- Optimizing physical and financial assets using SDDP
4d TA956 - 513C - Energy and Environment 1			Chair: HONGYU ZHANG
MINSEOK RYU- Differentially Private Distributed Optimization with Constraints	PAUL DOMMEL- Uniform Convergence of Kernel Estimators	HOJIN JUNG- An Interval-based Stochastic Unit Commitment Model Under Demand Uncertainty	HONGYU ZHANG- Integrated energy system planning under short-term and long-term uncertainty: Modelling and algorithms
4d TB292 - 513C - Optimization models for electric vehicle charging problems 2			Chair: MIGUEL ANJOS
NATHALIA WOLF- Improving Cloud Service Sustainability with Bilevel Optimization under Demand Uncertainty	KYRA ANN D'IGNAZIO- The Electric Bus Rostering and Charging Scheduling Problem with Uncertain Energy Consumption: a Two-Stage Stochastic Programming Approach	MIGUEL ANJOS- Integrated Location, Sizing and Pricing for Electric Vehicle Charging Stations	

4d TC142 - 513C - Optimal transition of energy systems/power grids 1			Chair: MICHEL DENAULT
MADELEINE SEATTLE- Multi-value benefits of transmission expansion in Western Canada	FLORIAN MITJANA- Reinforcement learning and large system operations	SOROOSH SABAEI- Generation Planning of Renewables Using Stochastic Dual Dynamic Programming: A Case Study on Quebec	
4d WA206 - 513D - Equilibrium modeling in energy markets			Chair: RYAN ENT
STEVEN GABRIEL- Equilibrium Modeling in Natural Gas Markets: A Theoretical Analysis and a Case Study for Brazil, A Method for Solving Gas Market and Water Market Equilibrium Problems	IBRAHIM ABADA- When market incompleteness is preferable to market power. Insights from power markets.	EDDIE ANDERSON- Connection queues and shared connection costs	RYAN ENT- Analysis of ISO New England's Energy Imbalance Reserve Using Mixed Complementarity Problems
4d WA957 - 513C - Energy and Environment 2			Chair: SEYYEDREZA MADANI
REHLICH LEA- Using structural properties to optimize operating strategies of district heating networks under nonlinear constraints	MAAIKE ELGERSMA- Obtaining the Convex Hull Formulation for Optimal Investment and Operation of Energy Storage Systems Including Reserves	LUIS LOPEZ- A Convex Model Predictive Control Optimization Model of Active Distribution Network Flexibility Regions	SEYYEDREZA MADANI- Smart Grids, Smart Pricing: Employing Reinforcement Learning for Prosumer-Responsive Critical Peak Pricing
4d WB958 - 513C - Energy and Environment 3			Chair: JORGE WESTON
ELISA CORBEAN- Robust Optimal Design of Renewable Energy Systems within an Energy Storage Process using Iron as Energy Carrier	ALEJANDRO ANGULO- A Robust Data-Driven Approach to Characterize Operational Flexibility Zones of Active Distribution Networks	JORGE WESTON- A Heuristic Adversarial Approach for the Two-Stage Robust AC OPF Problem	
4d WC171 - 513C - Optimization of Energy Communities			Chair: KRISTEN SCHELL
ALBERT SOLÀ VILALTA- Optimal Participation of Energy Communities in Electricity Markets under Uncertainty. A Multi-Stage Stochastic Programming Approach	LORELEY SEPPO- Hierarchical optimization of cooperative distributed energy resource aggregations in power distribution systems	KRISTEN SCHELL- Interdependent Network Optimization Modelling for Energy Network Design to Support Post Disaster Services	
4d ThA46 - 513C - Optimization Methods for Renewable Power Systems			Chair: ANTOINE LESAGE LANDRY
LESIA MITRIDATI- Privacy-Preserving Decentralized Local Electricity Markets	JUAN M. MORALES- Data Valuation in Data-driven Stochastic Optimization	JOHN W. SIMPSON PORCO- Hierarchical Feedback-Based Optimization of Distributed Energy Resources in Distribution Grids	ANTOINE LESAGE LANDRY- Online Interior-point Methods for Time-Varying Equality-constrained Optimization
4d ThA204 - 513D - Mobile and static battery systems operations and planning			Chair: ANDY SUN
ALGHUMAYJAN SAUD- Decision-Focused Learning for Energy Storage Control and Behavior Prediction	DIRK LAUINGER- Arbitrage and Regulation with Electricity Storage	MATTHEW BRUN- Optimal Facility Planning for Electric Vehicle Battery Recycling	ROBIN LEGAULT- Optimal Planning and Scheduling for a Large-Scale Electric Bus Fleet Transition Problem
4d ThB33 - 513C - Learning from Optimization in Energy			Chair: DAVI VALLADÃO
NATHAN CHO- On the Coupling Between Charging Scheduling of Electrified Bus Fleets and Grid Pricing	ANDRE FELIPE DE CARVALHO RAMOS- Advancing Renewable Energy Simulation: Integrating Time Variant Hidden Markov Models with SDDP for Enhanced Accuracy and Real-World Application	BRUNO FANZERES- Task-Based Prescriptive Trees for Two-Stage Linear Decision-Making Problems: Reformulations, Heuristic Strategies, and Applications	DAVI VALLADÃO- Contextually Robust Optimization: End-to-End Learning in Power System Applications
4d ThB118 - 511E - Optimization of Gas Networks			Chair: LARS SCHEWE
JULIA GRÜBEL- Adjustable Robust Nonlinear Network Design under Demand Uncertainties	PASCAL BÖRNER- Mixture of Gases on Networks	ROWAN TURNER- A tailored interior point method for fast optimization on gas networks	LARS SCHEWE- Optimization methods for the analysis of gas markets

4d ThC261 - 513C - Research at Hydro-Québec: Black-box Optimization			Chair: MIGUEL DIAGO
VINCENT MAI- Black box optimization with virtual sensors for turbine deformations during startup events	MIGUEL DIAGO- Risk-Based Optimization of Periodic Maintenance Policies for Electrical Transmission Equipment	ALEXIA MARCHAND- Comparison of optimization methods for hyperparameter tuning of a deep learning forecasting model	
4d FB959 - 513C - Energy and Environment 4			Chair: HANDE BENSON
TRULS FLATBERG- Efficient hydrogen infrastructure design using JuMP and convex approximations	HANDE BENSON- Applications of Optimization in Land Conservation		
Stream 4E: HEALTHCARE AND LIFE SCIENCES			
4e MB219 - 512E - Applications of Stochastic optimization in Healthcare			Chair: ANTOINE LEGRAIN
AMELIA LANE- A Network-Flow Approach to Increasing Inter-Facility Transfers for Nova Scotia Health Long-Term Care	SU LI- Quantifying the benefits of customized vaccination strategies: A network-based optimization approach	CAMILLE PINÇON- Multi-objective scheduling of chemotherapy drug preparation	ALI JAFARI- Real-time demand-driven inventory management in a hospital pharmacy
4e MC283 - 514A - Oncology and Clinical Decision Support			Chair: ADRIAN MANUEL REIMANN
CARL JULIUS MARTENSEN- Data-Driven Alternatives for Feedback Mechanisms in Acute Myeloid Leukaemia	KHANH DINH- Inference for parameters of selection and copy number aberrations from DNA-sequencing data	ADRIAN MANUEL REIMANN- Optimal Leukemia Treatment	
Stream 4F: FINANCE AND SOCIAL SCIENCES			
4f MB960 - 512G - Finance and social sciences 1			Chair: JOÃO VITOR PAMPLONA
JÖRGEN BLOMVAL- Optimal hedging of the interest rate swap book	FRANCIS HUOT CHANTAL- Predicting Limit Order Book Prices: A Machine Learning Approach	ROY KWON- Contextual Portfolio Optimization: Some Recent Results	JOÃO VITOR PAMPLONA- Fair Mixed Effects Support Vector Machine
Stream 4G: HIGH PERFORMANCE IMPLEMENTATION AND QUANTUM COMPUTING			
4g MB162 - 512B - Quantum Computing and Optimization I			Chair: BRANDON AUGUSTINO
YI ZHANG- Symplectic Extra-gradient Type Method for Solving General Non-monotone Inclusion Problem	STEFAN WOERNER- Quantum Algorithms for Multi-objective Optimization	GIACOMO NANNICINI- Solving the semidefinite relaxation of QUBOs in matrix multiplication time, and faster with a quantum computer	YUCHEN LU- Quantum Langevin Dynamics for Optimization
4g MB209 - 512A - A glimpse of computational optimization at Google I			Chair: HOSSEIN BATENI
MATTHEW FAHRBACH- Practical Performance Guarantees for Pipelined DNN Inference	MOHAMMAD HAJIAGHAYI- Contextual Nearest Neighbor Search in Sublinear Time	RAJESH JAYARAM- Multi-Vector Retrieval: The New Frontier of Similarity Search	HOSSEIN BATENI- Attention-based Model Structure Optimization
4g MC163 - 512E - Quantum Computing and Optimization II			Chair: BRANDON AUGUSTINO
BRANDON AUGUSTINO- The quantum central path algorithm	CHENYI ZHANG- Quantum speedups for stochastic optimization	ALICE ZHENG- Online learning of a panoply of quantum objects	
4g TC282 - 512E - Quantum Computing and Optimization III Chair: RODOLFO ALEXANDER QUINTERO OSPINA			
SVEN LEYFFER- Binary control pulse optimization for quantum systems	DAVID BERNAL NEIRA- Hybrid Quantum Branch-and-Bound Method for Quadratic Unconstrained Binary Optimization	AKSHAY BANSAL- Breaking barriers in two-party quantum cryptography via stochastic semidefinite programming	
4g WA961 - 512E - High performance implementation and quantum computing 1 Chair: PRAKASH GAWAS			
SHOJI SHIMIZU- Steepest-edge simplex algorithms for quadratic programming	PRAKASH GAWAS- An imitation-based learning approach using DAgger for the Casual Employee Call Timing Problem		

4g WB267 - 512E - Parallel solvers				Chair: YUJI SHINANO
DANIEL REHFELDT- QuBowI: An exact solver for MaxCut and QUBO problems	YUJI SHINANO- Progress on highly parallel ensemble solvers	KOICHI FUJII- Solving Large-Scale Quadratic Assignment Problems by a Parallelized Lagrangian-DNN-Based Branch-and-Bound	RYO KUROIWA- Parallel General-Purpose Dynamic Programming Solvers for Combinatorial Optimization	
4g WC191 - 512E - Optimization Modeling				Chair: CARL LAIRD
BRUNO VIEIRA- From Algorithms to Applications: developing and deploying analytical and optimization models with FICO® Xpress Insight	RUTH MISENER- OMLT: Solving inverse problems over trained graph neural networks	CARL LAIRD- Democratization of High-performance Computing Approaches for Optimization		
4g ThA181 - 512E - A glimpse of computational optimization at Google II				Chair: JUAN PABLO VIELMA
HOSSEIN ESFANDIARI- Active Learning via Density-based Space Transformation	DANIEL ALEJANDRO ESPINOZA GONZALEZ- An exponential penalization, network decomposition algorithm for large LPs	JOEY HUCHETTE- Capacity management via network flows with nonlinear transformations at the nodes	CHRISTIAN TJANDRAATMADJA- Network-aware scheduling of large ML workloads	
4g FB165 - 512E - A glimpse of computational optimization at Google III				Chair: JUAN PABLO VIELMA
TOBY DAVIES- ViolationLS: Constraint-Based Local Search in CP-SAT	ROSS ANDERSON- MathOpt: Solver independent modeling and execution independent solving in Google's OR-Tools	DANIEL DUQUE- Google's Operations Research API	JUAN PABLO VIELMA- Duality and decomposition in OR-Tools' MathOpt	
Stream 4H: GAME THEORY				
4h MB962 - 512C - Game Theory 1				Chair: AHMADREZA TAVASOLI
AURELIO OLIVEIRA- A new approach for solving linear programming problems with dense columns	NENA BATENBURG- A New Strategy for Solving Cournot Games	GABRIEL CUNHA- Mitigation mechanisms for externalities in intertemporal multistage markets: The case of cascaded hydropower plants	AHMADREZA TAVASOLI- Counterparty credit risk under netting agreements: A dynamic game Interpretation	
4h ThA256 - 512F - Applications of Nash Equilibrium Problems				Chair: ALEXANDRA SCHWARTZ
YUEH LIN- A Generalized Voting Game for Categorical Network Choices	RAUL GARCIA- Strategy Investments in Zero-Sum Games	DEVANSH JALOTA- When Simple is Near-Optimal in Security Games	ALEXANDRA SCHWARTZ- Dynamic pricing and strategic retailers in the energy sector: A multi-leader-follower approach	
4h FA330 - 512F - Computing Equilibria of Games				Chair: YICHUAN DING
FRANCESCO MORRI- Learning in Multi-Leader Single-Follower Stackelberg Games	ZI YI HUANG- Nash Equilibrium of Code-Share Mechanism under Airline Frequency Competition Game with Second Order Cone Programming	CHUANGYIN DANG- Computing Quasi-Proper Equilibria: From a Characterization to a Differentiable Path-Following Method	YICHUAN DING- Tight Bounds for the Price of Fairness	
Stream 4I: INDUSTRIAL APPLICATIONS				
4i TA276 - 512F - Large Scale Supply-Chain Network Design I				Chair: SEMIH ATAKAN
MATHIEU DAHAN- Relay-Hub Network Design for Consolidation Planning Under Demand Variability	JOCHEN KOENEMANN- Dynamic discretization discovery for region-based networks	MADISON VAN DYK- Fast Combinatorial Algorithms for Efficient Sortation	SEMIH ATAKAN- Designing Optimization Software for Use Case Scalability	
4i WA273 - 512G - Large Scale Supply-Chain Network Design II				Chair: ARASH HADDADAN
EUGENE ZAK- Multi-Modeling Approach to Real-World Optimization Problems	ARASH HADDADAN- A decomposition approach for large scale service network design	DAVID MILDEBRATH- Practical stochastic network design for long-term network planning	R RAVI- Velocity-aware inventory placement	
4i ThB963 - 512G - Industrial applications 1				Chair: THIBAUT BARBIER
JOHN CHINNECK- Faster Infeasibility Analysis for Linear Programs	DAVID LANDAY- Novel Branch & Bound Tree Management for Fast Convergence of MIPs on Parallel Computers	SIAMAK NADERI- Joint Transportation, Markdown, and Clearance Decisions at a Fast-Fashion Retailer	THIBAUT BARBIER- Optimizing Ticket Pricing with Expedia's AI ?	

4i FA288 - 512G - Optimization and Validation of Execution Models at Amazon Scale Chair: QI CHEN			
ZIHAO LI- Ideal inventory placement with demand spillover	QI CHEN- Validating Network Planning via Customer Centric Baselines	AKSHIT KUMAR- Dynamic Resource Allocation: Algorithmic Design Principles and Spectrum of Achievable Performances	OZLEM BILGINER- Solving Large Scale Linear Problems with Lagrangian Decomposition

CLUSTER 5: SPECIAL PRESENTATIONS

5c TA280 - 511F - A.W. Tucker Prize Chair: FRANK E. CURTIS

TO BE ANNOUNCED - A.W. Tucker Prize	TBA- Tucker Prize Finalists		
-------------------------------------	-----------------------------	--	--

5c WA316 - 511F - In memory of Jean-Louis Goffin 1: Recent advances in large scale optimization Chair: SAMIR ELHEDHLI

LUIS ZULUAGA- Lifting nonnegativity over general semialgebraic sets to nonnegativity over simple sets	RENATO MONTEIRO- A low-rank method for large-scale semidefinite programs	JIMING PENG- Rectified relaxation: A trustworthy and interpretable deep learning approach to linear optimization with complementary constraints	
---	--	---	--

5c WB317 - 511F - In memory of Jean-Louis Goffin 2: Analytic Centre Cutting Plane Methods: The contributions of J-L Goffin's work Chair: TAMÁS TERLAKY

TAMÁS TERLAKY- From the Relaxation Method to Chubanov's Algorithm, from ACCPM to Tall LPs.	LEVENT TUNCEL- A personal view of primal-dual interior-point algorithms for large-scale optimization and some recent developments	SAMIR ELHEDHLI- Analytic Centres and Mixed Integer Programming	MOHAMED EL FASSI- Local center cutting plane methods framework
--	---	--	--

5c WC81 - 511F - In memory of Jean-Louis Goffin 3: A tribute to J-L Goffin's academic influence Chair: FATMA GZARA

FATMA GZARA- The value of structure and decomposition in large scale optimization	MOHAMMAD OSKOOROUCHI- Beyond Academia: Remembering the Human Essence of a Great Scholar	MICHEL DENAULT- The power transition: optimization meets policy	FARANAK MOKHTARIAN- An Analytic Center self-Concordant Cut method for the Convex Feasibility Problem
---	---	---	--

5c ThC978 - 511F - Special sessions 1 Chair: TBA

TBA- Paul Y. Tseng Memorial Lecture			
-------------------------------------	--	--	--

Parallel sessions - ordered chronologically

Warning: The schedule is subject to changes. Please consult the [online version](#).

OC - 210A-F OPENING CEREMONY		MONDAY 9:00-10:30
CHARLES AUDET – MIGUEL ANJOS		
CLUSTER 1: CONTINUOUS OPTIMIZATION		MONDAY 14:00-16:00
1a MB39 - 510C - Advances in numerical methods for large scale nonlinear optimization and applications I		Chair: MARGHERITA PORCELLI
PAULO J. S. SILVA – MARGHERITA PORCELLI – FILIPPO ZANETTI – JACEK GONDZIO –		
1a MB11 - 516D - Advances in Large-Scale Nonlinear Optimization I		Chair: ALBERT BERAHAS
MERT GURBUZBALABAN – ALBERT BERAHAS – QI WANG – JELENA DIAKONIKOLAS –		
1a MB62 - 516A - Development of nonlinear optimization algorithms and applications		Chair: BIN GAO
ERMIN WEI – NICOLAS LE ROUX – XIN LIU – BIN GAO –		
1c MB71 - 510B - Inertial and stochastic algorithms		Chair: SEBASTIAN KASSING
SEBASTIAN KASSING – JUNWEN QIU – SAM DAVANLOO TAJBAKHSH – KUANGYU DING –		
1c MB92 - 511A - Douglas–Rachford Awesome Cones, Great Job! 1		Chair: LUIZ RAFAEL SANTOS
SCOTT BOIVIN LINDSTROM – VERA ROSHCHINA – FELIPE ATENAS – LUIZ RAFAEL SANTOS –		
1d MB6 - 510A - Convex Approaches for Quadratically Constrained Quadratic Programs		Chair: SAMUEL BURER
JIANZHE ZHEN – SAMUEL BURER – BOSHI YANG – ALEX WANG –		
1d MB29 - 510D - Recent advances in conic optimization - Part 1		Chair: MAKOTO YAMASHITA
KEIYU NOSAKA – AKIKO YOSHISE – SENA SAFARINA – CHARLES NAMCHASIRI –		
1e MB312 - 513D - Convex analysis and beyond		Chair: ADRIAN LEWIS
SHUBHANKAR BHATT – MICHAEL ORLITZKY – SEDI BARTZ – ADRIAN LEWIS –		
1g MB302 - 515A - Challenging DFO and SBO Applications 1		Chair: YVES LUCET
MORGANE MENZ – FILIPPO MARINI – LUCA BOISNEAULT – YVES LUCET –		
1h MB88 - 515C - Nonsmooth PDE Constrained Optimization 1		Chair: HARBIR ANTIL
HARBIR ANTIL – ANTON SCHIELA – DANIEL WACHSMUTH – TUOMO VALKONEN –		
CLUSTER 2: DISCRETE OPTIMIZATION		MONDAY 14:00-16:00
2a MB157 - 511F - Vehicle Routing Problems under Uncertainty		Chair: PEDRO MUNARI
MATHEUS JUN OTA – ALBERT SCHROTENBOER – RAFAEL AJUDARTE DE CAMPOS – PEDRO MUNARI –		
2b MB12 - 514C - Binary Quadratic Programming		Chair: SVEN MALLACH
LUCAS LÉTOCART – BORZOU ROSTAMI – LUCAS WADDELL – SVEN MALLACH –		
2c MB921 - 511B - Combinatorial Optimization and Graph Theory 1		Chair: TASUKU SOMA
STEPHAN MARNACH – VLADIMIR KOLMOGOROV – JENNY SEGSCHEIDER – TASUKU SOMA –		
2c MB223 - 511C - Complexity and algorithmic aspects of structured families of graphs 1		Chair: KATHIE CAMERON
AURÉLIE LAGOUTTE – MORITZ MÜHLENTHALER – CÉSAR HERNÁNDEZ CRUZ – KARTHEKEYAN CHANDRASEKARAN –		
2d MB248 - 513F - Distance Optimization in Data Science		Chair: EMILIO CARRIZOSA
JASONE RAMIREZ AYERBE – NURIA GÓMEZ VARGAS – THOMAS HALSKOV – EMILIO CARRIZOSA –		
2e MB153 - 513E - Stochastic Combinatorial Optimization I		Chair: NICOLE MEGOW
MARC UETZ – SHARAT IBRAHIMPUR – THOMAS KESSELHEIM – NICOLE MEGOW –		
2e MB239 - 514A - Network design		Chair: NEIL OLVER
LAURA VARGAS KOCH – ELLIS HERSHKOWITZ – NATHAN KLEIN – RHEA JAIN –		
2f MB344 - 513C - Modeling Software		Chair: DAN STEFFY
GREGOR HENDEL – DAN STEFFY – MILES LUBIN – CAN LI –		
2f MB932 - 513B - Computational Discrete and Integer Optimization 1		Chair: FRED GARDI
FREDERIC KRONER – SABRINA AMMANN – LISA MARIE MANKE – FRED GARDI –		

CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		MONDAY 14:00-16:00
3a MB107 - 515B - Multistage Stochastic Programs	Chair: ZHIPING CHEN	
ZHIPING CHEN – RAIAN NOUFEL LEFGOUM – LEONARD MACLEAN – JIA LIU –		
3b MB287 - 512H - Recent modeling language extensions: Stochastics and beyond	Chair: MICHAEL FERRIS	
FERRIS		
STEVEN DIRKSE – OLIVIER HUBER – ROBERT FOURER – MICHAEL FERRIS –		
3c MB108 - 511D - Applications of Robust and Distributionally Robust Optimization	Chair: ZHEN YANG	
ZHI CHEN – IAN YIHANG ZHU – LI CHEN – ZHEN YANG –		
3d MB167 - 513A - Reinforcement Learning in Industry	Chair: DANIEL JIANG	
ETHAN CHE – ALEC KOPPEL – ZHIWEI (TONY) QIN – ZHEQING ZHU –		
3e MB254 - 514B - Data-driven optimization in urban systems	Chair: MANXI WU	
HARIPRIYA PULYASSARY – JULIA YAN – ARON BRENNER – MANXI WU –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		MONDAY 14:00-16:00
4b MB190 - 511E - Network Design for Transportation Planning II	Chair: TEODOR GABRIEL CRAINIC	
JULIE KIENZLE – TEODOR GABRIEL CRAINIC – AMIRMOHAMMAD FATHOLLAHIFARD – EHSAN MIRZAEI –		
4d MB119 - 516E - Optimization models for electric vehicle charging problems 1	Chair: MIGUEL ANJOS	
NAGISA SUGISHITA – ISMAIL SEVIM – GAËL GUILLOT – TOMMASO SCHETTINI –		
4e MB219 - 512E - Applications of Stochastic optimization in Healthcare	Chair: ANTOINE LEGRAIN	
AMELIA LANE – SU LI – CAMILLE PINÇON – ALI JAFARI –		
4f MB960 - 512G - Finance and social sciences 1	Chair: JOÃO VITOR PAMPLONA	
JÖRGEN BLOMVALL – FRANCIS HUOT CHANTAL – ROY KWON – JOÃO VITOR PAMPLONA –		
4g MB209 - 512A - A glimpse of computational optimization at Google I	Chair: HOSSEIN BATENI	
MATTHEW FAHRBACH – MOHAMMAD HAJIAGHAYI – RAJESH JAYARAM – HOSSEIN BATENI –		
4g MB162 - 512B - Quantum Computing and Optimization I	Chair: BRANDON AUGUSTINO	
YI ZHANG – STEFAN WOERNER – GIACOMO NANNICINI – YUCHEN LU –		
4h MB962 - 512C - Game Theory 1	Chair: AHMADREZA TAVASOLI	
AURELIO OLIVEIRA – NENA BATENBURG – GABRIEL CUNHA – AHMADREZA TAVASOLI –		
CLUSTER 1: CONTINUOUS OPTIMIZATION		MONDAY 16:20-17:50
1a MC55 - 510C - From First- to High-order Methodologies: Recent Advances and Trends I	Chair: ARYAN MOKHTARI	
ARYAN MOKHTARI		
ERFAN YAZDANDOOST HAMEDANI – ARYAN MOKHTARI – MICHAL DEREZINSKI –		
1a MC902 - 516D - Nonlinear Programming 1	Chair: XINWEI LIU	
RACHAEL TAPPENDEN – PRASHANT KHANDURI – XINWEI LIU –		
1b MC14 - 513B - Global Optimization	Chair: YU HONG DAI	
ZHOU WEI – ZHILONG DONG – JOHN WARWICKER –		
1c MC44 - 510B - Design of New Nonsmooth Optimization Algorithms	Chair: BENJAMIN GRIMMER	
QIHANG LIN – ZHICHAO JIA – AHMET ALACAOGLU –		
1d MC908 - 510A - Semi-Definite Programming 1	Chair: ELLEN HIDEKI FUKUDA	
NATHAN KELLER – BENOÎT LEGAT – ELLEN HIDEKI FUKUDA –		

CLUSTER 2: DISCRETE OPTIMIZATION		MONDAY 16:20-17:50
2b	MC286 - 513F - MINLP Applications	Chair: AKSHAY GUPTE
OLHA SILINA – JUSTO PUERTO – DAVID REY –		
2c	MC250 - 511B - Graph Theory	Chair: FRÉDÉRIC MEUNIER
HADRIEN MÉLOT – ALAIN HERTZ – GRÉGOIRE BEAUDOIRE –		
2d	MC298 - 513A - Classification and regression via mathematical optimization	Chair: DOLORES ROMERO MORALES
ROMERO MORALES		
ANTONIO CONSOLO – MARIA EDUARDA PINHEIRO – DOLORES ROMERO MORALES –		
2f	MC339 - 513E - Extensions of Integer Programming	Chair: MARIANNA DE SANTIS
MARIANNA DE SANTIS – MARK TURNER – PEDRO MACIEL XAVIER –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		MONDAY 16:20-17:50
3a	MC934 - 515A - Continuous Stochastic Programming 1	Chair: JUNYU ZHANG
DI ZHANG – VINCENT LECLÈRE – JUNYU ZHANG –		
3b	MC57 - 512A - Modeling and Optimization under Uncertainty	Chair: SUVRAJEET SEN
DAVID MORTON – SUVRAJEET SEN – SIMGE KUCUKYAVUZ –		
3d	MC783 - 515B - Multi-stage Stochastic Programming and Reinforcement learning 1	Chair: ELIZABETH GLISTA
ELIZABETH GLISTA		
JEAN PAUL WATSON – ZITONG WANG – ELIZABETH GLISTA –		
3e	MC177 - 514C - Learning for Inventory	Chair: YAQI XIE
SEAN SINCLAIR – YAQI XIE – ZIJIN ZHANG – JIASHUO JIANG –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		MONDAY 16:20-17:50
4b	MC202 - 516E - Dynamic Facility and Network Design	Chair: SANJAY DOMINIK JENA
PETER SCHÜTZ – GABRIEL HOMSI – SANJAY DOMINIK JENA –		
4d	MC337 - 513C - Optimal transition of energy systems/power grids 2	Chair: MICHEL DENAULT
OLIVIER BAHN – SARA GHABOULIAN ZARE –		
4e	MC283 - 514A - Oncology and Clinical Decision Support	Chair: ADRIAN MANUEL REIMANN
CARL JULIUS MARTENSEN – KHANH DINH – ADRIAN MANUEL REIMANN –		
4g	MC163 - 512E - Quantum Computing and Optimization II	Chair: BRANDON AUGUSTINO
BRANDON AUGUSTINO – CHENYI ZHANG – ALICE ZHENG –		

–

CLUSTER 1: CONTINUOUS OPTIMIZATION		TUESDAY 08:30-10:30
1a	TA40 - 510C - Advances in Large-Scale Nonlinear Optimization II	Chair: RAGHU BOLLAPRAGADA
STEPHEN WRIGHT – RAGHU BOLLAPRAGADA – QUOC TRAN DINH – DANIEL ROBINSON –		
1a	TA31 - 516D - Advances in Nonconvex Optimization	Chair: SALAR FATTAHI
XUXING CHEN – RYAN CORY WRIGHT – WEIJUN XIE – SALAR FATTAHI –		
1a	TA59 - 516A - Advances in numerical methods for large scale nonlinear optimization and applications II	Chair: ERNESTO BIRGIN
FLORENTIN GOYENS – JEREMY RIEUSSEC – ERNESTO BIRGIN – ILARIA TROMBINI –		
1b	TA324 - 512H - Global Optimization at work	Chair: SERGIY BUTENKO
ROBERT GOTTLIEB – DMYTRO MATSYPURA – JOURDAIN LAMPERSKI – NIKOLAOS V. SAHINIDIS –		
1c	TA65 - 510B - Nonlinear optimization problems	Chair: WEI BIAN
WEI BIAN – WENJING LI – LEI YANG – HAO WANG –		
1c	TA112 - 511A - Advances in Variational Analysis and Nonsmooth Optimization I	Chair: JOHANNES ROYSET
YING CUI – JOHANNES ROYSET – JONG SHI PANG – DANIEL MIMOUNI –		
1d	TA138 - 510A - Moment, Polynomial and Tensor Optimization - Part 1	Chair: SUHAN ZHONG
CEDRIC JOSZ – KE YE – JIYOUNG CHOI – XINDONG TANG –		
1e	TA313 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 1	Chair: RADU IOAN BOT
HEINZ BAUSCHKE – SHAMBHAVI SINGH – RUSSELL LUKE – RADU IOAN BOT –		
1f	TA130 - 512E - Fast randomized algorithms for distributed optimization	Chair: LAURENT CONDAT
CORRADO COPPOLA – JIAXIN XIE – MING YAN – LAURENT CONDAT –		
1g	TA306 - 515C - Large-scale and Parallel DFO	Chair: YOUSSEF DIOUANE
PIERRE YVES BOUCHET – HAITIAN LI – JEFFREY LARSON – YOUSSEF DIOUANE –		
1h	TA90 - 512B - Nonsmooth PDE Constrained Optimization 2	Chair: ANNIKA SCHIEMANN
MICHAEL HINTERMÜLLER – PAUL MANNS – MARVIN SEVERITT – ANNIKA SCHIEMANN –		
CLUSTER 2: DISCRETE OPTIMIZATION		TUESDAY 08:30-10:30
2a	TA188 - 513B - Network-flow based MIP formulations	Chair: FRANÇOIS CLAUTIAUX
MIKE HEWITT – SEAN LO – FRANÇOIS CLAUTIAUX – LUIS MARQUES –		
2b	TA152 - 513F - Solving ACOPF problems	Chair: DANIEL BIENSTOCK
HASSAN HIJAZI – ANDY SUN – PASCAL VAN HENTENRYCK – MATIAS VILLAGRA –		
2c	TA263 - 511C - Graph Coloring	Chair: ECKHARD STEFFEN
JONAS ALKER – ISAAK H. WOLF – ALEXANDER CLOW – ECKHARD STEFFEN –		
2c	TA922 - 511B - Combinatorial Optimization and Graph Theory 2	Chair: ADOLFO ESCOBEDO
MALEK MOUHOU B – JOSE WALTEROS – EDWARD DUC HIEN NGUYEN – ADOLFO ESCOBEDO –		
2d	TA297 - 513A - Optimization for Safe and Private Machine Learning	Chair: THIBAUT VIDAL
FRAUKE LIERS – ROBERT BASSETT – THIBAUT VIDAL – MATTEO FRANCOBALDI –		
2e	TA229 - 512A - Algorithms for Pricing and Assortment Optimization	Chair: OMAR EL HOUSNI
RAJAN UDWANI – MIKA SUMIDA – SANJAY MEHROTRA – OMAR EL HOUSNI –		
2e	TA930 - 514B - Approximation and Online Algorithms 1	Chair: ALENA OTTO
DANISH KASHAEV – SANDER BORST – ANDREAS KARRENBAUER – ALENA OTTO –		
2f	TA269 - 513E - Solver Software I	Chair: TIMO BERTHOLD
TIMO BERTHOLD – MICHAEL WINKLER – JULIAN HALL – NILS CHRISTIAN KEMPKE –		

CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		TUESDAY 08:30-10:30
3a TA115 - 515A - Advances in First-Order Methods for Continuous and Stochastic Optimization Chair: TIANJIAO LI		
LIWEI JIANG – TIANJIAO LI – JIAMING LIANG – JIMMY ZHANG –		
3c TA99 - 510D - Robust Optimization and Machine Learning 1		Chair: ARAS SELVI
SERGIO CAMELO – MAN CHUNG YUE – ARAS SELVI – ABHILASH CHENREDDY –		
3c TA170 - 511E - Recent Advances in the Theory of Stochastic and Robust Optimization		Chair: HAOMING SHEN
ANGELOS GEORGHIOU – LOUIS CHEN – LI XIAOBO –		
3c TA938 - 511D - Robust and Distributionally Robust Optimization 1		Chair: MATTEO PETRIS
VITOR LUIZ PINTO DE PINA FERREIRA – ZUKUI LI – MATTEO PETRIS –		
3d TA53 - 515B - Approximation of Large Scale Structured Markov Decision Processes in Operations Applications		Chair: SELVAPRABU NADARAJAH
ANDRE AUGUSTO CIRE – SELVAPRABU NADARAJAH – DAVID BROWN – DANIEL JIANG –		
3e TA158 - 512C - Statistical and Computational Aspects of Robust Estimation		Chair: SOROOSH SHAFIEE
SOROOSH SHAFIEE – JONATHAN LI – TAOLI ZHENG – CYNTHIA RUSH –		
3e TA134 - 514C - Robust or Go Bust: Tales in Distributionally Robust and Data-Driven Optimization		Chair: WOLFRAM WIESEMANN
DAN IANCU – WOLFRAM WIESEMANN – DANIEL KUHN – CAGIL KOCYIGIT –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		TUESDAY 08:30-10:30
4b TA274 - 512G - Large Scale Optimisation in Transportation		Chair: FRANCOIS SOUMIS
NADIA RASOULI – FRÉDÉRIC QUESNEL – PHILIPPE RACETTE – MOHAND RYAD CHELBANI –		
4b TA192 - 516E - Network Design for Transportation Planning III		Chair: WALTER REI
TOM VAN WOENSEL – JULIA LANGE – FRANCESCO CONTU – WALTER REI –		
4d TA117 - 513D - Advances in generation, transmission and distribution modelling		Chair: ALEXANDRE BESNER
ALEXANDRE BESNER – SALVADOR PINEDA – SANJA BAJIC – YVES ALAIN MBEUTCHA –		
4d TA956 - 513C - Energy and Environment 1		Chair: HONGYU ZHANG
MINSEOK RYU – PAUL DOMMEL – HOJIN JUNG – HONGYU ZHANG –		
4i TA276 - 512F - Large Scale Supply-Chain Network Design I		Chair: SEMIH ATAKAN
MATHIEU DAHAN – JOCHEN KOENEMANN – MADISON VAN DYK – SEMIH ATAKAN –		
CLUSTER 5: SPECIAL PRESENTATIONS		TUESDAY 08:30-10:30
5c TA280 - 511F - A.W. Tucker Prize		Chair: FRANK E. CURTIS
To BE ANNOUNCED – –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		TUESDAY 14:00-16:00
1a TB67 - 510C - Numerical Optimization for Machine Learning 1	Chair: FREDERIK KUNSTNER	
AUDREY BAZERGHI – WU LIN – JONATHAN LAVINGTON – NICOLAS LOIZOU –		
1a TB136 - 511C - First-order Methods and Large-scale Optimization I	Chair: LEXIAO LAI	
ZIQI QIN – SIYU KONG – BOHAO MA – FENG YI LIAO –		
1a TB69 - 516A - From First- to High-order Methodologies: Recent Advances and Trends II	Chair: YANG LIU	
KATE WENQI ZHU – FADI HAMAD – SHUYAO LI – HANFENG ZENG –		
1a TB60 - 516D - Advances in numerical methods for large scale nonlinear optimization and applications III	Chair: STEPHANIA BELLAVIA	
WANPING DONG – STEPHANIA BELLAVIA – GIORGIA FRANCHINI – LUKA RUTESIC –		
1b TB5 - 512H - Polynomial and Conic Optimization	Chair: IMMANUEL BOMZE	
MARKUS GABL – IMMANUEL BOMZE – QUEENIE HUANG –		
1c TB207 - 511A - Large scale continuous optimization and first-order methods and related topics	Chair: HAIHAO LU	
ZIKAI XIONG – HAIHAO LU – TIANYUN TANG – DAVID APPELEGATE –		
1c TB63 - 510B - Nonsmooth Stochastic Optimization and Applications	Chair: HAILIN SUN	
INBAR SEROUSSI – JIE JIANG – HAILIN SUN –		
1d TB66 - 510A - SDP and Combinatorial Optimization	Chair: NATHAN KRISLOCK	
MICHAEL SMITH – TIMOTEJ HRGA – ANGELIKA WIEGELE – IAN SUGRUE –		
1d TB87 - 510D - Semidefinite Programming and Convex Algebraic Geometry	Chair: REKHA THOMAS	
TAKASHI TSUCHIYA – EITAN LEVIN – SUHAN ZHONG – MITCHELL HARRIS –		
1e TB311 - 514A - New insights into first- and second-order variational analysis	Chair: TIM HOHEISEL	
STEPHEN SIMONS – TING KEI PONG – HELMUT GFREERER – TIM HOHEISEL –		
1e TB314 - 514B - Algorithms for optimization, monotone inclusions and variational inequalities 2	Chair: ZI XU	
JIAYUAN WU – VIKTOR PAVLOVIC – EWA BEDNARCZUK – ZI XU –		
1f TB131 - 512C - Beyond Standard Assumptions in Optimization: Structured Non-Convexity, Non-Monotonicity, and Non-Smoothness	Chair: EDUARD GORBUNOV	
YI ZHOU – ILYAS FATKHULLIN – BRECHT EVENS – CONG FANG –		
1g TB322 - 515C - Recent Progresses in DFO 1	Chair: GIAMPAOLO LIUZZI	
GABRIEL JARRY BOLDUC – ZAIKUN ZHANG – SAMUEL MENDOZA – GIAMPAOLO LIUZZI –		
1h TB91 - 512B - Nonsmooth PDE Constrained Optimization 3	Chair: HENRIK WYSCHKA	
DREW KOURI – NICOLAI SIMON – ROHIT KHANDLWAL – HENRIK WYSCHKA –		
CLUSTER 2: DISCRETE OPTIMIZATION		TUESDAY 14:00-16:00
2a TB915 - 513B - Mixed Integer Linear Programming 1	Chair: QUOC TUNG LE	
NILS HASSAN QUTTINEH – BÁRBARA RODRIGUES – QUOC TUNG LE –		
2b TB195 - 513F - Optimization over integers, cones and polynomials	Chair: LUZE XU	
SHIXUAN ZHANG – KURT ANSTREICHER – LUZE XU – SHAONING HAN –		
2c TB86 - 511B - Spectral Graph Theory and Optimization	Chair: RENATA DEL VECCHIO	
HERMIE MONTERDE – LEONARDO LIMA – CYBELE VINAGRE – SHAHLA NASSERASR –		
2d TB199 - 513A - Learning to Select	Chair: VASSILIS DIGALAKIS	
GEORGE ORFANIDES – VASSILIS DIGALAKIS – MICHAEL LI – STEFANOS POULIDIS –		
2f TB340 - 513E - Solver Software II	Chair: PHILIPP CHRISTOPHEL	
SOFIANE TANJI – SVEN WIESE – JULIEN DARLAY – PHILIPP CHRISTOPHEL –		

CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		TUESDAY 14:00-16:00
3a	TB111 - 515A - PDE-constrained optimization under uncertainty	Chair: DANIEL WALTER
CAROLINE GEIERSBACH – JOHANNES MILZ – MATTHIAS HEINKENSCHLOSS MATTHIAS – DANIEL WALTER –		
3b	TB187 - 512A - Modeling and Algorithms for Stochastic Discrete Optimization	Chair: HAOXIANG YANG
ZHUO ZHANG – XINGYU LIN – CHENG GUO – LUHAO ZHANG –		
3c	TB100 - 511D - Robust Optimization and Machine Learning 2	Chair: MENG MENG LI
YUHAO WANG – CHIN PANG HO – DUO ZHOU – MENG MENG LI –		
3c	TB168 - 511E - Robust Optimization in practice	Chair: AHMADREZA MARANDI
HOURA MAHMOUDZADEH – JUSTIN STARREVELD – PERIKLIS PETRIDIS – MINAKSHI PUNAM MANDAL –		
3d	TB22 - 515B - Stochastic Dual Dynamic Programming : theory and applications	Chair: JOAQUIM DIAS GARCIA
BERNARDO FREITAS PAULO DA COSTA – GRANI A HANASUSANTO – ANDY PHILPOTT – JOAQUIM DIAS GARCIA –		
3e	TB73 - 514C - Data-driven Methods for Contextual Stochastic Optimization	Chair: MENG QI
RARES CRISTIAN – MINGLONG ZHOU – MENG QI – PAUL GRIGAS –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		TUESDAY 14:00-16:00
4a	TB943 - 512G - Networks and Telecom 1	Chair: CESAR A URIBE
CHARLES NOURRY – CESAR A URIBE –		
4b	TB166 - 513D - Network Design for Transportation Planning I	Chair: TEODOR GABRIEL CRAINIC
ALI ROUHANI – SAEDEH DEGHANI – BITA PAYAMI SHABESTARI – NADIA TAHMASBI –		
4b	TB945 - 511F - Transportation and logistics 1	Chair: JEAN FRANÇOIS CÔTÉ
ZAHRA ANSARILARI – LARS HELLEMO – NIMA MORADI – JEAN FRANÇOIS CÔTÉ –		
4b	TB946 - 516E - Transportation and logistics 2	Chair: SHEN PENG
TAHA VAROL – LAURA KOLCHEVA – SHEN PENG –		
4c	TB952 - 512E - Scheduling and manufacturing 1	Chair: MAURÍCIO GONÇALVES
ZOE FORNIER – LÉA RICARD – MAURÍCIO GONÇALVES –		
4d	TB292 - 513C - Optimization models for electric vehicle charging problems 2	Chair: MIGUEL ANJOS
NATHALIA WOLF – KYRA ANN D'IGNAZIO – MIGUEL ANJOS –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		TUESDAY 16:20-17:50
1d	TC23 - 510A - Semidefinite and Polynomial Optimization	Chair: FRANK VALLENTIN
WILLEM DE MUINCK KEIZER – DAVID GROSS – JEAN B. LASSERRE –		
1d	TC299 - 510D - Semidefinite Programming for Geometric Processing	Chair: AMIT SINGER
ANDREW MCRAE – AMIT SINGER – DAVID PALMER –		
1e	TC911 - 514A - Variational Analysis, Variational Inequalities and Complementarity 1	Chair: JEIN SHAN CHEN
MONIKA SYGA – EDOUARD OYALLON – JEIN SHAN CHEN –		
1g	TC305 - 515C - Multiobjective DFO	Chair: LUDOVIC SALOMON
EVERTON SILVA – LUDOVIC SALOMON – DIMO BROCKHOFF –		
1h	TC320 - 512B - Analysis of Algorithms for PDE Constrained Optimization 1	Chair: STEFAN ULBRICH
STEFAN ULBRICH – JESPER SCHRÖDER – ROBERT J BARALDI –		

CLUSTER 2: DISCRETE OPTIMIZATION		TUESDAY 16:20-17:50
2b TC920 - 513F - Mixed Integer Nonlinear Programming 1	Chair: ANDREA GHEZZI	
EHSAN YADEGARI – HENRI LEFEBVRE – ANDREA GHEZZI –		
2c TC359 - 511B - Advances in Combinatorial Optimization	Chair: TAMON STEPHEN	
STEFFEN BORGWARDT – AMY WIEBE – TAMON STEPHEN –		
2d TC289 - 513A - On Machine Learning for Mathematical Optimization	Chair: ANDREA LODI	
HONGYU CHENG – RICK WILLEMSSEN – ROSARIO MESSANA –		
2e TC358 - 514B - Approximation in Algorithmic Game Theory	Chair: CHAITANYA SWAMY	
POOJA KULKARNI – RIAN NEOGI – CHAITANYA SWAMY –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		TUESDAY 16:20-17:50
3b TC96 - 512A - Theory and Applications of Robust and Stochastic Integer Programming	Chair: WEIJUN XIE	
YONGJIA SONG – JUAN ALBERTO ESTRADA GARCIA – BRADLEY STURT –		
3d TC940 - 515B - Multi-stage Stochastic Programming and Reinforcement learning 1	Chair: GIULIA PEDRIELLI	
QIUSHI HAN – GIULIA PEDRIELLI –		
3e TC198 - 514C - Learning and Verification of Continuous Optimizers	Chair: BRANDON AMOS	
RAJIV SAMBHARYA – VINIT RANJAN – BRANDON AMOS –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		TUESDAY 16:20-17:50
4c TC953 - 513D - Scheduling and manufacturing 2	Chair: KENZA ALIOUI	
ASHLEY PEPER – MATHIEU GERVAIS DUBÉ – KENZA ALIOUI –		
4d TC142 - 513C - Optimal transition of energy systems/power grids 1	Chair: MICHEL DENAULT	
MADELEINE SEATLE – FLORIAN MITJANA – SOROOSH SABAEI –		
4g TC282 - 512E - Quantum Computing and Optimization III	Chair: RODOLFO ALEXANDER QUINTERO OSPINA	
SVEN LEYFFER – DAVID BERNAL NEIRA – AKSHAY BANSAL –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		WEDNESDAY 08:30-10:30
1a WA251 - 510C - Algebraic Methods in Nonlinear Programming	Chair: ALI MOHAMMAD NEZHAD	
THEO DIAMANDIS – MIN TAO – ALI MOHAMMAD NEZHAD – SHENG DING SUN –		
1a WA196 - 516D - Nonlinear stochastic optimization	Chair: MIAOLAN XIE	
YAO JI – MIAOLAN XIE – YUCHEN LOU – PUQIAN WANG –		
1a WA284 - 516A - Theory of constrained nonlinear optimization	Chair: YANGYANG XU	
FARZAD YOUSEFIAN – UDAY SHANBHAG – ZEV WOODSTOCK – YANGYANG XU –		
1b WA905 - 512H - Global Optimization 1	Chair: FABIO TARDELLA	
LEANDRO MAIA – HERMANN SCHICHL – XIANG LI – FABIO TARDELLA –		
1c WA285 - 510B - Efficient algorithms for large-scale sparse optimization problems and applications	Chair: DEFENG SUN	
YING LIN – GUOJUN ZHANG – KAIHUANG CHEN – DEFENG SUN –		
1c WA333 - 511A - Recent Advances in Regularized and Composite Optimization	Chair: ALEXANDER HSU	
AKATSUKI NISHIOKA – MOHAMED LAGHDAF HABIBOULLAH – JAMES BURKE – ALEXANDER HSU –		
1d WA97 - 510A - Semidefinite programming, sums of squares, and quantum information	Chair: HAMZA FAWZI	
SÉBASTIEN DESIGNOLLE – OJAS PAREKH – SUJIT RAO – ADAM BENE WATTS –		
1d WA208 - 510D - Low-rank semidefinite optimization, algorithms, software and applications	Chair: MICHAL KOCVARA	
TIANYUN TANG – HONG MING CHIU – MICHAL KOCVARA – ZHANWANG DENG –		
1e WA338 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 3	Chair: WALAA MOURSI	
BRUNO LOURENÇO – JONATHAN ECKSTEIN – BETHANY CALDWELL – WALAA MOURSI –		
1f WA334 - 512C - Random subspace methods for continuous optimization	Chair: PIERRE LOUIS POIRION	
MICHAEL ULBRICH – AKIKO TAKEDA – BENEDETTO MANCA – PIERRE LOUIS POIRION –		
1g WA309 - 515C - Multifidelity and Bilevel DFO	Chair: FRANCESCO RINALDI	
YUNSOO HA – XAVIER LEBEUF – SAEED GHADIMI – FRANCESCO RINALDI –		
1h WA914 - 512B - Optimal Control, PDE Constrained Optimization, and Multi-level Methods 1	Chair: SEPEHR MOALEMI	
JOHANNES HAUBNER – JAD WEHBEH – DOMINIK H. CEBULLA – SEPEHR MOALEMI –		

CLUSTER 2: DISCRETE OPTIMIZATION		WEDNESDAY 08:30-10:30
2a	WA231 - 511E - Bilevel Optimization with Discrete Decisions (I)	Chair: MARTINA CERULLI
MARTIN LABBE – YASMINE BECK – MARTIN SCHMIDT – FEDERICO BATTISTA –		
2a	WA916 - 513B - Mixed Integer Linear Programming 2	Chair: VANESSA ANN BEDDOE SAWKMIE
YOUNGJOO ROH – FLORIAN ROESEL – ENRICO MALAGUTI – VANESSA ANN BEDDOE SAWKMIE –		
2b	WA8 - 513F - Mixed Integer Quadratic/Conic Optimization	Chair: ATAMTURK ALPER
LINCHUAN WEO – FATMA KILINC KARZAN – ANDRES GOMEZ – SANTANU DEY –		
2c	WA169 - 511C - Complexity and algorithmic aspects of structured families of graphs 2	Chair: CÉSAR HERNÁNDEZ CRUZ
KATHIE CAMERON – JING HUANG – M. GABRIELA ARAUJO PARDO – JAN BOK –		
2c	WA923 - 511B - Combinatorial Optimization and Graph Theory 3	Chair: SATORU IWATA
ANJA FISCHER – FREIJA VAN LENT – FRANCISCO BARAHONA – SATORU IWATA –		
2d	WA927 - 513A - Machine Learning and Discrete Optimization 1	Chair: YASSINE YAAKOUBI
ROLAND WUNDERLING – JULIEN PALLAGE – DEFENG LIU – YASSINE YAAKOUBI –		
2e	WA931 - 514B - Approximation and Online Algorithms 2	Chair: JAROSLAV FOWKES
THOMAS ERLEBACH – ALBERTO LOCATELLI – ALANTHA NEWMAN – JAROSLAV FOWKES –		
2f	WA281 - 513E - New Algorithms for Large-Scale Optimization	Chair: DANIEL ALOISE
ANTONIO M. SUDOSO – ALPHA SALIOU BARRY – FRANCOIS SOUMIS – YOUSSEF EMINE –		
2g	WA243 - 512F - Constraints and Machine Learning	Chair: CHENBO SHI
RICARDO FUKASAWA – RAHUL PATEL – LOUIS MARTIN ROUSSEAU – CHENBO SHI –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		WEDNESDAY 08:30-10:30
3a	WA102 - 515A - Multistage Optimization Programs under Uncertainty	Chair: FRANCESCA MAGGIONI
DARINKA DENTCHEVA		
RUI GAO – ABDEL LISSER – DARINKA DENTCHEVA –		
3b	WA58 - 512A - Stochastic Integer Programming: Models and Algorithms	Chair: JIM LUEDTKE
ERFANEH NIKZAD – JEFF LINDEROTH – WOJIN KIM – EDUARDO MORENO –		
3c	WA106 - 511D - Robust Optimization and Machine Learning 3	Chair: ARAS SELVI
JOSE BLANCHET – JIE WANG – KARTHYEK MURTHY – CAIHUA CHEN –		
3d	WA148 - 515B - Advances in Approximate Dynamic Programming with Energy Applications	Chair: STEPHAN MEISEL
ALESSIO TRIVELLA – WOUTER VAN HEESWIJK – NEGAR SOHEILI – STEPHAN MEISEL –		
3e	WA941 - 514C - Data-driven optimization 1	Chair: SAEID ABBASIPARIZI
ADRIAN RIEKERT – PRIYANK AGRAWAL – HANSHENG JIANG – SAEID ABBASIPARIZI –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		WEDNESDAY 08:30-10:30
4b	WA947 - 516E - Transportation and logistics 3	Chair: DAVID SCHINDL
MICHEL GENDREAU – BAHMAN BORNAY – ROSEMARIE SANTA – DAVID SCHINDL –		
4d	WA957 - 513C - Energy and Environment 2	Chair: SEYYEDREZA MADANI
REHLICH LEA – MAAIKE ELGERSMA – LUIS LOPEZ – SEYYEDREZA MADANI –		
4d	WA206 - 513D - Equilibrium modeling in energy markets	Chair: RYAN ENT
STEVEN GABRIEL – IBRAHIM ABADA – EDDIE ANDERSON – RYAN ENT –		
4g	WA961 - 512E - High performance implementation and quantum computing 1	Chair: PRAKASH GAWAS
SHOJI SHIMIZU – PRAKASH GAWAS –		
4i	WA273 - 512G - Large Scale Supply-Chain Network Design II	Chair: ARASH HADDADAN
EUGENE ZAK – ARASH HADDADAN – DAVID MILDEBRATH – R RAVI –		

CLUSTER 5: SPECIAL PRESENTATIONS		WEDNESDAY 08:30-10:30
5c WA316 - 511F - In memory of Jean-Louis Goffin 1: Recent advances in large scale optimization Chair: SAMIR ELHEDHLI		
LUIS ZULUAGA – RENATO MONTEIRO – JIMING PENG –		
CLUSTER 1: CONTINUOUS OPTIMIZATION		WEDNESDAY 14:00-16:00
1a WB68 - 510C - Numerical Optimization for Machine Learning 2		Chair: MARK SCHMIDT
BETTY SHEA – FREDERIK KUNSTNER – VICTOR SANCHES PORTELLA – LEONARDO GALLI –		
1a WB41 - 516D - Advances in Large-Scale Nonlinear Optimization III		Chair: BAORYU ZHOU
FABIAN BASTIN – ANDREAS WAECHTER – JIAHAO SHI – KATYA SCHEINBERG –		
1a WB120 - 516A - Non-Convex, Non-Smooth, Stochastic, and High-Performance Computing Approaches		Chair: DONGDONG GE
JOACHIM DAHL – QI DENG – TIANYI LIN – DONGDONG GE –		
1b WB350 - 512H - Optimization on manifolds		Chair: DAVID GUTMAN
MIHAI I. FLOREA – YE HE – DAVID GUTMAN – MAN CHO SO –		
1c WB331 - 511A - Advances in First Order Methods		Chair: MARC TEBoulLE
JEROME BOLTE – SHOHAM SABACH – X.Y. HAN – MARC TEBoulLE –		
1c WB906 - 510B - Nonsmooth Optimization 1		Chair: ISABELLA KÄMING
JUYOUNG JEONG – POU GKAKIOTIS SPYRIDON – HONGZHI LIAO – ISABELLA KÄMING –		
1d WB18 - 510A - Semidefinite Programming Performance Estimation		Chair: ETIENNE DE KLERK
BENJAMIN GRIMMER – ETIENNE DE KLERK – HADI ABBASZADEHPEIVASTI – JASON ALTSCHULER –		
1d WB222 - 510D - Algorithms and Applications of Semidefinite Programming		Chair: CIFUENTES DIEGO
ZHUORUI LI – MONSE GUEDES AYALA – ARNESH SUJANANI – GEORGINA HALL –		
1e WB227 - 514A - Algorithms for optimization, monotone inclusions and variational inequalities 4		Chair: JUAN PABLO CONTRERAS
AFROOZ AFROOZ JALILZADEH – SIQI QU – ZHUOYU XIAO – JUAN PABLO CONTRERAS –		
1e WB329 - 515C - Conic optimization		Chair: JAKUB MARECEK
HECTOR RAMIREZ – ALES WODECKI – ANTONIO BELLON – JAKUB MARECEK –		
1f WB49 - 512C - Splitting and Randomized Methods 1		Chair: DEREN HAN
MEHIDDIN AL BAALI – LIPING WANG – ZIYAN LUO – DEREN HAN –		
1h WB321 - 512B - Challenging PDE Constrained Applications		Chair: ANTON BONDAREV
AKWUM ONWUNTA – JANNIK BREITKOPF – MARCEL MAXIMILIAN STEINHARDT – LUKAS KANNENGISSER –		

CLUSTER 2: DISCRETE OPTIMIZATION		WEDNESDAY 14:00-16:00
2a	WB236 - 513D - Bilevel Optimization with Discrete Decisions (II)	Chair: JOHANNES THÜRAUF
LEONARDO LOZANO – ANDREAS HORLÄNDER – NOAH WENINGER – MARGARIDA CARVALHO –		
2a	WB917 - 513B - Mixed Integer Linear Programming 3	Chair: DAVID WARME
JUAN FRANCISCO CORRECHER – JUANJO PEIRÓ – ERIK JANSEN – DAVID WARME –		
2b	WB7 - 513F - Algorithms for Global Mixed Integer Nonlinear Optimization	Chair: PIETRO BELOTTI
IMRE POLIK – JULIO GONZÁLEZ DÍAZ – ROBERT LUCE – PIETRO BELOTTI –		
2c	WB15 - 511B - Discrete Optimization	Chair: MARCIA FAMPA
GABRIEL PONTE – DANIELA LUBKE – YUAN ZHOU – YUSHAN QU –		
2c	WB185 - 511C - Complexity and algorithmic aspects of structured families of graphs 3	Chair: CHÍNH HOÀNG
ANGELE FOLEY – CHÍNH HOÀNG – R SRITHARAN – BEN CAMERON –		
2d	WB230 - 513A - Combinatorial Optimization Augmented Machine Learning and Decision Aware Learning	Chair: AXEL PARMENTIER
CHEIKH AHMED – AXEL PARMENTIER – ESTHER JULIEN – ELLEN VITERCIK –		
2e	WB146 - 514B - Stochastic Combinatorial Optimization II	Chair: VISWANATH NAGARAJAN
YIFAN SUN – ROHAN GHUGE – VISWANATH NAGARAJAN – KEVIN SCHEWIOR –		
2f	WB341 - 513E - Computational Mixed Integer Programming	Chair: PIERRE BONAMI
GIONI MEXI – PIERRE BONAMI – KRUNAL KISHOR PATEL – LUKAS SCHÜRMANN –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		WEDNESDAY 14:00-16:00
3a	WB232 - 515A - Decision-Focused Learning and Bilevel Models for Stochastic Optimization	Chair: TITO HOMEM DE MELLO
RUBEN VAN BEESTEN – FRANCISCO JARA MORONI – VÍCTOR BUCAREY – TITO HOMEM DE MELLO –		
3b	WB178 - 512A - Recent Advances to Solve Discrete Stochastic Problems	Chair: AYSENUR KARAGOZ
HARSHIT KOTHARI – WARD ROMELJNDERS – AYSENUR KARAGOZ – JANOSCH ORTMANN –		
3c	WB125 - 511E - Robust Optimization and Machine Learning 4	Chair: MENGMEG LI
SHIMRIT SHTERN – AHARON BEN TAL – VAMSI POTLURU – SHUO SUN –		
3c	WB939 - 511D - Robust and Distributionally Robust Optimization 2	Chair: FABRICIO OLIVEIRA
EVA LEY – KEVIN MARTIN AIGNER – TIANXIAO WANG – FABRICIO OLIVEIRA –		
3d	WB176 - 515B - Advances in Stochastic Programming	Chair: GUANGHUI LAN
GUANGHUI LAN – DOMINIC KEEHAN – YIFAN HU – ALEXANDER SHAPIRO –		
3e	WB82 - 514C - Data-driven Optimization Under Uncertainty: Theory and Applications	Chair: KARMEL SHEHADEH
HAO HAO – YILING ZHANG – KARMEL SHEHADEH – XIAN YU –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		WEDNESDAY 14:00-16:00
4a	WB944 - 512G - Networks and Telecom 2	Chair: SACHIN JAYASWAL
OLIVIER BÉLANGER – ANTOINE BERNARD – SACHIN JAYASWAL –		
4b	WB147 - 516E - Large-scale optimization in action: methods and applications	Chair: ALEXANDRE JACQUILLAT
SEBASTIEN MARTIN – ALEXANDRE JACQUILLAT – ARTHUR DELARUE – SIMON THOMÄ –		
4d	WB958 - 513C - Energy and Environment 3	Chair: JORGE WESTON
ELISA CORBEAN – ALEJANDRO ANGULO – JORGE WESTON –		
4g	WB267 - 512E - Parallel solvers	Chair: YUJI SHINANO
DANIEL REHFELDT – YUJI SHINANO – KOICHI FUJII – RYO KUROIWA –		
CLUSTER 5: SPECIAL PRESENTATIONS		WEDNESDAY 14:00-16:00
5c	WB317 - 511F - In memory of Jean-Louis Goffin 2: Analytic Centre Cutting Plane Methods: The contributions of J-L Goffin’s work	Chair: TAMÁS TERLAKY
TAMÁS TERLAKY – LEVENT TUNCEL – SAMIR ELHEDHLI – MOHAMED EL FASSI –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		WEDNESDAY 16:20-17:50
1a	WC903 - 516D - Nonlinear Programming 2	Chair: HUYEN TRANG TRAN
YASSINE KAMRI – NIZAR BOUSSELMY – HUYEN TRANG TRAN –		
1c	WC45 - 510B - Large-scale Iterative Optimization	Chair: MATEO DIAZ
ROBERT FREUND – CRISTÓBAL GUZMÁN – NAOMI GRAHAM –		
1d	WC909 - 510A - Semi-Definite Programming 2	Chair: DAVID PAPP
MANUEL RUIZ – DAVID PAPP –		
1g	WC323 - 515C - Recent Progresses in DFO 2	Chair: ANA LUISA CUSTODIO
CHRISTOPHE TRIBES – ANDREA BRILLI – ANA LUISA CUSTODIO –		
CLUSTER 2: DISCRETE OPTIMIZATION		WEDNESDAY 16:20-17:50
2a	WC160 - 513B - Network Design	Chair: NAVNEET VIDYARTHI
PABLO FELIPE OTEÍZA – JAEHEE JEONG – NAVNEET VIDYARTHI –		
2b	WC346 - 513F - MINLP meets Machine Learning	Chair: EMILIANO TRAVERSI
JOSÉ NETO – SHIQIANG ZHANG – EMILIANO TRAVERSI –		
2c	WC353 - 511B - Combinatorial Optimisation	Chair: FABIO FURINI
FRANKLIN DJEUMOU FOMENI – FABIO FURINI – ASVIN GOEL –		
2d	WC203 - 513A - Machine Learning and Combinatorial Optimization	Chair: BARTOLOMEO STELLATO
JIM LUEDTKE – SIMONE MILANESI – BARTOLOMEO STELLATO –		
2f	WC266 - 513E - Certified Optimization	Chair: JONAD PULAJ
JONAD PULAJ – ALEXANDER HOEN – ALDO KIEM –		
2g	WC241 - 514B - CPAIOR at ISMP: Mathematical Programming, Constraint Programming, and Hybrid Methods	Chair: LAURENT MICHEL
JOHN HOOKER – SEYEDE SAEDE HOSSEINI – LAURENT MICHEL –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		WEDNESDAY 16:20-17:50
3a	WC258 - 515A - Theory and Algorithms for Chance Constrained Optimization	Chair: ANIRUDH SUBRAMANYAM
DANIELA BERNHARD – SAI RAKSHITH – YUKI SHIRAI –		
3d	WC328 - 515B - Advances in Multi-stage Stochastic Programming and Reinforcement learning	Chair: YIZHE HUANG
YAN LI – XIN CHEN – YIZHE HUANG –		
3e	WC942 - 514C - Data-driven optimization 2	Chair: MOHAMMAD REZA BELBASI
NURSEN AYDIN – JAMES KOTARY – MOHAMMAD REZA BELBASI –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		WEDNESDAY 16:20-17:50
4a	WC212 - 514A - Telecommunication Network Optimization	Chair: DIMITRI PAPADIMITRIOU
BERNARD FORTZ – PEGAH ALIZADEH – DIMITRI PAPADIMITRIOU –		
4d	WC171 - 513C - Optimization of Energy Communities	Chair: KRISTEN SCHELL
ALBERT SOLÀ VILALTA – LORELEY SEPPO – KRISTEN SCHELL –		
4g	WC191 - 512E - Optimization Modeling	Chair: CARL LAIRD
BRUNO VIEIRA – RUTH MISENER – CARL LAIRD –		
CLUSTER 5: SPECIAL PRESENTATIONS		WEDNESDAY 16:20-17:50
5c	WC81 - 511F - In memory of Jean-Louis Goffin 3: A tribute to J-L Goffin's academic influence	Chair: FATMA GZARA
FATMA GZARA – MOHAMMAD OSKOOROUCHI – MICHEL DENAULT – FARANAK MOKHTARIAN –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		THURSDAY 08:30-10:30
1a ThA42 - 510C - Advances in Large-Scale Nonlinear Optimization IV	Chair: MIKE O'NEILL	
ION NECOARA – ANGELIA NEDICH – SHAGUN GUPTA – MIKE O'NEILL –		
1a ThA135 - 511C - From First- to High-order Methodologies: Recent Advances and Trends III	Chair: ANDRE MILZAREK	
CHUWEN ZHANG – ANDRE MILZAREK – RUICHEN JIANG – NACHUAN XIAO –		
1a ThA74 - 516A - The interplay between optimization, statistics and machine learning I	Chair: LIJUN DING	
LIJUN DING – JIALUN ZHANG – JINWEN YANG – CONG MA –		
1a ThA19 - 516D - Theoretical advances in nonconvex optimization	Chair: HIDEFUMI KAWASAKI	
CARLYLE DELIGNY – MOURAD BAIYOU – MARWA EL HALABI – HIDEFUMI KAWASAKI –		
1b ThA24 - 512H - Recent Advances in Global Optimization	Chair: ANA ROCHA	
ANA ROCHA – ELIGIUS MT HENDRIX – TUSHAR RATHI – FARUQUE HASAN –		
1c ThA30 - 510B - Nonsmooth nonconvex optimization with geometric structure	Chair: XIN LIU	
DOUGLAS SOARES GONCALVES – WEI LIU – QUENTIN REBJOCK –		
1c ThA56 - 511A - Optimization with Projections	Chair: XIAOQI YANG	
XIAOQI YANG – JINGLAI SHEN – XIAOXIA LIU –		
1d ThA214 - 510D - On SDP relaxations for combinatorial optimization problems	Chair: MANISH KRISHAN LAL	
JAN SCHWIDDESEN – NATHAN KRISLOCK – DANIEL BROSCHE – ANNIE RAYMOND –		
1d ThA139 - 510A - Moment, Polynomial and Tensor Optimization - Part 2	Chair: XINDONG TANG	
ZI YANG – JOE KILEEL – ELVIRA MORENO –		
1e ThA354 - 514C - Recent advances in Variational Inequalities and Complementarity	Chair: JEAN PIERRE DUSSAULT	
MATEO DIAZ – EBRAHIM SARABI – JEAN PIERRE DUSSAULT – BAPTISTE PLAQUEVENT JOURDAIN –		
1e ThA912 - 514A - Variational Analysis, Variational Inequalities and Complementarity 2	Chair: TATJANA CHAVDAROVA	
KRZYSZTOF RUTKOWSKI – LINGLING XU – FELIX SCHNEPPE – TATJANA CHAVDAROVA –		
1f ThA128 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization I	Chair: NICOLAS LOIZOU	
MARK SCHMIDT – RYAN D'ORAZIO – COURTNEY PAQUETTE – XIAOYU WANG –		
1g ThA325 - 515C - DFO based on Approximations to Derivatives 1	Chair: STEFAN M. WILD	
SARA SHASHAANI – KILLIAN WOOD – JIAQI WANG – STEFAN M. WILD –		
1h ThA361 - 512B - Analysis of Algorithms for PDE Constrained Optimization 2	Chair: ARSEN HNATIUK	
THOMAS LYNN – ANDREI DRAGANESCU – CHRISTOPH PLATE – ARSEN HNATIUK –		

CLUSTER 2: DISCRETE OPTIMIZATION		THURSDAY 08:30-10:30
2a	ThA237 - 513B - Recent Advances in Combinatorial Optimization	Chair: AHMAD ABDI
SAUMYA SINHA – KIRILL KUKHARENKO – MEIKE NEUWOHNER –		
2b	ThA77 - 513F - Polynomial Optimization Algorithms and Applications	Chair: BISSAN GHADDAR
ÁNGEL MANUEL GONZÁLEZ RUEDA – IGNACIO GÓMEZ CASARES – ROHIT KANNAN – BISSAN GHADDAR –		
2c	ThA924 - 511B - Combinatorial Optimization and Graph Theory 4	Chair: LUCAS MOSCHEN
ROLAND GRAPPE – VIET HUNG NGUYEN – SIYUE LIU – LUCAS MOSCHEN –		
2d	ThA211 - 513A - Explainable Artificial Intelligence and Mathematical Programming	Chair: JANNIS KURTZ
PATRICK VOSSLER – ILKER BIRBIL – NATHANAEL JO – JANNIS KURTZ –		
2e	ThA293 - 514B - Approximation Algorithms for the Traveling Salesman Problem	Chair: DAVID WILLIAMSON
BILLY JIN – JENS VYGEN – RICHARD SHAPLEY – DAVID WILLIAMSON –		
2f	ThA343 - 513E - Decomposition Methods	Chair: MARCO LÜBBECKE
KIBA EK KIM – MARCO LÜBBECKE – ANTOINE LEGRAIN – ELINA RÖNNBERG –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		THURSDAY 08:30-10:30
3a	ThA101 - 515A - Decision-making under Preference Uncertainty	Chair: WENJIE HUANG
SILVANA PESENTI – REZA VALIMORADI – HUIFU XU – WENJIE HUANG –		
3b	ThA937 - 512A - Discrete Stochastic Programming 1	Chair: MARYAM DARYALAL
EMANUELE CONCAS – PIETER KLEER – MARIA CAROLINA BAZOTTE CORGOZINHO – MARYAM DARYALAL –		
3c	ThA174 - 511D - Innovative RO Techniques in Transportation and Location Science	Chair: IHSAN YANIKOGLU
OMAR BENNOUNA – IHSAN YANIKOGLU – BESTE BASCIFTCI – FARZAD AVISHAN –		
3c	ThA277 - 511E - Optimization - Machine Learning Interfaces	Chair: CHAITHANYA BANDI
MIRAI TANAKA – TIM NIEMANN – DEEKSHA SINHA – CHAITHANYA BANDI –		
3d	ThA246 - 515B - Statistical Learning, Optimization and Stochastic Programming for Reinforcement Learning	Chair: MING YIN
YINGRU LI – DONGRUO ZHOU – HAQUE ISHFAQ – MING YIN –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		THURSDAY 08:30-10:30
4b	ThA949 - 516E - Transportation and logistics 5	Chair: CAGLA DURSUNOGLU
MOHAMMAD KARIMI – MALTE BILLEN – CAGLA DURSUNOGLU –		
4b	ThA948 - 511F - Transportation and logistics 4	Chair: ALI KERMANI
FAIZ HAMID – ALI KERMANI –		
4c	ThA954 - 512G - Scheduling and manufacturing 3	Chair: PAUL PATIENCE
ALIREZA ALAVI – XINYI GUO – PAUL PATIENCE –		
4d	ThA46 - 513C - Optimization Methods for Renewable Power Systems	Chair: ANTOINE LESAGE LANDRY
LESIA MITRIDATI – JUAN M. MORALES – JOHN W. SIMPSON PORCO – ANTOINE LESAGE LANDRY –		
4d	ThA204 - 513D - Mobile and static battery systems operations and planning	Chair: ANDY SUN
ALGHUMAYJAN SAUD – DIRK LAUNGER – MATTHEW BRUN – ROBIN LEGAULT –		
4g	ThA181 - 512E - A glimpse of computational optimization at Google II	Chair: JUAN PABLO VIELMA
HOSSEIN ESFANDIARI – DANIEL ALEJANDRO ESPINOZA GONZALEZ – JOEY HUCHETTE – CHRISTIAN TJANDRAATMADJA –		
4h	ThA256 - 512F - Applications of Nash Equilibrium Problems	Chair: ALEXANDRA SCHWARTZ
YUEH LIN – RAUL GARCIA – DEVANSH JALOTA – ALEXANDRA SCHWARTZ –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		THURSDAY 14:00-16:00
1a ThB37 - 510C - Sum-of-squares methods for continuous optimization and beyond	Chair: FRANCIS BACH	
GASPARD BEUGNOT – HAMZA FAWZI – GIOVANNI FANTUZZI – LUCAS SLOT –		
1a ThB137 - 511C - First-order Methods and Large-scale Optimization II	Chair: JUNWEN QIU	
PUYA LATAFAT – YASONG FENG – JIANTING PAN – NAOKI MARUMO –		
1a ThB93 - 516A - Nonconvex optimization: landscapes and dynamics I	Chair: CHRISTOPHER CRISCITIELLO	
CHRISTOPHER CRISCITIELLO – MATEO DIAZ – CIFUENTES DIEGO – YOSHI ARJEVANI –		
1a ThB35 - 516D - Nonlinear Programming for Data Science I	Chair: VIVAK PATEL	
VIVAK PATEL – PHILIPPE TOINT – KARL WELZEL – RAGHU PASUPATHY –		
1b ThB26 - 512H - Large-Scale and Structured Optimization	Chair: NAM HO NGUYEN	
SYUJI YAMADA – MILTIADES PARDALOS – QIAN ZHANG – NAM HO NGUYEN –		
1c ThB54 - 511A - Projection-free first-order methods for convex optimization	Chair: JAVIER PENA	
STEPHEN VAVASIS – RENBO ZHAO – ELIAS WIRTH – JAVIER PENA –		
1c ThB907 - 510B - Nonsmooth Optimization 2	Chair: ZEHUI LIU	
SIMONE REBEGOLDI – SAMIR ADLY – ZEHUI LIU – LEXIAO LAI –		
1d ThB123 - 510D - Online optimization and learning in quantum games	Chair: ANTONIOS VARVITSIOTIS	
CHRISTOF BRANDSTETTER – CONSTANTIN ICKSTADT – WAYNE LIN – EMMANOUIL VASILEIOS VLATAKIS GKARAGKOUNIS		
1d ThB910 - 510A - Semi-Definite Programming 3	Chair: TYLER WEAMES	
CHEE KHIAN SIM – JEFERSON ZAPATA – MAROUAN HANDA – TYLER WEAMES –		
1e ThB226 - 512E - Nonconvexity, stochasticity and hierarchy in optimization problems	Chair: MATHIAS STAUDIGL	
ALEXANDER BODARD – LUKE MARRINAN – YUYANG QIU – MATHIAS STAUDIGL –		
1e ThB913 - 514A - Variational Analysis, Variational Inequalities and Complementarity 3	Chair: ABDERRAHIM JOURANI	
XIAOXIAO MA – ABDERRAHIM JOURANI –		
1f ThB50 - 512C - Splitting and Randomized Methods 2	Chair: YANNAN CHEN	
ALOIS PICHLER – XUDONG LI – YANNAN CHEN –		
1g ThB300 - 515C - Stochastic DFO 1	Chair: MATT MENICKELLY	
ARMAND GISSLER – YIWEN CHEN – PENGCHENG XIE – MATT MENICKELLY –		
1h ThB318 - 512B - Shape Optimization	Chair: THEODOR KOMANN	
ZAKARIA HOUTA – KATHRIN WELKER – ROLAND HERZOG – THEODOR KOMANN –		

CLUSTER 2: DISCRETE OPTIMIZATION		THURSDAY 14:00-16:00
2a ThB161 - 513D - Advances in ILP: Branch and cut	Chair: ALINSON SANTOS XAVIER	
MARC PFETSCH – JOSEPH PAAT – ANTONIA CHMIELA – ALINSON SANTOS XAVIER –		
2a ThB918 - 513B - Mixed Integer Linear Programming 4	Chair: VENKAT AKHIL ANKEM	
RENAUD CHICOISNE – LOUISA SCHROEDER – FERNANDO DIAS – VENKAT AKHIL ANKEM –		
2b ThB34 - 513F - Advances on Multi-Objective and Separable Mixed Integer Nonlinear Programming		
Chair: MARIANNA DE SANTIS		
PHILIP DE CASTRO – MORITZ LINK – DANIELE PATRIA – LUCA MENCARELLI –		
2c ThB352 - 511B - Integer/Linear Optimisation	Chair: DANIELE CATANZARO	
YASUKO MATSUI – DANIELE CATANZARO – BENTO NATURA – ZHUAN KHYE KOH –		
2d ThB200 - 513A - Scaling Sparse Machine Learning	Chair: VASSILIS DIGALAKIS	
ISHY ZAGDOUN – GUANYI WANG – ILIAS ZADIK – ANNA DEZA –		
2e ThB244 - 514B - Advanced algorithms for chip design	Chair: STEPHAN HELD	
SUSANNE ARMBRUSTER – STEPHAN HELD – DANIEL BLANKENBURG – JAN MALTE SCHÜRKS –		
2f ThB342 - 513E - Combinatorial Optimization and Applications	Chair: JAAP PEDERSEN	
MOHAMMED GHANNAM – SHENG JIE CHEN – JAAP PEDERSEN – JULIAN YARKONY –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		THURSDAY 14:00-16:00
3a ThB935 - 515A - Continuous Stochastic Programming 2	Chair: DIONYSIS KALOGERIAS	
NICOLAS GAGNÉ – MARIUS ROLAND – CARLOS CARDONHA – DIONYSIS KALOGERIAS –		
3b ThB124 - 512A - Sequential Decision Making Under Uncertainty	Chair: RUI CHEN	
RUI CHEN – HAOMING SHEN – DAHYE HAN – ALFREDO TORRICO –		
3c ThB296 - 511D - Distributionally Robust Optimization: Theory Meets Applications	Chair: RUIWEI JIANG	
ALIAA ALNAGGAR – KANGLIN LIU – HOANG NAM NGUYEN – NOAM GOLDBERG –		
3d ThB186 - 515B - Efficient RL: Adaptive Experimentation and Estimation	Chair: IMON BANERJEE	
DAVID WOODRUFF – SOMAYEH MOAZENI – ETHAN CHE – IMON BANERJEE –		
3e ThB194 - 514C - Learning and Calibration in Data-Driven Optimization	Chair: TIANYU WANG	
MELVYN SIM – VISHAL GUPTA – CHUNLIN SUN – TIANYU WANG –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		THURSDAY 14:00-16:00
4b ThB155 - 511F - Models and Algorithms in Last-mile Delivery	Chair: OKAN ARSLAN	
HANG DONG – XIN WANG – MILAD ELYASI – OKAN ARSLAN –		
4b ThB950 - 516E - Transportation and logistics 6	Chair: ALEJANDRO TORIELLO	
SEYOUNG OH – ASEFEH HASSANI GOODARZI – MICHAEL LEVIN – ALEJANDRO TORIELLO –		
4d ThB118 - 511E - Optimization of Gas Networks	Chair: LARS SCHEWE	
JULIA GRÜBEL – PASCAL BÖRNER – ROWAN TURNER – LARS SCHEWE –		
4d ThB33 - 513C - Learning from Optimization in Energy	Chair: DAVI VALLADÃO	
NATHAN CHO – ANDRE FELIPE DE CARVALHO RAMOS – BRUNO FANZERES – DAVI VALLADÃO –		
4i ThB963 - 512G - Industrial applications 1	Chair: THIBAUT BARBIER	
JOHN CHINNECK – DAVID LANDAY – SIAMAK NADERI – THIBAUT BARBIER –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		THURSDAY 16:20-17:50
1a	ThC904 - 516D - Nonlinear Programming 3	Chair: IYAD WALWIL
MICHEL LAHMANN – ADEMIR ALVES RIBEIRO – IYAD WALWIL –		
1c	ThC38 - 510B - Nonsmooth calculus and optimization in machine learning	Chair: JALAL FADILI
ANTONIO SILVETI FALLS – SHOLOM SCHECHTMAN – VYACHESLAV KUNGURTSEV –		
1f	ThC129 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization II	Chair: FRANCESCO ORABONA
OLIVER HINDER – LOUIS SHARROCK – EVGENII CHZHEN –		
1g	ThC303 - 515C - Challenging DFO and SBO Applications 2	Chair: EDWARD HALLÉ HANNAN
SYLVERIO POOL MARQUEZ – CHARLOTTE BEAUTHIER – EDWARD HALLÉ HANNAN –		
1h	ThC319 - 512B - Multi-Level and Parallel Methods for PDE Constrained Optimization	Chair: WING HONG FELIX KWOK
RADOSLAV VUCHKOV – ED BUELER – WING HONG FELIX KWOK –		
CLUSTER 2: DISCRETE OPTIMIZATION		THURSDAY 16:20-17:50
2a	ThC180 - 513B - Advances in Mixed Integer Programming Theory	Chair: DANIEL DADUSH
THOMAS ROTHVOSS – STEFAN WELTGE – DANIEL DADUSH –		
2b	ThC159 - 513F - Discrete structures in non-convex problems	Chair: SONIA CAFIERI
KIM YU – GONZALO MUÑOZ – SONIA CAFIERI –		
2c	ThC355 - 511B - Kernels in Graphs	Chair: FRÉDÉRIC MEUNIER
HÉLÈNE LANGLOIS – FRÉDÉRIC MEUNIER – PÉTER ERDOS –		
2d	ThC928 - 513A - Machine Learning and Discrete Optimization 2	Chair: RAHUL MAZUMDER
JARED TANNER – ALEXANDRE FOREL – RAHUL MAZUMDER –		
2f	ThC270 - 513E - Augmentation Methods and Dynamic Programming	Chair: DOMINIC YANG
TOLGA BEKTAS – DOMINIC YANG – IBRAHIM DAN DIJE –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		THURSDAY 16:20-17:50
3a	ThC252 - 515A - Advances in Risk Optimization and Design	Chair: QUANYAN ZHU
YINGDONG LU – UTSAV SADANA – SHUTIAN LIU –		
3b	ThC116 - 512A - Advanced algorithms in stochastic mixed-integer programming	Chair: KIBAEK KIM
KAIWEN FANG – HANBIN YANG – GEUNYEONG BYEON –		
3d	ThC255 - 515B - Adaptive stochastic optimization with applications in energy system operations and transition	Chair: ANDY SUN
PETER BARKLEY – PANAGIOTIS ANDRIANESIS – THOMAS LEE –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		THURSDAY 16:20-17:50
4b	ThC951 - 516E - Transportation and logistics 7	Chair: ALI AKBAR SADAT ASL
STEFAN ROPKE – BIENVENU BAMBI – ALI AKBAR SADAT ASL –		
4d	ThC261 - 513C - Research at Hydro-Québec: Black-box Optimization	Chair: MIGUEL DIAGO
VINCENT MAI – MIGUEL DIAGO – ALEXIA MARCHAND –		

CLUSTER 1: CONTINUOUS OPTIMIZATION		FRIDAY 08:30-10:30
1a FA94 - 510C - Nonconvex optimization: landscapes and dynamics II		Chair: ANDREEA MUSAT
JU SUN – DAMIEN SCIEUR – AKSHAY RAMACHANDRAN – ANDREEA MUSAT –		
1a FA260 - 516A - Open-source software for nonlinear programming		Chair: CHARLIE VANARET
DAVID COLE – CHARLIE VANARET – MIHAI ANITESCU – DAVID KIESSLING –		
1a FA83 - 516D - Gradient-based methods: theory and advances		Chair: SHOTA TAKAHASHI
JINGYANG LI – SHOTA TAKAHASHI – KERRY HE –		
1b FA27 - 512H - Global Optimal Search: Unleashing Maximum Potential		Chair: FREDERIC MESSINE
TANMAYA KARMARKAR – ADRIAN GÖSS – FREDERIC MESSINE – KAMIL KAMIL KHAN –		
1c FA25 - 510B - Some recent advances in zeroth and first-order optimization methods		Chair: TING KEI PONG
CHING PEI LEE – ANDREAS THEMELIS – ANDERSEN (MAN SHUN) ANG – YAOLIANG YU –		
1c FA113 - 511A - Advances in Variational Analysis and Nonsmooth Optimization II		Chair: JOHANNES ROYSET
FRANK FISCHER – NGHIA TRAN – MEISAM RAZAVIYAYN – KRISHNAKUMAR BALASUBRAMANIAN –		
1d FA17 - 510A - Recent advances in conic optimization - Part 2		Chair: BRUNO LOURENÇO
CASEY GARNER – MITSUHIRO NISHIJIMA – MEHDI KARIMI – XIAOZHOU WANG –		
1d FA89 - 510D - Recent advances in quadratic and semidefinite optimization		Chair: FRANK PERMENTER
CHENYANG YUAN – RICHARD ZHANG – KEVIN SHU –		
1e FA103 - 514A - Recent advances in bilevel optimization		Chair: JAN HAROLD ALCANTARA
SHANGZHI ZENG – JEONGYEOL KWON – ALAIN ZEMKOHO – JAN HAROLD ALCANTARA –		
1f FA51 - 512C - Splitting and Randomized Methods 3		Chair: CHUNFENG CUI
CAIRONG CHEN – XINGJU CAI – CHUNFENG CUI –		
1g FA326 - 515C - DFO based on Approximations to Derivatives 2		Chair: ANNE AUGER
YANJUN LIU – TIANYU WANG – GEOVANI GRAPIGLIA – ANNE AUGER –		

CLUSTER 2: DISCRETE OPTIMIZATION		FRIDAY 08:30-10:30
2a	FA228 - 513B - Integer programming: structure and algorithms	Chair: MANUEL APRILE
ALEXANDRA ANNA LASSOTA – HYUNWOO LEE – LUCA FERRARINI – STEFAN KOBER –		
2a	FA262 - 513D - Aspects of Symmetry and Fairness in Integer Programming and Combinatorial Optimization	Chair: CHRISTOPHER HOJNY
ADRIAN VETTA – ANNIKA JAEGER – STEN WESSEL – CHRISTOPHER HOJNY –		
2b	FA348 - 513F - Convexification Techniques for MINLP	Chair: AMÉLIE LAMBERT
ABHISHEK PATHAPATI PATHAPATI – AMÉLIE LAMBERT – TOBIAS KUEN – MATTHIAS WALTER –		
2c	FA925 - 511B - Combinatorial Optimization and Graph Theory 5	Chair: AKIYOSHI SHIOURA
RICARDO EULER – YUICHIRO MIYAMOTO – MARTINA CERULLI – AKIYOSHI SHIOURA –		
2c	FA351 - 511C - Combinatorial Optimisation and Graph Theory	Chair: MARTIN DREES
RUI GONG – MARTIN DREES – VÍCTOR BLANCO – NIKLAS SCHLOMBERG –		
2d	FA173 - 513A - Progress on Machine Learning Approaches for Solving Combinatorial Optimization	
Chair: ZAIWEN WEN		
ZAIWEN WEN – AKANG WANG – JUNCHI YAN – JIANYONG SUN –		
2e	FA245 - 512A - TSP and vehicle routing	Chair: STEFAN HOUGARDY
LUISE PUHLMANN – STEFAN HOUGARDY – ELAHE AMIRI – DIRK MÜLLER –		
2e	FA150 - 514B - Submodularity and Beyond in Approximation Algorithms	Chair: KARTHEKEYAN CHANDRASEKARAN
ISHAN BANSAL – SHUBHANG KULKARNI – KENT QUANRUD – SABINE MÜNCH –		
2f	FA933 - 513E - Computational Discrete and Integer Optimization 2	Chair: QUENTIN LOUVEAUX
ANTHONY KARAHALIOS – ANDREAS TILLMANN – SHUNJI UMETANI – QUENTIN LOUVEAUX –		
2f	FA224 - 513C - Tightening MILPs by Reformulation	Chair: DOMINIK KAMP
JÖRG RAMBAU – JOÃO DIONÍSIO – JOHANNES SCHMIDT – DOMINIK KAMP –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		FRIDAY 08:30-10:30
3a	FA247 - 515A - Latest Developments in Stochastic Optimization	Chair: JONATHAN LI
DAVID SAUNDERS – ANDREW LIM – QINYU WU – TIAN TIAN MAO –		
3c	FA133 - 511D - Robust Optimization meets other domains	Chair: AHMADREZA MARANDI
JUDITH BRUGMAN – JOHANNES THÜRAUF – AYOUB FOUSSOUL – AHMADREZA MARANDI –		
3c	FA140 - 511E - Robust Optimization for Operations Problems	Chair: ZHENYU HU
MIGUEL CARRASCO – KAI PAN – CHUN CHENG – ZHENYU HU –		
3d	FA182 - 515B - Advances at the intersection of reinforcement learning and stochastic optimization	
Chair: SAEED GHADIMI		
DANIEL RANGEL MARTINEZ – WARREN POWELL – SWATI GUPTA – ANDRZEJ RUSZCZYNSKI –		
3e	FA259 - 512B - New models and applications of data-driven optimization	Chair: HODA BIDKHORI
MATEUS WAGA – CAGIN URU – HODA BIDKHORI – TRISTAN POLLNER –		
3e	FA234 - 514C - Data-driven decision making in high-stakes domains: notes on fairness and interpretability	Chair: SINA AGHAEI
AGHAEI		
LESIA SEMENOVA – SINA AGHAEI – JOHANN GAEBLER – SINA BAHARLOUEI –		

CLUSTER 4: APPLICATIONS AND COMPUTATION	FRIDAY 08:30-10:30
4b FA193 - 516E - Network Design for Transportation Planning IV	Chair: MARIA ELENA BRUNI
GUIDO PERBOLI – GIACOMO LANZA – MAHDI MOSTAJABDAVEH – MARIA ELENA BRUNI –	
4c FA955 - 512E - Scheduling and manufacturing 4	Chair: FARIN RASTGAR AMINI
TAZU NOMOTO – DEBORA PRETTI RONCONI – HARCÈNAGE DANSOU – FARIN RASTGAR AMINI –	
4h FA330 - 512F - Computing Equilibria of Games	Chair: YICHUAN DING
FRANCESCO MORRI – ZI YI HUANG – CHUANGYIN DANG – YICHUAN DING –	
4i FA288 - 512G - Optimization and Validation of Execution Models at Amazon Scale	Chair: QI CHEN
ZIHAO LI – QI CHEN – AKSHIT KUMAR – OZLEM BILGINER –	

CLUSTER 1: CONTINUOUS OPTIMIZATION	FRIDAY 14:00-16:00
1a FB75 - 510C - The interplay between optimization, statistics and machine learning II	Chair: ALEX WANG
SHIQIAN MA – FENG RUAN – YUETIAN LUO – VAIDEHI SRINIVAS –	
1a FB154 - 511C - First-order Methods and Large-scale Optimization III	Chair: LEXIAO LAI
L FELIPE BUENO – LI JIANG – PIERFRANCESCO BENEVENTANO – ZIJIAN LIU –	
1a FB114 - 516A - Computational advancement in nonconvex programming	Chair: ALEXIS MONTISON
JUN YA GOTOH – DIGVIJAY BOOB – MIJU AHN – ALEXIS MONTISON –	
1a FB70 - 516D - Nonlinear Programming for Data Science II	Chair: VIVAK PATEL
SERGE GRATTON – ROBERT WEBBER – MLADEN KOLAR – EDWARD TANSLEY –	
1b FB1 - 512H - Recent Advances in Global Optimization of Nonconvex Programs	Chair: HARSHA NAGARAJAN
ROBERT BURLACU – ARVIND RAGHUNATHAN – ERIN GEORGE – MOHIT TAWARMALANI –	
1c FB257 - 510B - Complexity of Stationarity Concepts in Non-Smooth Optimization	Chair: MAN CHO So
LAI TIAN – WENQING OUYANG – LINGLINGZHI ZHU –	
1d FB275 - 510D - Recent Advancements in the Theory and Applications of Semidefinite Programming	Chair: HAO HU
HAO HU – MANISH KRISHAN LAL – GABOR PATAKI – HENRY WOLKOWICZ –	
1d FB36 - 510A - SDP and Exact Relaxations of Hard Problems	Chair: YONG XIA
LOUIS LORTIE – SUNYOUNG KIM – SHUYANG LING – GODAI AZUMA –	
1e FB336 - 514A - Nonsmooth and hierarchical optimization in machine learning	Chair: RAHUL PARHI
ZHAOSONG LU – DAVID BENFIELD – RAHUL PARHI –	
1f FB52 - 512C - Splitting and Randomized Methods 4	Chair: JINGYA CHANG
ISABEL JACOB – TING LI – JINGYA CHANG –	
1g FB327 - 515C - Stochastic DFO 2	Chair: WARREN HARE
KWASSI JOSEPH DZAHINI – CUNXIN HUANG – CLÉMENT ROYER – WARREN HARE –	

CLUSTER 2: DISCRETE OPTIMIZATION		FRIDAY 14:00-16:00
2a FB151 - 513D - Recent Developments in MIP Theory and Applications	Chair: BURAK KOCUK	
BURAK KOCUK – GUSTAVO ANGULO – DAVID BERNAL NEIRA – DIEGO ALEJANDRO MORAN RAMIREZ –		
2a FB919 - 513B - Mixed Integer Linear Programming 5	Chair: JAKOB NORDSTRÖM	
NATHAN BENEDETTO PROENÇA – JAKOB NORDSTRÖM –		
2b FB9 - 513F - Theory of Mixed Integer Non-linear Programming	Chair: AIDA KHAJAVIRAD	
ROBERT HILDEBRAND – SILVIA DI GREGORIO – AIDA KHAJAVIRAD – BARBARA ANNA FIEDOROWICZ –		
2c FB926 - 511B - Combinatorial Optimization and Graph Theory 6	Chair: JOSEPH POREMBA	
HIROTAKA ONO – CLAUDIA O LÓPEZ SOTO – EL MEHDI OUDAUD – JOSEPH POREMBA –		
2d FB929 - 513A - Machine Learning and Discrete Optimization 3	Chair: FEDERICO D’ONOFRIO	
BRIAN LIU – PIERRE HOUDOUIN – STEVEN DAHDAH – FEDERICO D’ONOFRIO –		
2e FB295 - 512A - Algorithms with Predictions	Chair: CLIFFORD STEIN	
ERIC BALKANSKI – BENJAMIN MOSELEY – THOMAS LAVASTIDA – CLIFFORD STEIN –		
2e FB225 - 514B - Fairness via Optimization Modeling and Approximation Algorithms	Chair: DAVID SHMOYS	
VARUN SURIYANARAYANA – MAROUANE IBN BRAHIM – ARAVIND SRINIVASAN – DAVID SHMOYS –		
2f FB268 - 513E - Integer Programming Applications in Mobility and Traffic	Chair: NIELS LINDNER	
NIELS LINDNER – ENRICO BORTOLETTO – BERENIKE MASING – FABIAN LÖBEL –		
2g FB242 - 512G - Learning and Explaining Constraints	Chair: NADJIB LAZAAR	
J. CHRISTOPHER BECK – ALEXANDER TESCH – IAN MIGUEL – NADJIB LAZAAR –		
CLUSTER 3: OPTIMIZATION UNDER UNCERTAINTY		FRIDAY 14:00-16:00
3a FB936 - 515A - Continuous Stochastic Programming 3	Chair: SHIBSHANKAR DEY	
FEDERICO BOBBIO BOBBIO – SHIBSHANKAR DEY –		
3c FB122 - 511E - Optimization with Marginals and Moments	Chair: KARTHIK NATARAJAN	
KARTHIK NATARAJAN – ARJUN RAMACHANDRA – HAILONG SUN – HE SIMAI –		
3c FB104 - 511D - Data-driven Robust Optimization	Chair: JUSTIN DUMOCHELLE	
BO LIN – IRINA WANG – JUSTIN DUMOCHELLE –		
3c FB184 - 511A - Exact methods for robust nonlinear optimization	Chair: DICK DEN HERTOOG	
DANIQUE DE MOOR – DICK DEN HERTOOG – THORORIS KOUKOUVINOS –		
3d FB109 - 515B - Advances in multi-stage stochastic programming	Chair: BERNARDO PAGNONCELLI	
YONGJIA SONG – HAMED RAHIMIAN – MERVE BODUR – BERNARDO PAGNONCELLI –		
3e FB141 - 514C - Advances in data-driven decision making	Chair: VELIBOR MISIC	
NATHAN JUSTIN – YA CHI CHU – LONG HE – VELIBOR MISIC –		
CLUSTER 4: APPLICATIONS AND COMPUTATION		FRIDAY 14:00-16:00
4b FB172 - 511F - Demand-based and stochastic urban transportation planning	Chair: LÉA RICARD	
GIANMARCO ANDREANA – ÖYKÜ NAZ ATILA – JIATENG YIN – JIAQI LIANG –		
4b FB164 - 516E - Humanitarian Logistics	Chair: VITTORIO NICOLETTA	
HANI ZBIB – VITTORIO NICOLETTA – ALBORZ HASSANZADEH – MAEDEH SHARBAF –		
4d FB959 - 513C - Energy and Environment 4	Chair: HANDE BENSON	
TRULS FLATBERG – HANDE BENSON –		
4g FB165 - 512E - A glimpse of computational optimization at Google III	Chair: JUAN PABLO VIELMA	
TOBY DAVIES – ROSS ANDERSON – DANIEL DUQUE – JUAN PABLO VIELMA –		

Abstracts

1a - MB11 - 516D - **Advances in Large-Scale Nonlinear Optimization I** Chair: ALBERT BERAHAS 

Robustly Stable Accelerated Gradient Methods – MERT GURBUZBALABAN, RUTGERS UNIVERSITY –

We consider the problem of minimizing a strongly convex smooth function where the gradients are subject to additive worst-case deterministic errors that are square-summable. We study the trade-offs between the convergence rate and robustness to gradient errors when designing the parameters of a first-order algorithm. We focus on a general class of momentum methods (GMM) with constant stepsize and two momentum parameters which can recover gradient descent (GD), Nesterov's accelerated gradient (NAG), the heavy-ball (HB) and the triple momentum methods (TMM) as special cases. We measure the robustness of an algorithm in terms of the cumulative suboptimality over the iterations normalized by the squared 2 norm of the gradient errors. This quantity can be interpreted as the (squared) 2 gain of a dynamical system that represents the GMM iterations where the input is the gradient error sequence and the output is a weighted distance to the optimum. For quadratic objectives, we compute the 2 gain explicitly leveraging its representation as the H_∞ norm of the GMM system in the frequency domain and construct gradient errors that lead to worst-case performance explicitly. We also study the stability of GMM with respect to multiplicative errors by characterizing the structured real and stability radius of the GMM system through their connections to the H_∞ norm. This allows us to compare GD, HB, NAG methods in terms of robustness, and argue that HB is not as robust as NAG despite being the fastest in terms of the rate. We then develop the robustly stable heavy ball method that can be faster than NAG while being at the best robustness level possible. We also propose the robustly stable gradient descent that is the fastest version of GD with constant stepsize while being at the best robustness level. Finally, we extend our framework to general strongly convex smooth objectives, providing non-asymptotic rate results for inexact GMM methods and bounds on the 2 gain where we can choose the GMM parameters to systematically trade off the rate to robustness in a computationally tractable framework. Finally, we discuss extensions of our results and algorithms to the stochastic optimization setting where gradient errors are stochastic and to min-max optimization settings.

Advanced, Adaptive and Flexible Algorithms for Decentralized Optimization – ALBERT BERAHAS, UNIVERSITY OF MICHIGAN – Raghu Bollapragada, Shagun Gupta, Suhail Shah

The problem of optimizing an objective function by employing a decentralized procedure using multiple agents in a connected network has gained significant attention over the last decades. This is due to the wide applicability of decentralized optimization to many important science and engineering applications such as, optimal control, machine learning, robotics, sensor networks, and smart grids. Decentralized optimization prob-

lems come in diverse shapes and forms, and could have very different characteristics. In this talk, we discuss novel flexible approaches for solving decentralized optimization problems that adapt to problem characteristics. We present two unifying algorithmic frameworks that recover popular algorithms as special cases. We discuss the rationale behind our proposed techniques, convergence in expectation and complexity guarantees for our algorithms, and present encouraging numerical results.

Stochastic Interior-Point method for Inequality constrained optimization – QI WANG, LEHIGH UNIVERSITY –

An interior-point algorithm is proposed to solve the inequality constrained optimization problems where the objective and constraints functions are continuously differentiable. Such problems appear, for example, in the applications of machine learning problems with fairness constraints and physics constraints. It is assumed that only the stochastic objective gradient can be evaluated. The algorithm is unique in its use of inner neighborhoods of the feasible region—defined by a positive and vanishing neighborhood-parameter sequence—in which the iterates are forced to remain. For bound-constrained problems, it is shown that with a careful balance between the barrier, step-size, and neighborhood sequences, the proposed algorithm satisfies convergence guarantees in both deterministic and stochastic settings, showing promising experimental results. The convergence results are also extended to nonlinear-constrained problems under specific assumptions.

Nonsmooth Optimization on a Finer Scale: Bounded Local Subgradient Variation Perspective – JELENA DIAKONIKOLAS, UNIVERSITY OF WISCONSIN MADISON – Cristóbal Guzmán

Nonsmooth optimization problems are ubiquitous in industrial and machine learning applications, arising from the need to address tasks such as resource allocation, threat detection, and model training. Within the realm of mathematical optimization, nonsmooth problems stand as some of the most challenging; yet they often offer the possibility of developing efficient algorithms with provable guarantees. The complexity of these problems, encompassing both lower and upper bounds, has historically typically been examined under generic assumptions, bounding the growth of the objective functions by assuming Lipschitz continuity of either the objective itself or its gradient. In this talk, I will argue that these traditional assumptions defining classes of nonsmooth optimization problems inadequately capture local properties of problems that may make them amenable to more efficient algorithms. I will introduce a notion of local bounded variation of the (sub)gradient and discuss how, under this notion, we can obtain a more fine-grained characterization of the complexity of nonsmooth problems. One consequence of these results is that, contrary to prior belief, the complexity of nonsmooth optimization problems, such as those with piecewise linear objectives with polynomially many pieces, can be improved using parallelization even in high dimensions, either if the problem is unconstrained or if the function is allowed to be queried outside the feasible set.

1a - MB39 - 510C - **Advances in numerical methods for large scale nonlinear optimization and applications I** Chair: MARGHERITA PORCELLI

Exploring the Spectral Proximal Gradient Method and Applications – PAULO J. S. SILVA, UNIVERSIDADE ESTADUAL DE CAMPINAS –

The spectral proximal gradient method is an accelerated variant of the proximal gradient algorithm, leveraging second-order information obtained through the Barzilai-Borwein-Raydan stepsize. This talk concisely overviews its historical evolution, theoretical underpinnings, and key convergence results. We will also exemplify its behavior in many applications spanning different domains like projection onto the intersection of convex sets, low-rank plus sparse matrix decomposition, l_0 -regularized machine learning, and full wave inversion in Geophysics.

Regularized methods via cubic model subspace minimization for nonconvex optimization – MARGHERITA PORCELLI, UNIVERSITÀ DI FIRENZE – Stephania Bellavia, Davide Palitta, Valeria Simoncini

Adaptive cubic regularization methods for solving nonconvex problems need the efficient computation of the trial step, involving the minimization of a cubic model. We present new approach in which this model is minimized in a low dimensional subspace that, in contrast to classic approaches, is reused for a number of iterations. Whenever the trial step produced by the low-dimensional minimization process is unsatisfactory, we employ a regularized Newton step whose regularization parameter is a by-product of the model minimization over the low-dimensional subspace. We show that the worst-case complexity of classic cubic regularized methods is preserved, despite the possible regularized Newton steps. We focus on the large class of problems for which (sparse) direct linear system solvers are available and provide several experimental results showing the very large gains of our new approach when compared to standard implementations of adaptive cubic regularization methods based on direct linear solvers.

Scalable interior point methods for large scale discrete optimal transport problems – FILIPPO ZANETTI, UNIVERSITY OF EDINBURGH – Jacek Gondzio, Stefano Cipolla

In this talk, we address the solution of optimal transport problems using Interior Point Methods (IPM). We propose two approaches which exploit the expected sparsity of the optimal solution to enhance the efficiency of the numerical linear algebra required to solve the Newton system. The first approach deals with classical optimal transport over dense bipartite graphs, enforcing sparsity of the intermediate approximations by means of a column-generation-like approach. The numerical experiments show that this method can achieve scalable results for very large problems. The second approach deals with optimal transport over sparse graphs, using a proximal stabilized interior point method. The induced primal-dual regularization allows to use sparsified versions of the normal equations to inexpensively generate IPM search directions. The proposed method, despite a potential high level of inexactness in the Newton direction, retains the polynomial complexity of stan-

dard IPMs, for suitable choices of the sparsification parameters. Numerical results show that this approach is efficient and robust for large scale problems, and can outperform classical graph algorithms like network simplex and cost scaling.

Recent advances in preconditioners for interior point methods – JACEK GONDZIO, UNIVERSITY OF EDINBURGH –

We address challenges related to the application of iterative schemes when solving KKT systems arising in interior point methods. We demonstrate that certain undesirable features of these systems can be exploited to design preconditioners for iterative methods. Recent advances including the new theoretical insights and the new computational results will be presented.

1a - MB62 - 516A - **Development of nonlinear optimization algorithms and applications** Chair: BIN GAO

Improved lower bounds for stochastic optimization under Markovian sampling – ERMIN WEI, NORTHWESTERN UNIVERSITY – Zhenyu Sun

Stochastic optimization with Markovian sampling allows the sampling scheme following a Markov chain, which generalizes the usual i.i.d. sampling assumption. This problem encompasses various applications that range from asynchronous distributed optimization to online reinforcement learning. Given any Markovian sampling process, we give a lower bound on the sample complexity of finding ϵ -approximate stationary points for any first-order methods. In particular, for samples drawn from a countable Markov process, any algorithm that accesses smooth, non-convex functions through queries to a stochastic gradient oracle with bounded noise, requires at least ϵ^{-2} samples to find an ϵ -approximate stationary point. The lower bound matches the upper bound of a recently proposed algorithm, MC-SAG, hence establishing its min-max optimality.

Designing surrogate functions for reinforcement learning – NICOLAS LE ROUX, MICROSOFT RESEARCH –

We propose a series of theoretical results related to designing surrogate functions for Reinforcement Learning. We cover links between reinforcement learning and maximum likelihood, discuss choices of Bregman divergences when doing mirror descent, and introduce accelerated methods in function space.

This talk aims at explaining the various possible choices and their consequences when designing off-policy RL algorithms, for instance for reinforcement learning from human feedback.

The Exactness of the L1 Penalty Function for a Class of Mathematical Programs with Generalized Complementarity Constraints – XIN LIU, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES –

In a Mathematical Program with Generalized Complementarity Constraints (MPGCC), complementarity relationships are imposed between each pair of variable blocks. MPGCC includes the traditional Mathematical Program with Complementarity Constraints (MPCC) as a special case. On account of the disjunctive feasible region, MPCC and MPGCC are generally difficult to handle. The L1 penalty method, often

adopted in computation, opens a way of circumventing the difficulty. Yet it remains unclear about the exactness of the L1 penalty function, namely, whether there exists a sufficiently large penalty parameter so that the penalty problem shares the optimal solution set with the original one. In this paper, we consider a class of MPGCCs that are of multi-affine objective functions. This problem class finds applications in various fields, e.g., the multi-marginal optimal transport problems in many-body quantum physics and the pricing problems in network transportation. We first provide an instance from this class, the exactness of whose L1 penalty function cannot be derived by existing tools. We then establish the exactness results under rather mild conditions. Our results cover those existing ones for MPCC and apply to multi-block contexts.

Low-rank optimization on Tucker tensor varieties – BIN GAO, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE – Renfeng Peng, Ya Xiang Yuan

In the realm of tensor optimization, low-rank tensor decomposition, particularly Tucker decomposition, stands as a pivotal technique for reducing the number of parameters and for saving storage. We embark on an exploration of Tucker tensor varieties—the set of tensors with bounded Tucker rank—in which the geometry is notably more intricate than the well-explored geometry of matrix varieties. We give an explicit parametrization of the tangent cone of Tucker tensor varieties and leverage its geometry to develop provable gradient-related line-search methods for optimization on Tucker tensor varieties. The search directions are computed from approximate projections of antigradient onto the tangent cone, which circumvents the calculation of intractable metric projections. To the best of our knowledge, this is the first work concerning geometry and optimization on Tucker tensor varieties. In practice, low-rank tensor optimization suffers from the difficulty of choosing a reliable rank parameter. To this end, we incorporate the established geometry and propose a Tucker rank-adaptive method that is capable of identifying an appropriate rank during iterations while the convergence is also guaranteed. Numerical experiments on tensor completion with synthetic and real-world datasets reveal that the proposed methods are in favor of recovering performance over other state-of-the-art methods. Moreover, the rank-adaptive method performs the best across various rank parameter selections and is indeed able to find an appropriate rank.

1a - MC55 - 510C - **From First- to High-order Methodologies: Recent Advances and Trends I** Chair: ARYAN MOKHTARI ↩

Efficient Methods for Solving Bilevel-structure Optimization Problem – ERFAN YAZDANDOOST HAMEDANI, THE UNIVERSITY OF ARIZONA –

Bilevel optimization is an important class of hierarchical decision-making where one optimization problem is nested within another. This modeling framework finds extensive applications across diverse domains, including machine learning, energy systems, and economics. However, as the scale and complexity of these systems continue to increase, there is a critical need to develop scalable and reliable optimization algorithms for bilevel optimization. For example, model-agnostic

meta-learning in machine learning can be cast as a bilevel optimization where the follower computes task-specific parameters using a set of meta-parameters (leader’s decision variable) and the goal is to learn meta-parameters that produce appropriate task-specific parameters after adaptation. Other examples in this area include few-shot learning, reinforcement learning, and hyperparameter learning.

In this talk, we discuss various bilevel-structure problems including constrained bilevel optimization and bilevel saddle point problems. For each problem structure, we introduce a novel efficient first-order algorithm with a convergence rate guarantee. Several methods have been developed for tackling unconstrained bilevel optimization problems, but there is limited work on methods for constrained and min-max settings. Our proposed methods have an improved per-iteration complexity, surpassing existing methods. We also present numerical experiments to showcase the superior performance of our methods compared with state-of-the-art methods.

Non-asymptotic Global Convergence Rates of BFGS with Exact Line Search – ARYAN MOKHTARI, THE UNIVERSITY OF TEXAS AT AUSTIN – Qiujiang Jin, Ruichen Jiang

In this talk, we explore the non-asymptotic global convergence rates of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method implemented with exact line search. Notably, due to Dixon’s equivalence result, our findings are also applicable to other quasi-Newton methods in the convex Broyden class employing exact line search, such as the Davidon-Fletcher-Powell (DFP) method. Specifically, we focus on problems where the objective function is strongly convex with Lipschitz continuous gradient and Hessian. Our results hold for any initial point and any symmetric positive definite initial Hessian approximation matrix. The analysis unveils a detailed three-phase convergence process, characterized by distinct linear and superlinear rates, contingent on the iteration progress. Additionally, our theoretical findings demonstrate the trade-offs between linear and superlinear convergence rates for BFGS when we modify the initial Hessian approximation matrix, a phenomenon further corroborated by our numerical experiments.

Second-order Information Promotes Mini-Batch Robustness in Variance-Reduced Gradients – MICHAL DEREZINSKI, UNIVERSITY OF MICHIGAN – Sachin Garg, Albert Berahas

We show that, for finite-sum minimization problems, incorporating partial second-order information of the objective function can dramatically improve the robustness to mini-batch size of variance-reduced stochastic gradient methods, making them more scalable while retaining their benefits over traditional Newton-type approaches. We demonstrate this phenomenon on a prototypical stochastic second-order algorithm, called Mini-Batch Stochastic Variance-Reduced Newton (Mb-SVRN), which combines variance-reduced gradient estimates with access to an approximate Hessian oracle. In particular, we show that when the data size n is sufficiently large, i.e., $n \gg \alpha^2 \kappa$, where κ is the condition number and α is the Hessian approximation factor, then Mb-SVRN achieves a fast linear convergence rate that is independent of the gradient mini-batch size b , as long as b is in the range between 1 and $b_{\max} = O(n/(\alpha \log n))$. Only after increasing the mini-batch size past this critical point $b_{\max}$, the method begins

to transition into a standard Newton-type algorithm which is much more sensitive to the Hessian approximation quality. We demonstrate this phenomenon empirically on benchmark optimization tasks showing that, after tuning the step size, the convergence rate of Mb-SVRN remains fast for a wide range of mini-batch sizes, and the dependence of the phase transition point $b_{\max}$ on the Hessian approximation factor α agrees with our theory.

1a - MC902 - 516D - Nonlinear Programming 1

Chair: XINWEI LIU

Connections between Gradient Descent and the Power Method – RACHAEL TAPPENDEN, UNIVERSITY OF CANTERBURY –

Applying Gradient Descent (GD) with a fixed step size to minimize a quadratic function is equivalent to running the Power Method (PM) on the gradients. The connection between GD with a fixed step size and the PM, both with and without fixed momentum, is thus established. Consequently, valuable eigen-information is available via GD. By considering the dynamics of the gradients and iterates, new step size strategies are proposed to improve the practical performance of GD.

A Discretization Approach for Low-Dimensional Non-Convex Bilevel Optimization – PRASHANT KHANDURI, WAYNE STATE UNIVERSITY – Xiaotian Jiang, Ioannis Tsaknakis, Mingyi Hong

Bilevel optimization has become an invaluable tool in areas like hyper-parameter optimization, meta-learning, and reinforcement learning. However, most modern bilevel algorithms are developed for problems with strongly convex and/or unconstrained lower-level tasks. On the contrary, in our work, we deal with a problem where the lower-level task is non-convex and constrained. To solve such a challenging problem, we propose a novel value-function-based approach in which the value function and the respective problem reformulation are discretized. Under this approach, the value function's optimization problem is transformed into an easier convex optimization task. We tackle the discretized problem using a penalty approach and establish the relation between the KKT points, and the local and global minima of the two problems. For the solution of the penalty reformulated problem, we develop a gradient descent-based method and calculate its convergence rate. Our experimental results confirm the effectiveness of the proposed method.

D-stationary points for optimization with least constraint violation and the convergence of an SQP method – XINWEI LIU, HEBEI UNIVERSITY OF TECHNOLOGY – Yu Hong Dai

We consider the nonlinear optimization problem with least ℓ_1 -norm measure of constraint violations and introduce the concepts of the D-stationary points with the help of exact penalty function. If the stationary point is feasible, they correspond to the Fritz-John stationary point, the KKT stationary point and the singular stationary point, respectively. A new exact penalty SQP method with inner and outer iterations is proposed which admits the convergence to a D-stationary point and rapid infeasibility detection without driving the penalty parameter to zero. It demonstrates the commentary given in [SIAM J. Optim., 20 (2010), 2281–2299] and can be thought

to be a supplement of the theory of nonlinear optimization on rapid detection of infeasibility. Some preliminary numerical results demonstrate that the proposed method is robust and efficient in solving infeasible nonlinear problems and a degenerate problem without LICQ in the literature.

1a - TA31 - 516D - Advances in Nonconvex Optimization Chair: SALAR FATTAHI

From Stability to Chaos: Analyzing Gradient Descent Dynamics in Quadratic Regression – XUXING CHEN, UNIVERSITY OF CALIFORNIA, DAVIS – Krishnakumar Balasubramanian, Promit Ghosal, Bhavya Agrawalla

Iterative algorithms like the gradient descent and its stochastic variants are widely used to train deep neural networks. Classical optimization theory operates under small-order step-sizes which guarantees convergence. In stark contrast to traditional optimization, recent empirical studies in deep learning have revealed that training deep neural networks with large-order step-sizes yields superior generalization performance. Motivated by this line of work, we conduct a comprehensive investigation into the dynamics of gradient descent using large-order constant step-sizes in the context of quadratic regression models. Within this framework, we reveal that the dynamics can be encapsulated by a specific cubic map, naturally parameterized by the step-size. Through a fine-grained bifurcation analysis concerning the step-size parameter, we delineate five distinct training phases: (1) monotonic, (2) catapult, (3) periodic, (4) chaotic, and (5) divergent, precisely demarcating the boundaries of each phase. As illustrations, we provide examples involving phase retrieval and two-layer neural networks employing quadratic activation functions and constant outer-layers, utilizing orthogonal training data. Our simulations indicate that these five phases also manifest with generic non-orthogonal data. We also empirically investigate the generalization performance when training in the various non-monotonic (and non-divergent) phases. In particular, we observe that performing an ergodic trajectory averaging stabilizes the test error in non-monotonic (and non-divergent) phases. This is joint work with Krishnakumar Balasubramanian, Promit Ghosal, and Bhavya Agrawalla.

Optimal Low-Rank Matrix Completion: Semidefinite Relaxations and Eigenvector Disjunctions – RYAN CORY WRIGHT, IMPERIAL COLLEGE LONDON – Jean Pauphilet, Dimitris Bertsimas, Sean Lo

Low-rank matrix completion consists of computing a matrix of minimal complexity that recovers a given set of observations as accurately as possible, and has numerous applications such as product recommendation. Unfortunately, existing methods for solving low-rank matrix completion are heuristics that, while highly scalable and often identifying high-quality solutions, do not possess any optimality guarantees. We reexamine matrix completion with an optimality-oriented eye, by reformulating low-rank problems as convex problems over the non-convex set of projection matrices and implementing a disjunctive branch-and-bound scheme that solves them to certifiable optimality. Further, we derive a novel and often tight class of convex relaxations by decomposing a low-rank matrix as a sum of

rank-one matrices and incentivizing, via a Shor relaxation, that each two-by-two minor in each rank-one matrix has determinant zero. In numerical experiments, our new convex relaxations decrease the optimality gap by two orders of magnitude compared to existing attempts. Moreover, we showcase the performance of our disjunctive branch-and-bound scheme and demonstrate that it solves matrix completion problems over 150x150 matrices to certifiable optimality in hours, constituting an order of magnitude improvement on the state-of-the-art for certifiably optimal methods.

On the Partial Convexification of the Low-Rank Constrained Optimization – WEIJUN XIE, GEORGIA INSTITUTE OF TECHNOLOGY –

The low-rank constrained optimization arises in various machine learning and operations research problems. It minimizes a linear objective function subject to one or multiple two-sided linear matrix inequalities and a low-rank domain set. Although the low-rank constrained optimization is, in general, NP-hard, a viable approach is to convexify the domain set (i.e., replace the domain set with its convex hull), known as “partial convexification.” Partial convexification often leads to a tractable convex relaxation; however, its solution quality lacks theoretical guarantees. To fill this gap, we first establish the necessary and sufficient condition under which the partial convexification matches the original low-rank constrained optimization. Next, we derive an upper bound on the minimum rank among all the optimal solutions of the partial convexification. We prove that the proposed rank upper bound is tight, i.e., it is the best that one could expect. To efficiently solve the partial convexification, we develop a column generation algorithm combined with a rank-reduction algorithm. This combination ensures that the output solution satisfies the theoretical rank upper bound guarantees. Finally, our various numerical experiments validate the strength of the partial convexification and the effectiveness of the proposed algorithms.

A Parametric Approach for Solving Convex Quadratic Optimization with Indicators Over Trees – SALAR FATTAHL, UNIVERSITY OF MICHIGAN – Aaresh Bhathena, Andres Gomez, Simge Kucukyavuz

In this talk, we discuss convex quadratic optimization problems involving indicator variables, each associated with a continuous variable, particularly focusing on scenarios where the matrix defining the quadratic term is positive definite and its sparsity pattern corresponds to the adjacency matrix of a tree graph. We introduce a graph-based dynamic programming algorithm that solves this problem in quadratic time and memory. Central to our algorithm is a precise parametric characterization of the cost function across various nodes of the graph corresponding to distinct variables. Our computational experiments conducted on both synthetic and real-world datasets demonstrate the superior performance of our proposed algorithm compared to existing algorithms and state-of-the-art mixed-integer optimization solvers. An important application of our algorithm is in the real-time inference of Gaussian hidden Markov models from data affected by outlier noise. Using a real on-body accelerometer dataset, we solve instances of this problem with over 30,000 variables in under a minute, and its online variant within milliseconds on a standard computer.

1a - TA40 - 510C - **Advances in Large-Scale Nonlinear Optimization II** Chair: RAGHU BOLLAPRAGADA

First-Order Algorithms for Nonconvex Min-Max Problems with Cohypomonotonicity –

STEPHEN WRIGHT, UNIVERSITY OF WISCONSIN MADISON – Ahmet Alacaoglu, Donghwan Kim

We focus on constrained, L-smooth, nonconvex-nonconcave min-max problems either satisfying rho-cohypomonotonicity or admitting a solution to the rho-weakly Minty Variational Inequality (MVI), where larger values of the parameter rho correspond to a greater degree of nonconvexity. Relevant problem classes include two player reinforcement learning and interaction dominant min-max problems. It has been conjectured that first-order methods can tolerate values of rho no larger than 1/L, but results until now have stagnated at the tighter requirement rho ≤ 0.5/L. We obtain optimal or best-known complexity guarantees with cohypomonotonicity or weak MVI conditions for rho ≤ 1/L, using inexact variants of Halpern and Krasnoselskii-Mann (KM) iterations. We also provide algorithms and complexity guarantees in the stochastic case with the same range on rho. Our improvements come from harnessing the recently proposed “conic nonexpansiveness” property of operators. Finally, we provide a refined analysis for inexact Halpern iteration and propose a stochastic KM iteration with a multilevel Monte Carlo estimator.

Efficient Gradient Tracking Algorithmic Frameworks for Decentralized Optimization

– RAGHU BOLLAPRAGADA, UNIVERSITY OF TEXAS AT AUSTIN – Albert Berahas, Shagun Gupta

Gradient tracking optimization algorithms have received significant attention over the past years for distributed optimization over networks due to their ability to converge to the solution under a constant step size. At every iteration, these algorithms require a computation step of calculating gradients at each node and a communication step to share the iterate and a gradient estimate among nodes. The complexity of these two steps varies significantly across different applications of distributed optimization. In this talk, we present an algorithmic framework that decomposes these two steps, provides flexibility among these steps at each iteration, and is robust to the stochasticity inherent in these steps. We provide optimal theoretical convergence and complexity guarantees, and illustrate its performance on quadratic and classification problems.

Randomized and Stochastic Methods for Fixed-Point and Root-Finding Problems –

QUOC TRAN DINH, DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH, THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL –

In this talk, we present some recent randomized and stochastic algorithms to approximate a solution of fixed-point problems, or equivalently, a solution of root-finding problems. Our algorithms rely on new variants of well-known stochastic variance-reduced estimators, but applying to a different operator. We prove the convergence of these algorithms for two classes of problems: co-hypomonotone and co-coercive settings. We show that our algorithms achieve some improvement in terms of oracle complexity compared to deterministic methods. Our algorithms are also verified via concrete numerical experiments

and comparison to existing methods.

Stochastic Proximal-Gradient Methods with Nonlinear Constraints – DANIEL ROBINSON, LEHIGH UNIVERSITY – Frank E. Curtis, Xiaoyi Qu

We present a stochastic proximal-gradient method for solving regularized optimization problems with nonlinear constraints. The method may also be used when exact (deterministic) function and gradient evaluations are available. A complexity analysis and computational results for the proposed method will be presented that highlight the strengths of our approach.

1a - TA59 - 516A - **Advances in numerical methods for large scale nonlinear optimization and applications II** Chair: ERNESTO BIRGIN

←

Riemannian trust-region methods for strict saddle functions with complexity guarantees – FLORENTIN GOYENS, UNIVERSITY PARIS DAUPHINE PSL – Clément Royer

The difficulty of minimizing a nonconvex function is in part explained by the presence of saddle points, which slows down the algorithm. This also directly impacts the worst-case complexity guarantees. However, a number of nonconvex problems of interest possess a favorable structure for optimization, in the sense that saddle points can be escaped efficiently by an appropriate algorithm. This so-called strict saddle property has been extensively used in data science to derive good properties for first-order algorithms, such as convergence to second-order critical points. However, the analysis and the design of second-order algorithms in the strict saddle setting have received significantly less attention.

We consider second-order trust-region methods for a class of strict saddle functions defined on Riemannian manifolds. These functions exhibit (geodesic) strong convexity around minimizers and negative curvature at saddle points. We show that the standard trust-region method with exact subproblem minimization finds an approximate local minimizer in a number of iterations that depends logarithmically on the accuracy parameter, which significantly improves known results for general nonconvex optimization. We also propose an inexact variant of the algorithm that explicitly leverages the strict saddle property to compute the most appropriate step at every iteration. Our bounds for the inexact variant also improve over the general nonconvex case, and illustrate the benefit of using strict saddle properties within optimization algorithms.

Keywords: Riemannian optimization, strict saddle function, second-order method, complexity guarantees.

A stochastic Hessian-free trust-region method for statistical learning – JEREMY RIEUSSEC, UNIVERSITY OF MONTREAL – Fabian Bastin, Jean Laprès Chartrand

We investigate an adaptive Hessian-free trust-region method for statistical learning. A sampling strategy based on dependant random variables is designed to reduce the sample size at each iteration as much as possible while conserving, in this stochastic setting, the trust-region convergence properties to first-order, and potentially second-order, critical points which are well established for the deterministic case.

Taking advantage of the outer product Hessian approximation with truncated conjugate gradient, this technique is applicable in maximum likelihood estimation as well as in nonlinear least squares regression. We illustrate the potential gains in computation time and memory consumption on various numerical experiments.

Implementation and experiments of an augmented Lagrangian method with first-order information – ERNESTO BIRGIN, UNIVERSITY OF SÃO PAULO –

Algencan is a safeguarded augmented Lagrangian method for minimization with general nonlinear constraints. At each iteration, a bound-constrained subproblem is solved approximately. The approach used for the minimization with bound constraints is an active set strategy. A projected spectral gradient step is used to leave the faces. To minimize inside the faces, essentially any unconstrained minimization method can be used. Currently, taking advantage of the modular structure of augmented Lagrangian methods, we are revisiting first-order matrix-free methods for the minimization inside the faces. The current implementation of Algencan, when second-order information is not available, uses inside the faces truncated Newton directions in which the Newtonian system is solved with conjugate gradients. Inspired by recent work, we analyzed the possibility of using MINRES. Driven by a relatively recent comparison of solvers for bound-constrained minimization, we also considered the possibility of using nonlinear conjugate gradients. The work is in progress and in this talk we show the numerical results we have obtained so far.

Spectral Stochastic Gradient Method with Additional Sampling – ILARIA TROMBINI, UNIVERSITÀ DEGLI STUDI DI FERRARA – Valeria Ruggero, Natasa Krejic

In this talk, we propose a new stochastic gradient method for numerical minimization of finite sums. We also propose a modified version of this method applicable on more general problems referred to as infinite sum problems, where the objective function is in the form of mathematical expectation. The method is based on a strategy to exploit the effectiveness of the well-known Barzilai-Borwein (BB) rules [1] or variants of these (BB-like rules) for updating the step length in the standard gradient method. Very effective improvements with respect to the standard BB rules have been obtained using the Adaptive Barzilai-Borwein (ABB) strategy [2] and its modification ABBmin [3]. Starting from these observations and following the approaches developed in Algorithms LSOS [4] and SLiSeS [5], the proposed method adapts the aforementioned strategy into the stochastic framework by exploiting the same SAA estimator of the objective function for several iterations. Furthermore, the sample size is controlled by an additional sampling which also plays a role in accepting the proposed iterate point. Moreover, the number of "inner" iterations with the same sample is also controlled by an adaptive rule which prevents the method from getting stacked with the same estimator for too long. The method is described and analyzed first for the minimization of the finite sum case, giving rise to the scheme named LSNM-BB; then, it is generalized to the more general case, leading to the version named LSNM-BB-G. Convergence results are discussed for the finite and infinite sum version, for general and strongly convex objective functions. Numerical experiments on well-known datasets for binary classifications show very promising

performance of the method, without the need to provide special values for hyperparameters on which the method depends. We tested on the datasets w8a, IJCNN, RCV1 and MNIST for the finite sum problem and we simulate the infinite sum problem with the dataset CIFAR10.

[1] IMA J.Numer.Anal., 8, 141–148 (1988). [2] Comput. Optim. Appl. 35(1), 69–86 (2006). [3] J. Ind. Manag. Optim. 4(2), 299–312 (2008). [4] LSOS: Line-search Second-Order Stochastic optimization methods for nonconvex finite sums (2022). [5] SLiSeS: Subsampled Line Search Spectral Gradient Method for Finite Sums (2023).

1a - TB60 - 516D - **Advances in numerical methods for large scale nonlinear optimization and applications III** Chair: STEPHANIA BELLAVIA ↩

Exploiting Negative Curvature in Conjunction with Adaptive Sampling: Theoretical Results and a Practical Algorithm – WANPING DONG, UNIVERSITY OF MICHIGAN –

With the emergence of deep learning in recent years, nonconvex optimization has gained more interest and focus, but its nonconvex properties bring us great challenges in algorithm design. Our goal is to develop algorithms that have second-order convergence and affordable complexity so that they can be used for large-scale problems. Therefore, we'll exploit negative curvature with adaptive sampling strategy. First, we provide the convergence result for the deterministic setting where the gradients and Hessians are inexact. In the second part, I want to show our result in the stochastic setting where the gradients are constructed by random samples. Except for the theory part, I'll also talk about our practical algorithm which is the Newton-CG method with negative curvature detection and adaptive sampling. Our experiment results on different nonconvex problems are shown after that.

Solving large-scale nonlinear least-squares with random Gauss-Newton models – STEPHANIA BELLAVIA, UNIVERSITY OF FLORENCE – Greta Malaspina, Benedetta Morini

We address the solution of large-scale nonlinear least-squares problems by stochastic Gauss-Newton methods combined with a line-search strategy. The algorithms proposed have per-iteration computational complexity lower than classical deterministic methods, due to the employment of random models inspired by randomized linear algebra tools. Under suitable assumptions, the stochastic optimization procedures can achieve a desired level of accuracy in the first-order optimality condition. We discuss the construction of the random models and the iteration complexity results to drive the gradient below a prescribed accuracy, then we present results from our computational experience.

A stochastic gradient method with variance control and variable learning rate for Deep Learning – GIORGIA FRANCHINI, UNIMORE –

In this talk we present a stochastic gradient algorithm which rules the increase of the mini-batch size in a predefined fashion and automatically adjusts the learning rate by means of a monotone or non-monotone line search procedure. The mini-batch size is incremented at a suitable a priori rate throughout

the iterative process in order that the variance of the stochastic gradients is progressively reduced. The a priori rate is not subject to restrictive assumptions, allowing for the possibility of a slow increase in the mini-batch size. On the other hand, the learning rate can non-monotonically vary along the iterations, as long as it is appropriately bounded. Convergence results for the proposed method are provided for both convex and non convex objective functions. Moreover it can be proved that the algorithm enjoys a global linear rate of convergence on strongly convex functions. The low per-iteration cost, the limited memory requirements and the robustness against the hyperparameters setting make the suggested approach well-suited for implementation within the deep learning framework, also for GPGPU-equipped architectures. Numerical results on training deep neural networks for multi-class image classification show a promising behaviour of the proposed scheme with respect to similar state of the art competitors.

Probabilistic Trust Region method for solving Multi-objective Problems – LUKA RUTESIC, UNIVERSITY OF NOVI SAD, FACULTY OF SCIENCES – Natasa Krejic, Natasa Krklec Jerinkic

The problem considered is a multi-objective optimization problem, in which the goal is to find an optimal value of a vector function representing various criteria. An algorithm which utilizes trust region framework with probabilistic model functions able to cope with noisy problems and approximate functions and their derivatives is derived and analysed. We prove the almost sure convergence of the proposed algorithm to a Pareto critical point if the model functions are good approximations in probabilistic sense. Numerical results demonstrate effectiveness of the probabilistic trust region by comparing it to competitive stochastic multi-objective solvers. The application in supervised machine learning is showcased by training non discriminatory Logistic Regression models on different size data groups. Additionally, we use several test examples with irregularly shaped fronts to exhibit the efficiency of the algorithm.

1a - TB67 - 510C - **Numerical Optimization for Machine Learning 1** Chair: FREDERIK KUNSTNER ↩

Relative Monte Carlo for Reinforcement Learning – AUDREY BAZERGH, NORTHWESTERN UNIVERSITY –

We introduce relative Monte Carlo (rMC), a new general purpose policy gradient algorithm for reinforcement learning with discrete action space. The policy is improved in real time using relative returns between a root sample path and counterfactual simulated paths instantiated by taking a different action from the root. The method is compatible with any differentiable policy, including the leading choice of neural network parametrization. It is guaranteed to converge for episodic and average reward tasks. Unlike traditional Monte Carlo, rMC policy gradient steps are performed throughout the rollout using a memory-efficient update decomposition, inspired by eligibility traces. Strong couplings between root and counterfactual paths further contribute to low data generation and memory requirements. We test rMC with a policy network in a two-tiered inventory fulfillment problem. Numerical results show it performs well compared to related algorithms.

Structured Inverse-free Matrix Adaptive Methods for Large Neural Networks – WU LIN, VECTOR INSTITUTE – Felix Dangel, Runa Eschenhagen, Juhan Bae, Richard E. Turner, Alireza Makhzani

Optimization is an essential ingredient of deep learning. Many optimization problems can be reformulated from a probabilistic perspective to exploit the Fisher-Rao geometric structure of a probability family. In this talk, we show how to design new quasi-Newton methods for large-scale neural network (NN) training by leveraging the geometric structure. We first establish a second-order view of adaptive methods such as RMSProp and full-matrix AdaGrad when the square root is removed from their preconditioned gradient step. We then introduce and exploit preconditioner invariance to make such square-root-free (non-diagonal) adaptive methods inverse-free while maintaining their preconditioner structures for modern low-precision training with mini-batches. Finally, we propose Kronecker-factored adaptive methods without the root to bridge the computation gap between non-diagonal and diagonal adaptive methods, and demonstrate advantages of our methods over square-root-based counterparts like Shampoo for training large NNs such as Transformers and Mambas in half-precision by removing numerically unstable matrix square-root decompositions and inversions.

Target-based Surrogates for Stochastic Optimization – JONATHAN LAVINGTON, UNIVERSITY OF BRITISH COLUMBIA – Vaswani Sharan, Babanezhad Reza, Mark Schmidt, Nicolas Le Roux

We consider minimizing functions for which it is expensive to compute the (possibly stochastic) gradient. Such functions are prevalent in reinforcement learning, imitation learning and adversarial training. Our target optimization framework uses the (expensive) gradient computation to construct surrogate functions in a target space (e.g. the logits output by a linear model for classification) that can be minimized efficiently. This allows for multiple parameter updates to the model, amortizing the cost of gradient computation. In the full-batch setting, we prove that our surrogate is a global upper-bound on the loss, and can be (locally) minimized using a black-box optimization algorithm. We prove that the resulting majorization-minimization algorithm ensures convergence to a stationary point of the loss. Next, we instantiate our framework in the stochastic setting and propose the SSO algorithm, which can be viewed as projected stochastic gradient descent in the target space. This connection enables us to prove theoretical guarantees for SSO when minimizing convex functions. Our framework allows the use of standard stochastic optimization algorithms to construct surrogates which can be minimized by any deterministic optimization method. To evaluate our framework, we consider a suite of supervised learning and imitation learning problems. Our experiments indicate the benefits of target optimization and the effectiveness of SSO.

Stochastic Extragradient Methods: New Convergence Guarantees for Monotone and Non-monotone Variational Inequalities – NICOLAS LOIZOU, JOHNS HOPKINS UNIVERSITY –

Stochastic extragradient methods (SEG) have garnered significant interest in recent years and rank among the most efficient algorithms for solving large-scale min-max optimization and variational inequalities problems (VIP) that occur in various machine learning tasks. Despite their undeniable popularity,

current convergence analyses of SEG require strong assumptions like bounded variance or growth conditions. Moreover, several important questions regarding the convergence properties of these methods remain unanswered, including mini-batching, efficient step-size selection, and convergence guarantees under various sampling strategies. In this talk, we address these questions and provide novel convergence guarantees for several variants of SEG, including Same-sample SEG (S-SEG), Stochastic Past Extragradient (SPEG), and SEG with Random Reshuffling (SEG-RR). We complement our convergence guarantees with experiments that empirically validate our theoretical findings.

1a - TB69 - 516A - From First- to High-order Methodologies: Recent Advances and Trends II Chair: YANG LIU

Tensor Methods for Nonconvex Optimization using Cubic-quartic regularization models – KATE WENQI ZHU, UNIVERSITY OF OXFORD –

High-order tensor methods for solving both convex and nonconvex optimization problems have recently generated significant research interest, due in part to the natural way in which higher derivatives can be incorporated into adaptive regularization frameworks, leading to algorithms with optimal global rates of convergence and local rates that are faster than Newton's method. On each iteration, to find the next solution approximation, these methods require the unconstrained local minimization of a (potentially nonconvex) multivariate polynomial of degree higher than two, constructed using third-order (or higher) derivative information, and regularized by an appropriate power of the change in the iterates. Developing efficient techniques for the solution of such subproblems is currently, an ongoing topic of research, and this talk addresses this question for the case of the third-order tensor subproblem. In particular, we propose the CQR algorithmic framework, for minimizing a nonconvex Cubic multivariate polynomial with Quartic Regularisation, by sequentially minimizing a sequence of local quadratic models that also incorporate both simple cubic and quartic terms. The role of the cubic term is to crudely approximate local tensor information, while the quartic one provides model regularization and controls progress. We provide necessary and sufficient optimality conditions that fully characterise the global minimizers of these cubic-quartic models. We then turn these conditions into secular equations that can be solved using nonlinear eigenvalue techniques. We show, using our optimality characterisations, that a CQR algorithmic variant has the optimal-order evaluation complexity of $O(\epsilon^{-3/2})$ when applied to minimizing our quartically-regularised cubic subproblem, which can be further improved in special cases. We propose practical CQR variants that judiciously use local tensor information to construct the local cubic-quartic models. We test these variants numerically and observe them to be competitive with ARC and other subproblem solvers on typical instances and even superior on ill-conditioned subproblems with special structure.

A consistently adaptive trust-region method – FADI HAMAD, UNIVERSITY OF PITTSBURGH – Oliver Hinder

Adaptive trust-region methods attempt to maintain strong convergence guarantees without depending on conservative es-

timates of problem properties such as Lipschitz constants. However, on close inspection, one can show existing adaptive trust-region methods have theoretical guarantees with severely suboptimal dependence on problem properties such as the Lipschitz constant of the Hessian. For example, TRACE developed by Curtis et al. obtains a $O(\Delta_f L^{3/2} \epsilon^{-3/2}) + \tilde{O}(1)$ iteration bound where L is the Lipschitz constant of the Hessian. Compared with the optimal $O(\Delta_f L^{1/2} \epsilon^{-3/2})$ bound this is suboptimal with respect to L . We present the first adaptive trust-region method which circumvents this issue and requires at most $O(\Delta_f L^{1/2} \epsilon^{-3/2}) + \tilde{O}(1)$ iterations to find an ϵ -approximate stationary point, matching the optimal iteration bound up to an additive logarithmic term. Our method is a simple variant of a classic trust-region method and in our experiments performs competitively with both ARC and a classical trust-region method.

A randomized algorithm for nonconvex minimization with inexact evaluations and complexity guarantees – SHUYAO LI, UNIVERSITY OF WISCONSIN MADISON – Stephen Wright

We consider minimization of a smooth nonconvex function with inexact oracle access to gradient and Hessian (without assuming access to the function value) to achieve approximate second-order optimality. A novel feature of our method is that if an approximate direction of negative curvature is chosen as the step, we choose its sense to be positive or negative with equal probability. We allow gradients to be inexact in a relative sense and relax the coupling between inexactness thresholds for the first- and second-order optimality conditions. Our convergence analysis includes both an expectation bound based on martingale analysis and a high-probability bound based on concentration inequalities. We apply our algorithm to empirical risk minimization problems and obtain improved gradient sample complexity over existing works.

A MINRES-based Algorithm Framework for Unconstrained Nonconvex Optimization with Non-positive Curvature Detection – HANFENG ZENG, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Yang Liu, Wenqing Ouyang, Andre Milzarek

In this talk, we present a MINRES-based algorithm framework for unconstrained nonconvex optimization problems. Our approach utilizes the minimal residual method (MINRES), a well-known solver for indefinite symmetric linear systems, to compute descent directions that leverage higher order and non-positive curvature (NPC) information. We derive and discuss asymptotic convergence properties in both linesearch and trust region settings. Moreover, under the Kurdyka-Lojasiewicz property, we prove that the proposed approach can avoid strict saddle points and converges to second order optimal points. This is primarily achieved by using proper regularizations of the Newton system and forward linesearch along NPC directions. Finally, we consider potentially indefinite variants of several limited memory quasi-Newton matrices (including BFGS, DFP, SR1 and PSB updates) and show how such updates can be applied and incorporated in our algorithm framework. Numerical experiments on the CUTEst test collection and on a deep auto-encoder problem illustrate the efficiency of the proposed methodology.

1a - TB136 - 511C - First-order Methods and Large-scale Optimization I Chair: LEXIAO LAI →

Activity Identification in Non-smooth Optimization: A Set-Valued Operator Perspective

– ZIQI QIN, SHANGHAI JIAO TONG UNIVERSITY – Jingwei Liang

In the last ten years, the concept of partial smoothness has played a crucial role in various areas such as local convergence analysis, sensitivity analysis, and algorithm acceleration. However, one significant limitation of partial smoothness is its applicability only to functions in finite dimensions. In this presentation, I aim to address this limitation by introducing an extension of partial smoothness from functions to set-valued operators. By extending partial smoothness to set-valued operators, we can achieve a more comprehensive understanding and analysis of these operators. This extension enables us to gain an intuitive geometric interpretation of finite activity identification in proximal operator splitting algorithms. Specifically, we can identify the subsets of the domain where the set-valued operator exhibits finite activity, which has important implications for the convergence behavior and computational efficiency of proximal operator splitting algorithms. Additionally, this extension allows us to provide an upper bound for the finite activity property, whereas existing results have only considered the lower bound. By obtaining both upper and lower bounds, we can obtain a more complete characterization of the finite activity property in proximal operator splitting algorithms. This comprehensive understanding can lead to improved algorithm design and performance in practical optimization problems. Furthermore, the extension of partial smoothness to set-valued operators also allows us to derive identification results for degenerate problems. Degeneracy often arises in optimization problems when certain constraints or variables become redundant or highly correlated. By incorporating the concept of partial smoothness into the analysis of degenerate problems, we can gain insights into the identification of critical points and the behavior of optimization algorithms in such scenarios. In summary, the extension of partial smoothness from functions to set-valued operators opens up new avenues for analysis and understanding in optimization. It provides a geometric interpretation of finite activity identification, enables the derivation of upper bounds, and facilitates the analysis of degenerate problems. These insights can contribute to the development of more efficient and effective optimization algorithms in both theory and practice.

Goldstein modulus in Lipschitz Minimization

– SIYU KONG, CORNELL UNIVERSITY – Adrian Lewis

In 1977, Goldstein proposed a simple conceptual procedure for minimizing Lipschitz functions. From each current iterate x , we move a fixed step length in the direction opposite to the “Goldstein subgradient” - the shortest convex combination of gradients at points within that step length of x . How to choose the step length is unclear, and in any case, the Goldstein subgradient is not computable in practice. However, recent clever approximation schemes have resulted in practical versions with complexity bounds, and even near-linear convergence on large classes of objectives. We introduce a robust measure of the slope of a Lipschitz function, based on Goldstein subgradients. By exploring and leveraging linear growth of this “Goldstein

modulus” around minimizers, we propose a conceptual subgradient algorithm with adapted radius selection, and thus explain the phenomenon of near-linear convergence of subgradient methods.

Convergence of heavy-ball SGD in the nonconvex case: a novel time window-based analysis – BOHAO MA, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Junwen Qiu, Andre Milzarek

In this talk, we propose a novel time window-based analysis technique to investigate the convergence behavior of the stochastic heavy-ball method (SHB) in nonconvex settings. Despite its popularity, the asymptotic convergence properties of SHB remain less understood in nonconvex scenarios. This is primarily due to the absence of sufficient descent and challenges in controlling stochastic errors in an almost sure sense. To address these challenges, we analyze the potential approximate descent of SHB across specific time windows - rather than examining the behavior between consecutive iterates as done in classical studies. This time window-based approach simplifies the convergence analysis and enables us to establish the first iterate convergence result for SHB under the Kurdyka-Lojasiewicz (KL) property. Depending on the underlying step size dynamics and the KL exponent, we further characterize the local convergence rates of SHB.

Error bounds, PL condition, and quadratic growth for weakly convex functions, and linear convergences of proximal point methods – FENG YI LIAO, UNIVERSITY OF CALIFORNIA, SAN DIEGO, ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT – Lijun Ding, Yang Zheng

Many machine learning problems lack strong convexity properties. Fortunately, recent studies have revealed that first-order algorithms also enjoy linear convergences under various weaker regularity conditions. While the relationship among different regularity conditions for convex and smooth functions is well understood, it is not the case for the nonsmooth setting. In this talk, we go beyond convexity and smoothness, and clarify the connections among common regularity conditions (including strong convexity, restricted secant inequality, subdifferential error bound, Polyak-Lojasiewicz inequality, and quadratic growth) in the class of weakly convex functions. Our analysis is simple and easy to follow. We use the notion of slope and Ekeland’s variational principle as the key tools to establish the relationships among those common regularity conditions. In addition, we present a simple and modular proof for the linear convergence of the proximal point method (PPM) for convex (possibly nonsmooth) optimization using these regularity conditions. Our convergence results require weaker conditions while having simpler proofs than many existing characterizations. Linear convergence also holds when the subproblems of PPM are solved inexactly with a proper control of inexactness. We expect practical applications of these fast linear convergences of PPM to large-scale, sparse conic optimization with the connection to the augmented Lagrangian method.

1a - WA196 - 516D - **Nonlinear stochastic optimization** Chair: MIAOLAN XIE

Distributed Composite Stochastic Mirror Descent for Stochastic Optimization – YAO JI,

PURDUE UNIVERSITY –

We study statistical estimation of high-dimensional sparse parameters over a network. The estimation problem is formulated as Stochastic Approximation. Assuming agents only have access to unbiased estimates of the gradients of their local cost functions, we propose composite stochastic optimization problem on each stage of a multistage algorithm; each stage being solved to a prescribed accuracy by the non-Euclidean composite distributed stochastic mirror descent method (DSMD) in the adapt than combine (ATC) and combine than adapt (CTA) forms. We show that, with high probability, the iterates generated by each agent linearly converge during the first “preliminary” phase of the routine and is sublinear during the second “asymptotic” phase. Further, we characterize the transient time needed for DSMD to approach the asymptotic convergence rate. Numerical experiments demonstrate the tightness of the theoretical results.

Stochastic Adaptive Regularization Method with Cubics - A High Probability Complexity Bound – MIAOLAN XIE, CORNELL UNIVERSITY –

We present a high probability complexity bound for a stochastic adaptive regularization method with cubics, also known as regularized Newton method. The method makes use of stochastic zeroth-, first- and second-order oracles that satisfy certain accuracy and reliability assumptions. Such oracles have been used in the literature by other stochastic adaptive methods, such as trust region and line search. These oracles capture many settings, such as expected risk minimization, stochastic zeroth-order optimization, and others. In this paper, we give the first high probability iteration bound for stochastic cubic regularization, and show that just as in the deterministic case, it is superior to other stochastic adaptive methods.

Design Guidelines of Noise-Tolerant Optimization Methods – YUCHEN LOU, NORTHWESTERN UNIVERSITY – Jorge Nocedal

The development of nonlinear optimization algorithms capable of performing reliably in the presence of noise has garnered considerable attention lately. This work advocates for strategies to create noise-tolerant nonlinear optimization algorithms by adapting classical deterministic methods. These adaptations follow certain design guidelines described here, which make use of estimates of the noise level in the problem. The application of our methodology is illustrated by the development of a line search gradient projection method and a robust quasi-Newton method. The algorithms are tested on an engineering design problem, where the propagation of noise is unknown. It is shown that a new self-calibrated line search, noise-aware finite-difference techniques, and a lengthening L-BFGS-B method can be effective even in the high noise regime. Numerical experiments investigate the resiliency of key algorithmic components. Finally, convergence analysis of the proposed methods establishes convergence to a neighborhood of the solution.

Robustly Learning Single-Index Models via Alignment Sharpness – PUQIAN WANG, UNIVERSITY OF WISCONSIN, MADISON – Zarifis Nikos, Ilias Diakonikolas, Jelena Diakonikolas

We study the problem of learning Single-Index Models under the L_2^2 loss in the agnostic model. We give an efficient learning algorithm, achieving a constant factor approximation to the optimal loss, that succeeds under a range of distributions (includ-

ing log-concave distributions) and a broad class of monotone and Lipschitz link functions. This is the first efficient constant factor approximate agnostic learner, even for Gaussian data and for any nontrivial class of link functions. Prior work for the case of unknown link function either works in the realizable setting or does not attain constant factor approximation. The main technical ingredient enabling our algorithm and analysis is a novel notion of a local error bound in optimization that we term alignment sharpness and that may be of broader interest.

1a - WA251 - 510C - **Algebraic Methods in Nonlinear Programming** Chair: ALI MOHAMMAD NEZHAD 

Convex Network Flows – THEO DIAMANDIS, MIT – Guillermo Angeris, Alan Edelman

We introduce a general framework for flow problems over hypergraphs. In our problem formulation, which we call the convex flow problem, we have a concave utility function for the net flow at every node and a concave utility function for each edge flow. The objective is to minimize the sum of these utilities, subject to constraints on the flows allowed at each edge, which we only assume to be a convex set. This framework not only includes many classic problems in network optimization, such as max flow, min-cost flow, and multi-commodity flows, but also generalizes these problems to allow, for example, concave edge gain functions. In addition, our framework includes applications spanning a number of fields: optimal power flow over lossy networks, routing and resource allocation in ad-hoc wireless networks, Arrow-Debreu Nash bargaining, and order routing through financial exchanges, among others. We show that the convex flow problem has a dual with a number of interesting interpretations, and that this dual decomposes over the edges of the hypergraph. Using this decomposition, we propose a fast solution algorithm that parallelizes over the edges and admits a clean problem interface. We provide an open source implementation of this algorithm in the Julia programming language, which we show is significantly faster than the state-of-the-art commercial convex solver Mosek on the optimal power flow problem and the decentralized exchange order routing problem.

A full splitting algorithm for fractional programs with structured numerators and denominators – MIN TAO, NANJING UNIVERSITY – Radu Ioan Bot, Guoyin Li

In this paper, we consider a class of nonconvex and nonsmooth fractional programming problems, which involve the sum of a convex, possibly nonsmooth function composed with a linear operator and a differentiable, possibly nonconvex function in the numerator and a convex, possibly nonsmooth function composed with a linear operator in the denominator. These problems have applications in various fields, including CT reconstruction and sparse signal recovery. We propose an adaptive full-splitting proximal subgradient algorithm with an extrapolated step that addresses the challenge of evaluating the composition in the numerator by decoupling the linear operator from the nonsmooth component. We specifically evaluate the nonsmooth function using its proximal operator, while the linear operator is assessed through forward evaluations. Furthermore, the smooth component in the numerator is evaluated

through its gradient, the nonsmooth component in the denominator is managed using its subgradient, and the linear operator in the denominator is also assessed through forward evaluations. We demonstrate subsequential convergence toward an approximate lifted stationary point and ensure global convergence under the Kurdyka-Lojasiewicz property, all achieved *without relying on any full-row rank assumptions regarding the linear operators*. We further explain the reasoning behind aiming for an approximate lifted stationary point. This is exemplified by constructing a scenario illustrating that the algorithm could diverge when seeking exact solutions. Lastly, we present a practical iteration of the algorithm incorporating a nonmonotone line search, significantly enhancing its convergence performance. Our theoretical findings are validated through simulations involving limited-angle CT reconstruction and the robust sharp ratio minimization problem.

O-minimal structures and central paths in nonlinear optimization – ALI MOHAMMAD NEZHAD, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL – Saugata Basu

A central path lies at the heart of interior point methods for nonlinear optimization. We prove quantitative results about central paths in nonlinear optimization with definable functions in an o-minimal structure over $\mathbb{R}$. Using results from singularity theory and o-minimal geometry, we establish conditions for the existence, convergence, and also smoothness of a central path at the limit point.

On describing convex hull of quadratic sets via aggregations – SHENG DING SUN, UNIVERSITY OF CAMBRIDGE – Grigoriy Blekherman, Santanu Dey

Obtaining efficient and concise description for convex hull of sets defined by quadratic inequalities is a fundamental task in optimization, often appearing as a subroutine (e.g. trust region subproblem) of some larger problem. A particular case is when the convex hull can be obtained by taking certain nonnegative linear combinations of the defining constraints, which we call aggregations. For arbitrary number of (open) quadratic constraints, we propose the condition of hidden hyperplane convexity (HHC) which is sufficient for special aggregations to describe the convex hull. Under a further condition we show that the number of aggregations needed in the description is at most quadratic in the number of constraints. This answers a question asked by Dey, Munoz and Serrano. We also discuss a condition under which our results generalizes to closed constraints. We do not need the assumption that all connected components are full dimensional, which was needed in the work by Modaresi and Vielma.

1a - WA284 - 516A - **Theory of constrained nonlinear optimization** Chair: YANGYANG XU 

Complexity Guarantees for Optimal Nash Equilibrium Seeking and Bilevel Variational Inequalities – FARZAD YOUSEFIAN, RUTGERS UNIVERSITY –

In noncooperative game theory, understanding the quality of Nash equilibrium has been a storied area of research and has led to the emergence of popular metrics including the Price of Anarchy and the Price of Stability. The evaluation of these ratios is complicated by the need to compute the worst and the best

equilibrium, respectively. In this talk, we present a class of iteratively regularized extragradient methods with performance guarantees for computing the optimal equilibrium. To this end, we consider optimization problems with merely monotone variational inequality (VI) constraints when the objective function is either (i) merely convex, (ii) strongly convex, or (iii) nonconvex. In (i), we considerably improve the existing complexity guarantees. In (ii) and (iii), we derive new complexity guarantees for solving this class of problems. Notably, this appears to be the first time that nonconvex optimization with monotone VI constraints is addressed with complexity guarantees. We further show that our results in (i) and (ii) can be generalized to address a class of bilevel VIs that subsumes challenging problem classes such as the generalized Nash equilibrium problem. We will then discuss extensions to stochastic and large-scale settings. We validate the theoretical findings using preliminary numerical experiments for computing the best and the worst Nash equilibria.

Global Resolution of Chance-Constrained Optimization Problems: Minkowski Functionals and Monotone Inclusions – UDAY SHANBHAG, PENNSYLVANIA STATE UNIVERSITY –

Chance-constrained optimization problems, an important subclass of stochastic optimization problems, are often complicated by nonsmoothness, and nonconvexity. Thus far, non-asymptotic rates and complexity guarantees for computing an approximate global minimizer remain open questions. We consider a subclass of problems in which the probability is defined as $P\{\zeta | \zeta \in K(x)\}$, where K is a set defined as $K(x) = \{\zeta \in K | c(x, \zeta) \leq 1\}$, $c(x, \cdot)$ is a positively homogenous function for any $x \in X$, and K is a nonempty and convex set, symmetric about the origin. We make two contributions in this context. (i) First, when ζ admits a log-concave density on K , the probability function is equivalent to an expectation of a nonsmooth Clarke-regular integrand, allowing for the chance-constrained problem to be restated as a convex program. Under a suitable regularity condition, the necessary and sufficient conditions of this problem are given by a monotone inclusion with a compositional expectation-valued operator. (ii) Second, under differing requirements on ζ , we present a variance-reduced proximal scheme and provide amongst the first rate and complexity guarantees for resolving chance-constrained optimization problems. This is joint work with Peixuan Zhang (Penn State), Constantino Lagoa (Penn State), and Ibrahim Bardakci (Bartın University, Turkey).

Breaking the cycle: Deterministic block-iterative analysis for the Frank-Wolfe algorithm – ZEV WOODSTOCK, ZUSE INSTITUTE BERLIN (ZIB) – Gábor Braun, Sebastian Pokutta

To-date, Block-coordinate Frank-Wolfe (BCFW) algorithms are proven to converge under restrictive assumptions on how the block-coordinates are updated, namely, (A) single-coordinate updates (via stochastic and greedy selection), and (B) cyclic/full update schemes which require the evaluation of all LMOs (linear minimization oracles) in the Cartesian product for every outer iteration. These requirements are particularly disadvantageous when one LMO is far more computationally expensive than the others (e.g., requiring an eigencomputation or a large LP solve). Drawing inspiration from the literature on prox-based algorithms, we show that convergence for BCFW is guaranteed under a much milder assumption on

the updates. This allows one to activate the expensive LMO(s) less frequently, which also leads to concrete gains in algorithmic performance.

Damped Proximal Augmented Lagrangian Method for weakly-Convex Problems with Convex Constraints – YANGYANG XU, RENSSELAER POLYTECHNIC INSTITUTE – Hari Dahal, Wei Liu

In this talk, I will give a damped proximal augmented Lagrangian method (DPALM) for solving problems with a weakly-convex objective and convex linear/nonlinear constraints. Instead of taking a full stepsize, DPALM adopts a damped dual stepsize to ensure the boundedness of dual iterates. DPALM can produce a (near) eps-KKT point within eps^{-2} outer iterations if each DPALM subproblem is solved to a proper accuracy. In addition, I will show overall iteration complexity of DPALM when the objective is either a regularized smooth function or in a regularized compositional form. For the former case, DPALM achieves the complexity of $\text{eps}^{-2.5}$ to produce an eps-KKT point by applying an accelerated proximal gradient (APG) method to each DPALM subproblem. For the latter case, the complexity of DPALM is eps^{-3} to produce a near eps-KKT point by using an APG to solve a Moreau-envelope smoothed version of each subproblem. Our outer iteration complexity and the overall complexity either generalize existing best ones from unconstrained or linear-constrained problems to convex-constrained ones, or improve over the best-known results on solving the same-structured problems. Furthermore, numerical experiments on linearly/quadratically constrained non-convex quadratic programs and linear-constrained robust nonlinear least squares are conducted to demonstrate the empirical efficiency of the proposed DPALM over several state-of-the-art methods.

1a - WB41 - 516D - Advances in Large-Scale Nonlinear Optimization III Chair: BAORYU ZHOU



Dynamic optimal combination of forecasting models – FABIAN BASTIN, UNIVERSITÉ DE MONTRÉAL – Clémence Granade, Stéphane Dellacherie, Alexandre Deslauriers

In many real-world forecasting problems, we have access to several error-minimization models whose performance may depend on the current state of the studied system. It is often possible to improve prediction performance by merging these models together in an appropriate way, which depends on the system conditions. Therefore, the optimal model combination must be dynamically updated according to the available data and target.

We review various approaches that allow us to reduce prediction error, measured by a given performance indicator such as the root mean-squared error, and formulate a general mathematical programming framework to dynamically find the optimal combination of forecasting models with respect to the chosen indicator. We then analyze the mathematical properties of the resulting optimization problem and the solution techniques required to obtain solutions. Since the data distribution is expected to evolve over time, we further discuss how to efficiently construct the training sample along with the training and validation samples.

We illustrate the proposed methodology using short-

term electricity demand forecasting, relying on real data and machine-learning models provided by Hydro-Québec.

On the global convergence of a decomposition algorithm for nonconvex two-level problems – ANDREAS WAECHTER, NORTHWESTERN UNIVERSITY – Yuchen Lou, Ermin Wei

We discuss the global convergence of a decomposition algorithm for nonlinear nonconvex continuous two-stage problems. To render the second-stage response function smooth, a smoothing technique based on a barrier function is used so that standard nonlinear second-order optimization methods can be utilized for both the first-stage and the second-stage problems. The challenge in this context is the existence of local minima and other stationary points in the subproblem besides a global minimum. We introduce a new definition of a local minimum and show how global convergence to such points can be proven.

Modified Line Search Sequential Quadratic Methods for Equality-Constrained Optimization with Unified Global and Local Convergence Guarantees – JIAHAO SHI, UNIVERSITY OF MICHIGAN – Albert Berahas, Raghu Bollapragada

In this paper, we propose a method that has foundations in the line search sequential quadratic programming paradigm for solving general nonlinear equality constrained optimization problems. The method employs a carefully designed modified line search strategy that utilizes second-order information of both the objective and constraint functions, as required, to mitigate the Maratos effect. Contrary to classical line search sequential quadratic programming methods, our proposed method is endowed with global convergence and local superlinear convergence guarantees. Moreover, we extend the method and analysis to the setting in which the constraint functions are deterministic but the objective function is stochastic or can be represented as a finite-sum. We also design and implement a practical inexact matrix-free variant of the method. Finally, numerical results illustrate the efficiency and efficacy of the method.

Comparing stochastic oracles arising in robust gradient estimation – KATYA SCHEINBERG, CORNELL UNIVERSITY –

We will discuss several robust gradient estimators, such as median of means, and the properties of the stochastic oracles based on these estimators. We will demonstrate how these properties fit into convergence analysis of stochastic optimization methods that rely on such oracles and we will compare their performance with the standard stochastic gradient based on a minibatch.

1a - WB68 - 510C - Numerical Optimization for Machine Learning 2 Chair: MARK SCHMIDT ↩

Why Line-Search When You Can Plane-Search? – BETTY SHEA, UNIVERSITY OF BRITISH COLUMBIA –

The practical performance of an optimization method depends on details such as using good step sizes. Strategies for setting step sizes are generally limited to hyperparameter tuning (for a fixed step size), step size schedules and line searches. For many common machine learning problems, line optimization and subspace optimization find accurate step sizes for asymp-

totically the same cost as using a fixed step size. In some cases, line optimization may find step sizes that are ruled out by the standard Armijo condition. For optimization methods that use multiple search directions, such as gradient descent with momentum, using subspace optimization instead of fixed step size schedules allow for better adaptivity and potentially faster convergence. In the case of some neural networks, subspace optimization allows the use of different step sizes for different layers that could decrease the amount of training time needed, as well as reducing the dependence on hyperparameter tuning.

Searching for optimal per-coordinate step-sizes with multidimensional backtracking – FREDERIK KUNSTNER, UBC –

The backtracking line-search is an effective technique to automatically tune the step-size in smooth optimization. It guarantees similar performance to using the theoretically optimal step-size. Many approaches have been developed to instead tune per-coordinate step-sizes, also known as diagonal preconditioners, but none of the existing methods are provably competitive with the optimal per-coordinate step-sizes. We propose multidimensional backtracking, an extension of the backtracking line-search to find good diagonal preconditioners for smooth convex problems. Our key insight is that the gradient with respect to the step-sizes, also known as hypergradients, yields separating hyperplanes that let us search for good preconditioners using cutting-plane methods. As black-box cutting-plane approaches like the ellipsoid method are computationally prohibitive, we develop an efficient algorithm tailored to our setting. Multidimensional backtracking is provably competitive with the best diagonal preconditioner and requires no manual tuning.

When Online Learning Meets Stochastic Calculus – VICTOR SANCHES PORTELLA, UNIVERSITY OF BRITISH COLUMBIA – Nick Harvey, Christopher Liaw

Online learning is a theoretical framework for learning and optimization under adversarial circumstances. Optimization methods developed in this setting are usually robust and have formal notions of beyond worst-case (or adaptive) performance. Often, the development and analysis of such methods can often be challenging and unintuitive. A recent line of work has looked at online learning through the lens of differential equations and continuous-time analysis. This viewpoint has yielded optimal results for several problems in online learning. In this talk I will overview a few uses of stochastic calculus as a guide in the design and analysis of online learning methods, with a particular focus on the fundamental problem of prediction with experts' advice.

Don't be so Monotone: Relaxing Stochastic Line Search in Over-Parameterized Models – LEONARDO GALLI, LMU MUNICH – Holger Rauhut, Mark Schmidt

Recent works have shown that line search methods can speed up SGD and Adam in modern over-parameterized settings. However, existing line searches may take steps that are smaller than necessary since they require a monotone decrease of the (mini-)batch objective function. We explore nonmonotone line search methods to relax this condition and possibly accept larger step sizes. Despite the lack of a monotonic decrease,

we prove the same fast rates of convergence as in the monotone case. These are the first rates proving the convergence of nonmonotone line search methods in the stochastic setting under interpolation. Our experiments show that nonmonotone methods improve the speed of convergence and generalization properties of SGD/Adam even beyond the previous monotone line searches. We propose a POLyak NONmonotone Stochastic (PoNoS) method, obtained by combining a nonmonotone line search with a Polyak initial step size. Furthermore, we develop a new resetting technique that in the majority of the iterations reduces the amount of backtracks to zero while still maintaining a large initial step size. To the best of our knowledge, a first runtime comparison shows that the epoch-wise advantage of line-search-based methods gets reflected in the overall computational time.

1a - WB120 - 516A - **Non-Convex, Non-Smooth, Stochastic, and High-Performance Computing Approaches** Chair: DONGDONG GE

Recent advances in the COPT conic optimizer – JOACHIM DAHL, CARDINAL OPERATIONS –

The COPT conic optimizer is a state-of-the-art solver for a linear conic problems. This is a well-studied problem for the special case of linear conic optimization over the symmetric cones (that is, linear optimization, second-order cone optimization and semidefinite optimization), and many high-performance implementations are available in both commercial and non-commercial packages. In this talk we give an overview of the COPT conic optimizer, and we discuss recent extensions of the COPT conic optimizer for non-symmetric cones, in particular the exponential cone.

Whereas the last decade has seen abundant research into first-order methods, where non-symmetric cones pose little additional difficulty over symmetric cones (provided first-order derivatives are readily available), then the extension of second-order methods to non-symmetric cones has received considerable less attention, both in academia and industry. A standard, and very successful second-order method for symmetric cones is based on Nesterov-Todd primal-dual scalings, which map primal and dual iterates to the same point, and similar for the derivatives of the barrier functions. As the Nesterov-Todd scalings are defined exclusively for symmetric cones, in the COPT conic optimizer we consider different primal-dual scalings proposed in 2001 by Tuncel. These primal-dual scalings also map primal and dual iterates to the same point, and similarly for the derivatives of the barrier functions. This results in an algorithm with many similarities to the symmetric counter-part based on Nesterov-Todd scalings.

We present numerical results of the algorithm for a range of different problems, including both symmetric cones and exponential cones. We also compare the quality and efficiency of the exponential cone solver with standard algorithms for non-linear optimization, for a range of applications, including logistic regression and geometric programming problems.

Stochastic Structured Nonconvex Optimization in the Absence of Smoothness – QI DENG, SHANGHAI JIAO TONG UNIVERSITY –

In this talk, we will present new algorithm developments for

structured nonconvex stochastic optimization that do not satisfy the standard smoothness conditions. Specifically, we will focus on scenarios where the objective function exhibits weak convexity but lacks smoothness. Through a model-based approach, we will analyze how stochastic optimization can still benefit from minibatching and distributed computing, which are two key techniques traditionally associated with smooth optimization. In the minibatch setting, we will demonstrate that many stochastic algorithms achieve linear speedup over the batch size, even for nonsmooth weakly convex problems. Our analysis is based on a novel sensitivity analysis of the proximal mapping used in each algorithm iteration. In a centralized distributed environment, we will develop an asynchronous stochastic prox-linear method. We will show that the delays only affect the high-order term in the complexity rate and, therefore, are negligible in the long run.

Furthermore, we will extend our study to more general nonsmooth problems without the standard Lipschitz continuity assumption. Based on new adaptive regularization (step-size) strategies, we will demonstrate that it is still possible to preserve the standard convergence rate with a constant failure rate. Our analyses are based on rather weak assumptions: the Lipschitz parameter can be either bounded by a general growth function of $\text{norm}(x)$ or locally estimated through independent random samples.

New complexity results on nonconvex non-smooth optimization – TIANYI LIN, MASSACHUSETTS INSTITUTE OF TECHNOLOGY –

In this talk, we present several new results on minimizing a non-smooth and nonconvex function under a Lipschitz condition. Recent works show that while the classical notion of Clarke stationarity is computationally intractable up to some sufficiently small constant tolerance, the randomized first-order algorithms find an approximate Goldstein stationary point with the finite-time convergence guarantee, which is independent of problem dimension. First of all, we present several motivating examples and explain why the notion of Goldstein stationary is important to the finite-time convergence guarantee. Second, we show that the randomization is necessary to obtain a dimension-independent guarantee, by proving a lower bound of $\Omega(d)$ for any deterministic algorithm that has access to both first-order and zeroth-order oracles. We also show that the zeroth-order oracle is essential to obtain a finite-time convergence guarantee, by showing that any deterministic algorithm with only the first-order oracle can not find an approximate Goldstein stationary point within a finite number of iterations up to sufficiently small constant parameter and tolerance. Finally, we establish the relationship between the notion of Goldstein stationary and the uniform smoothing, and propose several new zeroth-order methods with solid finite-time convergence guarantee.

GPU-Centric Algorithm for Non-Linear Programming Problems – DONGDONG GE, SHANGHAI JIAO TONG UNIVERSITY –

Over the past year, a large number of first-order and higher-order algorithms based on GPU and CUDA architectures have emerged for solving large-scale linear and nonlinear programming problems. There are many algorithms that have significant advantages over traditional CPU architecture algorithms in certain large-scale problems. In this report, we will provide

a brief summary of this trend, including some new developments in hardware vendors such as Nvidia and library function construction. We will also present several new algorithms for linear programming, quadratic programming, and semi definite programming, and demonstrate their significant computational advantages demonstrated through COPT(Cardinal Solver) solver on suitable GPU hardware.

1a - WC903 - 516D - **Nonlinear Programming**
2 Chair: HUYEN TRANG TRAN ↩

Numerical design of optimized first-order methods – YASSINE KAMRI, UCLOUVAIN – François Glineur, Julien M. Hendrickx

Large-scale optimization is central to many engineering applications such as machine learning, signal processing, and control. First-order optimization methods are a popular choice for solving such problems due to their performance and low computational cost. However, in many applications, such as machine learning, their performance is highly dependent on the choice of step sizes. Tuning the step sizes of first-order methods is a notoriously hard nonconvex problem. In this talk, we present several approaches to optimizing the step sizes of first-order optimization methods, relying on recent developments in Performance Estimation Methods. We compare our results to existing optimized step sizes and provide new results, namely numerically optimized steps for cyclic coordinate descent and inexact gradient descent, leading to accelerated convergence rates with and without momentum.

Analysis of Optimization Methods via non-convex Performance Estimation – NIZAR BOUSSELMY, UCLOUVAIN –

The “Performance Estimation Problem framework” computes numerically the exact worst-case performance of a given optimization method on a given function class. It allowed to determine the convergence rate of numerous first-order methods and also to design optimal methods. The convex formulation of the Performance Estimation Problems can be efficiently solved but it restricts the analysis to certain types of methods. Recently, some works have proposed to tackle a non-convex formulation of Performance Estimation Problems allowing more flexibility in the analysis but at a computational cost. In this work, we propose to exploit the non-convex Performance Estimation Problem to analyze new optimization schemes that were out of the scope of the classical convex Performance Estimation Problem framework like zeroth, second-order, and adaptive methods.

Stochastic ISTA/FISTA Adaptive Step Search Algorithms for Convex Composite Optimization – HUYEN TRANG TRAN, CORNELL UNIVERSITY – Lam Nguyen, Katya Scheinberg

We develop and analyze stochastic variants of ISTA and a full backtracking FISTA algorithms [Beck and Teboulle, 2009, Scheinberg et al., 2014] for composite optimization without the assumption that stochastic gradient is an unbiased estimator. This work extends analysis of inexact fixed step ISTA/FISTA in [Schmidt et al., 2011] to the case of stochastic gradient estimates and adaptive step-size parameter chosen by backtracking. It also extends the framework for analyzing stochastic line-search method in [Cartis and Scheinberg, 2018] to the proximal gradient framework as well as to the accelerated first order

methods.

1a - ThA19 - 516D - **Theoretical advances in nonconvex optimization** Chair: HIDEFUMI KAWASAKI ↩

Practical session : Model and solve a non-linear optimization problem with Artelys Knitro – CARLYLE DELIGNY, ARTELYS CANADA INC – Notes: This session presents practical exercises to be performed by the participants. A computer and an Internet connection are required to participate.

The challenges faced by operational research (OR) practitioners, whether in operational scenarios or during technoeconomic analyses, are typically intricate and inherently non-linear. A skilled modeler’s task is to simplify this complexity and identify the core issue whose resolution untangles the rest. A common strategy is to linearize the problem to leverage the wealth of theoretical and practical tools provided by linear programming, such as duality, sensitivity analysis, and efficient algorithms. While this approach is often effective, it’s important to be acquainted with non-linear optimization methods to avoid overly relying on linearization out of habit. In many instances, problems can be efficiently solved directly without linearization, utilizing appropriate non-linear methods and tools. This tutorial will showcase several instances of non-linear problems, typically addressed through linear relaxation, yet warranting a direct approach. It will delve into the methodologies and tools employed to tackle these challenges. This tutorial aims at an audience familiar with combinatorial and linear programming, focusing on practical examples and discussions of both linear and non-linear approaches. The focus will be on practical application, with participants implementing various examples in a Python notebook using the Artelys Knitro solver. This session aims to empower participants to recognize the type of non-linear problem they encounter, determine appropriate methods and tools for solving it, and learn how to effectively utilize them.

The k-cut problem in hypergraphs, a computational study – MOURAD BAILOU, CNRS – Francisco Barahona, Hélène Toussaint

Consider a hypergraph $H = (V, E)$, with a weights vector $w \in \mathbb{R}_+^E$, and a fixed number k . The k -cut problem consists of finding a partition $\{S_1, \dots, S_k\}$ of V that minimizes $w(\delta(S_1, \dots, S_k))$, where $\delta(S_1, \dots, S_k)$ is the set of hyperedges having at least two elements into two different sets of the partition. Based on the lower bound of Narayanan (1996) we develop a combinatorial branch and bound algorithm for computing the optimal solution for the k -cut problem. We also show that a direct extension of Naor’s (2001) linear integer program for the k -cut problem in graphs will not holds in the case of hypergraphs. For this, we produce a new linear integer program that we use to develop a branch and cut algorithm using Cplex. We also consider several heuristics for computing an upper bound. Finally, we give experimental results comparing all these approaches.

Fairness in Submodular Maximization over a Matroid Constraint – MARWA EL HALABI, SAMSUNG SAIT AI LAB – Jakub Tarnawski, Ashkan Norouzi Fard, Thuy Duong Vuong

Submodular maximization over a matroid constraint is a fundamental problem with various applications in machine learning. Some of these applications involve decision-making over data-points with sensitive attributes such as gender or race. In such settings, it is crucial to guarantee that the selected solution is fairly distributed with respect to this attribute. Recently, fairness has been investigated in submodular maximization under a cardinality constraint in both the streaming and offline settings, however the more general problem with matroid constraint has only been considered in the streaming setting and only for monotone objectives. This work fills this gap. We propose various algorithms and impossibility results offering different trade-offs between quality, fairness, and generality.

An antipodal theorem for nonlinear optimization problems – HIDEFUMI KAWASAKI, KYUSHU UNIVERSITY –

In this talk, we investigate nonlinear optimization problems using algebraic topology. Specifically, we utilize Borsuk-Ulam’s theorem, which states that for any continuous mapping from the n -sphere to n -dimensional Euclidean space, there exists a pair of antipodal points that are mapped to the same point. We apply this theorem to the optimal-value functions of an n -tuple of parametric optimization problems and present an antipodal theorem for them. Furthermore, by considering the support function of an n -tuple of convex sets as the objective functions, we introduce a ham-sandwich-type theorem.

1a - ThA42 - 510C - **Advances in Large-Scale Nonlinear Optimization IV** Chair: MIKE O’NEILL

A stochastic moving ball approximation method for smooth convex constrained minimization – ION NECOARA, NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY POLITEHNICA BUCHAREST –

In this talk, we consider constrained optimization problems with convex, smooth objective and constraints. We propose a new stochastic gradient algorithm, called the Stochastic Moving Ball Approximation (SMBA) method, to solve this class of problems, where at each iteration we first take a gradient step for the objective function and then perform a projection step onto one ball approximation of a randomly chosen constraint. The computational simplicity of SMBA, which uses first-order information and considers only one constraint at a time, makes it suitable for large-scale problems with many functional constraints. We provide a convergence analysis for the SMBA algorithm using basic assumptions on the problem, that yields new convergence rates in both optimality and feasibility criteria evaluated at some average point. Our convergence proofs are novel since we need to deal properly with infeasible iterates and with quadratic upper approximations of constraints that may yield empty balls. We derive convergence rates when the objective function is convex or strongly convex. Preliminary numerical experiments on convex quadratically constrained quadratic problems demonstrate the viability and performance of our method when compared to some existing state-of-the-art optimization methods and software.

Randomized Algorithms for Variational Inequality Problems – ANGELIA NEDICH, ARIZONA STATE UNIVERSITY – Abhishek Chakraborty

We consider a variational inequality problem arising from a game among multiple agents, where each agent aims to minimize its own cost function subject to its constrained set represented as the intersection of a (possibly infinite) number of convex functional level sets. A direct projection-based approach or Lagrangian-based techniques for such a problem can be computationally expensive if not impossible to implement. To deal with the problem, we consider randomized methods that avoid the projection step on the whole constraint set by employing random feasibility updates. In particular, we propose and analyze such random methods for solving VIs based on the projection method, Korpelevich method, and Popov method. We establish the almost sure convergence of the methods and, also, provide their convergence rate guarantees. We illustrate the performance of the methods in simulations for two-agent games.

Retrospective Approximation for Stochastic Constrained Problems Using Sequential Quadratic Programming – SHAGUN GUPTA, UNIVERSITY OF TEXAS AT AUSTIN – Albert Berahas, Raghu Bollapragada

Sequential Quadratic Programming (SQP) is one of the state-of-the-art algorithms used to solve deterministic constrained nonlinear optimization problems. In recent years, the framework has been extended to solve deterministic equality and inequality constrained problems with stochastic objective functions. In response to the challenges posed by stochasticity, various schemes have been incorporated into SQP algorithms to adapt key parameters, such as step size and merit parameter, from the deterministic setting to the stochastic setting. These include stochastic line search, Lipschitz constant estimation, and Hessian averaging. In our work, we leverage SQP algorithms within the innovative framework of Retrospective Approximation. This framework introduces a novel approach to solving stochastic constrained problems by allowing the SQP algorithm to solve a series of subsampled deterministic subproblems. Each deterministic subproblem is solved not to optimality, but to a specified accuracy with an increasing sample size for the subsampled deterministic problems. This strategic decoupling of stochasticity from the SQP algorithm proves instrumental, enabling the utilization of legacy deterministic solvers. Thus, by decoupling stochasticity, the Retrospective Approximation framework facilitates the integration of legacy deterministic solvers, reducing requirements for hyperparameter tuning in stochastic settings. We provide theoretical convergence requirements for the increase in the subsampling batch size and required solution accuracy for deterministic subproblems. We also conduct numerical experiments to showcase the utilization of legacy deterministic solvers for stochastic constrained problems.

A Double Stepsize Stochastic SQP Method with Complexity Guarantees – MIKE O’NEILL, UNIVERSITY OF NORTH CAROLINA –

Stochastic gradient algorithms play a key role in the training of large-scale machine learning models. These algorithms can be readily extended to problems with simple constraints, such as when projections onto the feasible region can be computed efficiently. Recently, there has been a surge of interest in stochastic gradient-like algorithms for stochastic optimization problems with nonlinear and nonconvex constraints due to modern machine learning applications such as fair machine learning and

physics informed neural networks. In this talk, we present a new stochastic Sequential Quadratic Programming (SQP) for deterministically constrained stochastic optimization. Unlike previous work, this algorithm utilizes a different step size for each component of the orthogonal decomposition of the SQP step, which enables faster convergence with respect to the constraint violation and an improved worst-case complexity result. In addition to improving the worst-case complexity, this algorithmic approach also enables significant flexibility with respect to computing step sizes, such as employing a (safeguarded) line search on the constraint violation. Some preliminary numerical experiments will also be presented.

1a - ThA74 - 516A - The interplay between optimization, statistics and machine learning I
Chair: LIJUN DING

Sharpness and well-conditioning of nonsmooth convex formulations in statistical signal recovery – LIJUN DING, UNIVERSITY OF WISCONSIN MADISON –

We study a sample complexity vs. conditioning tradeoff in modern signal recovery problems where convex optimization problems are built from sampled observations. We begin by introducing a set of condition numbers related to sharpness in ℓ_p norm or Schatten- p norms based on nonsmooth reformulations of a class of convex optimization problems, including sparse recovery, low-rank matrix sensing, covariance estimation, and (abstract) phase retrieval. In each of the recovery tasks, we show that the condition numbers become dimension independent constants once the sample size exceeds some constant multiple of the recovery threshold. Structurally, this result ensures that the inaccuracy in the recovered signal due to both observation noise and optimization error is well-controlled. Algorithmically, such a result ensures that a new first-order method for solving the class of sharp convex functions in a given ℓ_p norm or Schatten- p norm, when applied to the nonsmooth formulations, achieves nearly-dimension-independent linear convergence. Time permitting, we will give a detailed explanation of (i) the statistical model, (ii) how the model implies the restricted isometry property (RIP), (iii) how RIP implies good conditioning of the nonsmooth reformulations, and (iv) the convergence proof of the new first-order algorithm.

Fast Nonconvex Matrix Estimation with Global Optimality Certification – JIALUN ZHANG, UNIVERSITY OF ILLINOIS URBANA CHAMPAIGN –

We consider using gradient descent to minimize the nonconvex function $f(X) = \phi(XX^T)$ over an $n \times r$ factor matrix X , in which ϕ is an underlying smooth convex cost function defined over $n \times n$ matrices. While only a second-order stationary point X can be provably found in reasonable time, if X is additionally *rank deficient*, then its rank deficiency certifies it as being globally optimal. This way of certifying global optimality necessarily requires the search rank r of the current iterate X to be *overparameterized* with respect to the rank r^* of the global minimizer X^* . Unfortunately, overparameterization significantly slows down the convergence of gradient descent, from a linear rate with $r = r^*$ to a sublinear rate when $r > r^*$, even when ϕ is strongly convex. In this paper, we propose an in-

expensive preconditioner that restores the convergence rate of gradient descent back to linear in the overparameterized case, while also making it agnostic to possible ill-conditioning in the global minimizer X^* .

cuPDLP: A GPU implementation of restarted primal-dual hybrid gradient for linear programming – JINWEN YANG, UNIVERSITY OF CHICAGO – Haihao Lu

We provide an affirmative answer to the long-standing question: Are GPUs useful in solving linear programming? We present cuPDLP, a GPU implementation of restarted primal-dual hybrid gradient (PDHG) for solving linear programming (LP). We show that this prototype implementation has comparable numerical performance on standard LP benchmark sets as Gurobi, a highly optimized implementation of the simplex and interior-point methods. Furthermore, we present the superior performance of cuPDLP with its CPU counterpart. This demonstrates the power of using GPUs in the optimization solvers.

Top-K Ranking with a Monotone Adversary – CONG MA, UNIVERSITY OF CHICAGO –

In this talk, we address the top- K ranking problem with a monotone adversary. We consider the scenario where a comparison graph is randomly generated and the adversary is allowed to add arbitrary edges. The statistician's goal is then to accurately identify the top- K preferred items based on pairwise comparisons derived from this semi-random comparison graph. The main contribution of this paper is to develop a weighted maximum likelihood estimator (MLE) that achieves near-optimal sample complexity, up to a $\log^2(n)$ factor, where n denotes the number of items under comparison. This is made possible through a combination of analytical and algorithmic innovations. On the analytical front, we provide a refined ℓ_∞ error analysis of the weighted MLE that is more explicit and tighter than existing analyses. It relates the ℓ_∞ error with the spectral properties of the weighted comparison graph. Motivated by this, our algorithmic innovation involves the development of a semi-definite-program-based (SDP-based) approach to reweight the semi-random graph and meet specified spectral properties. Additionally, we propose a first-order method based on the Matrix Multiplicative Weight Update (MMWU) framework. This method efficiently solves the resulting SDP in nearly-linear time relative to the size of the semi-random comparison graph.

1a - ThA135 - 511C - From First- to High-order Methodologies: Recent Advances and Trends III Chair: ANDRE MILZAREK

HSODF: A Homogeneous Framework for Second-Order Methods – CHUWEN ZHANG, SHANGHAI UNIVERSITY OF FINANCE AND ECONOMICS – Chang He, Yuntian Jiang, Dongdong Ge, Bo Jiang, Yinyu Ye

We introduce a homogeneous second-order descent framework (HSODF) for nonconvex and convex optimization based on the generalized homogeneous model (GHM). In comparison to the Newton steps, the GHM can be solved by extremal symmetric eigenvalue procedures and thus grant an advantage in ill-conditioned problems. We show that HSODF is able to recover some well-known second-order methods, such as trust-

region methods and gradient regularized methods while maintaining comparable iteration complexity bounds. We propose specific variants in the framework. Finally, we show HSODM can be used in reinforcement learning, linear systems and interior point methods.

Descent properties of an Anderson accelerated gradient method with restarting

– ANDRE MILZAREK, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Wenqing Ouyang, Yang Liu

Anderson Acceleration (AA) is a popular acceleration technique to enhance the convergence of fixed-point iterations. The analysis of AA approaches typically focuses on the convergence behavior of a corresponding fixed-point residual, while the behavior of the underlying objective function values along the accelerated iterates is currently not well understood. In this talk, we present new results showing that AA with restarting is a local descent method which can decrease the objective function faster than the gradient method. These insights provide a novel perspective on the convergence analysis of AA in non-convex optimization. Numerical experiments are conducted to illustrate our theoretical findings.

Online Learning Guided Quasi-Newton Methods: Improved Global Non-asymptotic Guarantees

– RUICHEN JIANG, UT AUSTIN – QIUJIANG JIN, ARYAN MOKHTARI

Quasi-Newton (QN) methods are popular iterative algorithms known for their superior practical performance compared to Gradient Descent (GD)-type methods. However, the existing theoretical results for this class of algorithms do not sufficiently justify their advantage over GD-type methods. In this talk, we discuss our recent efforts to address this issue. Specifically, in the strongly convex setting, we propose the first “globally” convergent QN method that achieves an explicit “non-asymptotic superlinear” rate. We show that the rate presented for our method is provably faster than GD after at most $O(d)$ iterations, where d is the problem dimension. Additionally, in the convex setting, we present an accelerated variant of our proposed method that provably outperforms the accelerated gradient method and converges at a rate of $O(\min\{1/k^2, \sqrt{d \log k}/k^{2.5}\})$, where k is the number of iterations. To attain these results, we diverge from conventional approaches and construct our QN methods based on the Hybrid Proximal Extragradient (HPE) framework and its accelerated variants. Furthermore, a pivotal algorithmic concept underpinning our methodologies is an online learning framework for updating the Hessian approximation matrices. Specifically, we relate our method’s convergence rate to the regret of a specific online convex optimization problem in the matrix space and choose the sequence of Hessian approximation matrices to minimize its overall regret.

Adam-family Methods for Nonsmooth Optimization with Convergence Guarantees

– NACHUAN XIAO, NATIONAL UNIVERSITY OF SINGAPORE – Xiaoyin Hu, Xin Liu, Kim Chuan Toh

In this talk, we present a comprehensive study on the convergence properties of Adam-family methods for nonsmooth optimization, especially in the training of nonsmooth neural networks. We introduce a novel two-timescale framework that adopts a two-timescale updating scheme, and prove its convergence properties under mild assumptions. We prove that under

mild conditions, any cluster point of the sequence generated by our proposed framework is a stationary point of the objective function in the sense of the conservative field. Furthermore, we establish that under mild conditions with randomly chosen initial points and stepsizes, almost surely, any cluster point of the sequence is a Clarke stationary point of the objective function, independent of the chosen conservative field. Our proposed framework encompasses various popular Adam-family methods, providing convergence guarantees for these methods in training nonsmooth neural networks. Furthermore, we develop stochastic subgradient methods that incorporate gradient clipping techniques for training nonsmooth neural networks with heavy-tailed noise. Through our framework, we prove that under heavy-tailed evaluation noises that are only assumed to be integrable, our proposed gradient clipping methods conform to the proposed framework. As a result, the convergence properties of these gradient clipping methods directly follow those established for our proposed framework under mild conditions. Extensive numerical experiments demonstrate the high efficiency and robustness of our proposed methods.

1a - ThB35 - 516D - Nonlinear Programming for Data Science I Chair: VIVAK PATEL

Reliable Optimization Algorithms when Objective Function Evaluations are Expensive – VIVAK PATEL, UNIVERSITY OF WISCONSIN MADISON –

Gradient methods are experiencing a growth in methodological and theoretical developments owing to the challenges of optimization problems arising in data science. Focusing on data science applications with expensive objective function evaluations yet inexpensive gradient function evaluations (e.g., identification of conservative estimating equation models), gradient methods that never make objective function evaluations are either being rejuvenated or actively developed. However, as we show, such gradient methods are all susceptible to catastrophic divergence under realistic conditions for data science applications. In light of this, gradient methods which make use of objective function evaluations become more appealing, yet, as we show, can result in an exponential increase in objective evaluations between accepted iterates. As a result, existing gradient methods are poorly suited to the needs of optimization problems arising from data science. In this talk, we present a novel, generic methodology that economically uses objective function evaluations in a problem-driven manner to prevent catastrophic divergence and avoid an explosion in objective evaluations between accepted iterates. Our methodology allows for specific procedures that can make use of specific step size selection methodologies or search direction strategies, and we develop a novel step size selection methodology that is well-suited to data science applications. We show that a procedure resulting from our methodology is highly competitive with standard optimization methods on CUTEst test problems. We then show a procedure resulting from our methodology is highly favorable relative to standard optimization methods on optimization problems arising in our target data science applications. Thus, we provide a novel gradient methodology that is better suited to optimization problems arising in data science.

Yet another fast variant of Newton’s method for nonconvex optimization – PHILIPPE TOINT,

UNIVERSITY OF NAMUR (PAST) – Serge Gratton, Jerad Sadok

A second-order algorithm is proposed for minimizing any smooth nonconvex function that suitably alternates between regularized Newton step and negative curvature steps. At most iterations, the Hessian matrix is regularized with the square root of the norm of the current gradient and an additional term taking moderate negative curvature into account, while a negative curvature step is taken only exceptionally otherwise. As a consequence, the proposed method has the feature that it only requires the solution of a single linear system at nearly all iterations., making is ideally suited to applications where of-the-shelf direct or preconditioned iterative linear solvers are available. We establish that , under standard assumptions, at most $O(\log(\epsilon)|\epsilon|^{-3/2})$ evaluations of the problem’s objective function , gradient and Hessian are needed for this algorithm to obtain an epsilon-approximate first-order critical point, and at most $O(|\log(\epsilon)|\epsilon^{-3})$ to obtain an epsilon-approximate second-order one. Initial numerical experiments with two variants of the new method are finally presented. The first variant uses direct linear solvers and the second uses subspace-based ones, like truncated conjugate-gradients. The performance of these variants on a set of standard test problems is shown to be competitive.

Approximate Derivatives for Tensor Methods – KARL WELZEL, UNIVERSITY OF OXFORD – Raphael Hauser, Yang Liu, Coralia Cartis

It is well known that second-order methods converge much faster than first-order methods. Based on the theory one should expect another speed-up (in terms of function evaluations) when incorporating third-order tensor information. Third-order oracles, however, are almost never available. We therefore propose a generalization of the quasi-Newton approach to higher-order methods. Just like quasi-Newton methods extract second-order information from repeated gradient evaluations, this method extracts third-order information from repeated Hessian evaluations. In each step the new approximation is one that satisfies a higher-order analogue of the secant equation and simultaneously is closest possible to the previous approximation in a weighted Frobenius norm. It can be shown that such approximations recover the true derivative in the limit under certain assumptions and that, in this case, tensor methods using these approximations achieve super-quadratic local convergence. We present small-scale numerical experiments that show the advantages and disadvantages of using tensor methods over second-order models in practice. For this we focus on the ARp method which uses a regularized Taylor expansion as a local model and is able to incorporate both second- and third-order methods in the same framework.

Deterministic and Stochastic Frank Wolfe Recursion on Probability Spaces – RAGHU PASUPATHY, PURDUE UNIVERSITY – Di Yu, Shane Henderson

Motivated by emergency response applications, we consider smooth optimization problems over the space of non-negative Borel measures supported on a convex compact set $X \subset \mathbb{R}^d$. Through direct application of variational calculus, we devise a deterministic Frank-Wolfe (DFW) first-order recursion using the influence function, to generate iterates exhibiting $O(\epsilon - 1)$ convergence in function value. Importantly, the recursion we present is made possible by a simple result that expresses the

solution to the infinite-dimensional Frank- Wolfe sub-problem as the solution to a finite-dimensional optimization problem in $\mathbb{R}^d$. Such a result allows writing the Frank-Wolfe recursion as a convex combination of the incumbent iterate and a Dirac measure concentrating on the minimum of the influence function at the incumbent iterate. To account for common application contexts where only Monte Carlo observations of the objective and influence function are available, we construct the stochastic Frank-Wolfe (SFW) variation that generates a sequence of random measures constructed using minima of increasingly accurate estimates of the influence function. We demonstrate that the SFW sequence exhibits $O(1/k)$ complexity almost surely and in expectation. Furthermore, we show that an easy-to-implement fixed-step, fixed-sample version of SFW exhibits exponential convergence to ϵ -optimality. We end with a central limit theorem on the observed objective values at the sequence of generated random measures. We include a number of illustrative example problem classes along with exact calculations of the influence function.

1a - ThB37 - 510C - **Sum-of-squares methods for continuous optimization and beyond**
Chair: FRANCIS BACH

GloptiNets: Scalable Non-Convex Optimization with Certificates – GASPARD BEUGNOT, INRIA –

We present a novel approach to non-convex optimization with certificates, which handles smooth functions on the hypercube or on the torus. Unlike traditional methods that rely on algebraic properties, our algorithm exploits the regularity of the target function intrinsic in the decay of its Fourier spectrum. By defining a tractable family of models, we allow at the same time to obtain precise certificates and to leverage the advanced and powerful computational techniques developed to optimize neural networks. In this way the scalability of our approach is naturally enhanced by parallel computing with GPUs. Our approach, when applied to the case of polynomials of moderate dimensions but with thousands of coefficients, outperforms the state-of-the-art optimization methods with certificates, as the ones based on Lasserre’s hierarchy, addressing problems intractable for the competitors.

Convex relaxations for Gibbs states – HAMZA FAWZI, UNIVERSITY OF CAMBRIDGE –

A fundamental question in quantum information is to compute expectation values under the Gibbs state of a given local Hamiltonian. I will present in this talk an algorithm for this problem based on hierarchies of convex relaxations over the positive semidefinite cone and the matrix relative entropy cone. I will also discuss convergence guarantees. Based on joint work with Omar Fawzi and Samuel Scalet (arXiv:2311.18706).

Moment-SOS relaxations for variational problems – GIOVANNI FANTUZZI, FAU ERLANGEN NÜRNBERG –

Moment-SOS relaxations are an established technique to compute converging sequences of lower bounds on the global minimum of finite-dimensional polynomial optimization problems. In this talk, I will present two recent extensions of moment-SOS relaxations to infinite-dimensional variational problems, where an integral functional is to be minimized over functions from a Sobolev space. The first extension optimizes so-called "null

Lagrangian translations” and returns certified lower bounds on the global minimum of the variational problem. The second extension, instead, produces upper bounds by approximating minimizers of finite element discretizations of the variational problem. Conditions that ensure these upper and lower bounds converge to the desired global minimum will be discussed, and current gaps between theory and practice will be illustrated by means of examples.

Nonconvergence of some sum-of-squares bounds for global polynomial optimization – LUCAS SLOT, ETH ZURICH – Manuel Wiedmer

Consider the problem of computing the minimum of a polynomial f on a closed set X . Given a measure μ supported on X , Lasserre proposes a decreasing sequence of upper bounds on this minimum, each of which may be computed by solving a semidefinite program. When X is compact, these bounds converge to the minimum under minor assumptions on μ . Later, Lasserre introduces a related, but far more economical sequence of upper bounds which rely on the push-forward measure of μ by f . While these new bounds are weaker a priori, they actually achieve similar asymptotic convergence rates on compact sets. We show that no such free lunch exists in the non-compact setting. While convergence for the standard bounds is guaranteed when $X = \mathbb{R}^n$ and μ is the Gaussian distribution, we prove that the bounds relying on the push-forward measure fail to converge in that setting already for polynomials of degree 6.

1a - ThB93 - 516A - **Nonconvex optimization: landscapes and dynamics I** Chair: CHRISTOPHER CRISCITIELLO

The sensor network localization problem has benign landscape under mild rank relaxation – CHRISTOPHER CRISCITIELLO, EPFL – Andrew Mcrae, Quentin Rebjock, Nicolas Boumal

We consider the sensor network localization problem (also called metric multidimensional scaling): one observes some pairwise distances between n ground-truth points in $\mathbb{R}^d$, and the goal is to recover this cloud of ground-truth points (up to translation and rotation). The corresponding optimization problem is nonconvex, and we show that it can have spurious local minima. However, inspired by numerical experiments, we argue that if one relaxes the problem by optimizing over clouds of n points in dimension k greater than d , then all second-order critical points of the problem are global minima. Specifically, we show this for two settings: (1) for arbitrary ground-truth points, when all pairwise distances are known, and $k = O(\sqrt{nd})$, and: (2) for isotropic random ground-truth points, when most (but not necessarily all) pairwise distances are known, and $k = O(d \log(n))$. To the best of our knowledge, these are the first landscape results for this nonconvex version of sensor network localization.

Stochastic approximation with decision-dependent distributions: asymptotic normality and optimality – MATEO DIAZ, JOHNS HOPKINS UNIVERSITY – Joshua Cutler, Dmitriy Drusvyatskiy

We analyze a stochastic approximation algorithm for decision-dependent problems, wherein the data distribution used by the algorithm evolves along the iterate sequence. The primary

examples of such problems appear in performative prediction and its multiplayer extensions. We show that under mild assumptions, the deviation between the average iterate of the algorithm and the solution is asymptotically normal, with a covariance that nicely decouples the effects of the gradient noise and the distributional shift. Moreover, building on the work of Hajek and Le Cam, we show that the asymptotic performance of the algorithm is locally minimax optimal.

If you’ve reached this point, you should know that I am not into long abstracts, but the system is forcing me to add one. So here is an AI-generated, nonsensical, grandiose paragraph (My coauthors had nothing to do with this): In the splendid twilight of verdurous philosophies, where the symphony of eloquence pirouettes upon the delicate breeze of imagination, there exists a realm ethereally tethered to the quintessence of grandiloquence. Herein, the vernacular tapestry weaves its intricate dance, embroidering the ether with the silk of opulent lexicons, each thread a testament to the magniloquent. Amidst the labyrinthine gardens of verbal extravagance, where words bloom like the rarest of orchids under the luminous gaze of the moon’s embrace, one finds the essence of conversation elevated to an art of celestial craftsmanship.

HALLaR: A new Large-Scale SDP Solver and application for Maximum Stable Set – CIFUENTES DIEGO, GEORGIA INSTITUTE OF TECHNOLOGY –

HALLaR is a new first-order method for solving large-scale semidefinite programs (SDPs) with bounded domain. It relies on a novel hybrid method that uses both nonconvex low-rank optimization and the convex Frank-Wolf method. Computational results show that HALLaR finds highly accurate solutions in substantially less time than state-of-the-art SDP solvers. In particular, we solve the stable set SDP for graphs with up to 10^7 vertices using a regular laptop. We also introduce a novel rounding procedure in order to construct good solutions to the original (discrete) stable set problem from the SDP solution.

Symmetry and Critical Points – YOSSI ARJEVANI, THE HEBREW UNIVERSITY –

Given a nonconvex function invariant under the action of some group G , it is natural to ask how the critical points reflect this symmetry. It turns out that in various cases of interest, notably fitting two-layer ReLU networks and symmetric tensor decomposition, critical points are fixed by (typically large) subgroups of G , and so are symmetry breaking.

In this talk, we will see how symmetry breaking phenomena of this nature allow new ways of recognizing, differentiating and understanding local minima. Specific topics include: – Sharp estimates on the loss, the Hessian spectrum and higher-order derivatives holding in finite, arbitrarily large, dimensionality (through Puiseux series representation) – Analytic results pertaining to foundational questions in optimization, curvature and generalization – Annihilation of spurious minima – Symmetry breaking arcs connecting critical points (time permitting)

Our analytic approach makes extensive use of methods from o-minimal theory and representation theory of groups. No special background will be assumed but a basic familiarity with group theory .

1a - ThB137 - 511C - **First-order Methods and Large-scale Optimization II** Chair: JUNWEN QIU

Adaptivity in convex optimization beyond minimization – PUYA LATAFAT, KU LEUVEN – Andreas Themelis, Patrinos Panagiotis

Bilevel optimization is a comprehensive framework that bridges single- and multi-objective optimization. It encompasses many general formulations, such as, but not limited to, standard nonlinear programs. This work demonstrates how elementary proximal gradient iterations can be used to solve a wide class of convex bilevel optimization problems without involving sub-routines. Compared to and improving upon existing methods, ours (1) can handle a much wider class of problems, including both constraints and nonsmooth terms, (2) requires mere local Lipschitz gradient continuity of the differentiable terms without imposing any strong convexity assumptions, and (3) provides a systematic adaptive stepsize selection strategy that leads to large and nonmonotonic stepsize sequences while being insensitive to the choice of parameters.

The Anytime Convergence of Stochastic Gradient Descent with Momentum: From a Continuous-Time Perspective – YASONG FENG, FUDAN UNIVERSITY – Jiang Yifan, Tianyu Wang, Zhiliang Ying

In this talk, we focus on the stochastic optimization problem and introduce a continuous-time perspective. We propose a stochastic first-order algorithm, called Stochastic Gradient Descent with Momentum (SGDM), and show that the trajectory of SGDM, despite its stochastic nature, converges to a deterministic second-order Ordinary Differential Equation (ODE) in L2-norm, as the stepsize goes to zero. The connection between the ODE and the algorithm results in delightful patterns in the discrete-time convergence analysis. More specifically, we develop convergence results for the ODE through a Lyapunov function, and translate the whole argument to the discrete-time case. This approach yields a novel anytime convergence guarantee for stochastic gradient methods. Our contribution is significant in that it better captures the convergence behavior across the entire trajectory of the algorithm, rather than at a single iterate.

Efficient sparse probability measures recovery via Bregman gradient – JIANTING PAN, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Ming Yan

This paper presents an algorithm tailored for the efficient recovery of sparse probability measures incorporating l0-sparse regularization within the probability simplex constraint. Employing the Bregman proximal gradient method, our algorithm achieves sparsity by explicitly solving underlying subproblems. We rigorously establish the convergence properties of the algorithm, showcasing its capacity to converge to a local minimum with a convergence rate of $O(1/k)$ under mild assumptions. To substantiate the efficacy of our algorithm, we conduct numerical experiments, offering a compelling demonstration of its efficiency in recovering sparse probability measure.

Universal heavy-ball method for nonconvex optimization under Hölder continuous Hess-

sians – NAOKI MARUMO, UNIVERSITY OF TOKYO – Akiko Takeda

We propose a new first-order method for minimizing nonconvex functions with Lipschitz continuous gradients and Hölder continuous Hessians. The proposed algorithm is a heavy-ball method equipped with two particular restart mechanisms. The algorithm enjoys an oracle complexity bound for finding a solution where the gradient norm is less than ϵ ; the exponent for ϵ in the complexity bound depends on the Hölder exponent ν of Hessians. Our complexity result covers the classical bound for $\nu = 0$ (i.e., Lipschitz continuous gradients) and a state-of-the-art bound for $\nu = 1$ (i.e., Lipschitz continuous gradients and Hessians). Our algorithm is ν -independent and thus universal; it automatically achieves the complexity bound with the optimal ν without knowledge of the Hölder continuity. In addition, the algorithm does not require other problem-dependent parameters as input, including the gradient's Lipschitz constant or the target accuracy ϵ . Numerical results illustrate that the proposed method is promising.

1a - ThC904 - 516D - **Nonlinear Programming**
3 Chair: IYAD WALWIL

A parallelizable ADMM approach for block-structured quadratic programs – MICHEL LAHMANN, TU BRAUNSCHWEIG – Christian Kirches

Block-structured quadratic programs (QPs) arise frequently in the context of optimal control problems. The goal of our study is to provide a fast solution to these QPs which is crucial for the successful application of optimal control algorithms to many real-world problems. Existing ADMM-based QP solvers usually split the problem into an equality-constrained QP and a projection onto a box. In general, this is a proven and widely used method. Nevertheless, splitting the problem in this way does not allow the block structure of QPs arising from optimal control problems to be optimally utilized. With our new splitting of the problem we are able to utilize this block structure and solve parts of the problem in parallel. We will introduce the resulting method and show first numerical results.

Fuzzy clustering with capacity constraints: theory, algorithm, convergence analysis and numerical experiments – ADEMIR ALVES RIBEIRO, FEDERAL UNIVERSITY OF PARANÁ (UFPR) – Lucas Pedroso, Kleber Benatti

In this talk we present a theoretical analysis on fuzzy centroid-based clustering methods. We consider constraints that may be useful in some practical applications, such as restrictions on the number of points in each group, and methods that deal with these constraints. We propose a more general formulation to the constrained clustering problem, where each point has an associated weight, and the sum of the weights of the points that compose each group is established a priori. We discuss existence and uniqueness of solutions of the involved problems, providing mathematical foundations for the established formulas. Besides, we propose a practical algorithm for solving this problem and present its convergence analysis. This algorithm follows an alternate minimization scheme, wherein a given iteration addresses first the problem in terms of the probabilities of each point belonging to each cluster and then finds the position of the centroids. The computation of the probabilities is performed by solving a linear system derived from the Karush-Kuhn-Tucker conditions. With the aim of validating

our algorithm, we present numerical tests with synthetic data to demonstrate its performance for the given problems.

The Smoothed Duality Gap as a Stopping Criterion – IYAD WALWIL, TÉLÉCOM PARIS – Olivier Fercoq

We optimize the running time of the primal-dual algorithms by optimizing their stopping criteria for solving convex optimization problems under affine equality constraints, which means terminating the algorithm earlier with fewer iterations. We study the relations between four stopping criteria and show under which conditions they are accurate to detect optimal solutions. The uncomputable one: "Optimality gap and Feasibility error", and the computable ones: the "Karush-Kuhn-Tucker error", the "Projected Duality Gap", and the "Smoothed Duality Gap". Assuming metric sub-regularity or quadratic error bound, we establish that all of the computable criteria provide practical upper bounds for the optimality gap, and approximate it effectively. Furthermore, we establish comparability between some of the computable criteria under certain conditions. Numerical experiments on basis pursuit, and quadratic programs with(out) non-negative weights corroborate these findings and show the superior stability of the smoothed duality gap over the rest.

1a - FA83 - 516D - **Gradient-based methods: theory and advances** Chair: SHOTA TAKAHASHI [↩](#)

Nonconvex Stochastic Bregman Proximal Gradient Method with Application to Deep Learning – JINGYANG LI, NATIONAL UNIVERSITY OF SINGAPORE – Kuangyu Ding, Kim Chuan Toh

The widely used stochastic gradient methods for minimizing nonconvex composite objective functions require the Lipschitz smoothness of the differentiable part. But the requirement does not hold true for problem classes including quadratic inverse problems and training neural networks. To address this issue, we investigate a family of stochastic Bregman proximal gradient (SBPG) methods, which only require smooth adaptivity of the differentiable part. SBPG replaces the upper quadratic approximation used in SGD with the Bregman proximity measure, resulting in a better approximation model that captures the non-Lipschitz gradients of the nonconvex objective. We formulate the vanilla SBPG and establish its convergence properties under nonconvex setting without finite-sum structure. Experimental results on quadratic inverse problems testify the robustness of SBPG. Moreover, we propose a momentum-based version of SBPG (MSBPG) and prove it has improved convergence properties. We apply MSBPG to the training of deep neural networks with a polynomial kernel function, which ensures the smooth adaptivity of the loss function. Experimental results on representative benchmarks demonstrate the effectiveness and robustness of MSBPG in training neural networks. Since the additional computation cost of MSBPG compared with SGD is negligible in large-scale optimization, MSBPG can potentially be employed as an universal open-source optimizer in the future.

Approximate Bregman Proximal Gradient Algorithm for Relatively Smooth Nonconvex Optimization – SHOTA TAKAHASHI, THE UNIVERSITY OF TOKYO – Akiko Takeda

In applications to machine learning and signal processing, regularization is useful to avoid overfitting or to impose the model structure. An optimization problem with a regularization term can be formulated as a composite nonconvex optimization problem. To address the composite nonconvex optimization problem, we propose the approximate Bregman proximal gradient algorithm (ABPG). ABPG employs a new distance that approximates the Bregman distance, making the subproblem of ABPG simpler to solve compared to that of existing Bregman-type algorithms. The subproblem of ABPG often has a closed-form solution due to the approximate Bregman distance. Similarly to existing Bregman-type algorithms, ABPG does not require the global Lipschitz continuity for the gradient of the smooth part. Instead, assuming the smooth adaptable property, we establish the global subsequential convergence under standard assumptions. Additionally, assuming that the objective function is a Kurdyka-Łojasiewicz function, we prove the global convergence of ABPG for a special case. In numerical experiments, we have applied ABPG to the convex lp regularized least squares problem, the lp loss problem, and the nonnegative linear system. Our numerical results show that ABPG outperforms existing algorithms especially when the gradient of the smooth part is not globally Lipschitz or even local Lipschitz continuous.

Bregman Proximal Methods for Quantum Information Theoretic Problems – KERRY HE, MONASH UNIVERSITY – James Saunderson, Hamza Fawzi

Convex optimization problems arise naturally when studying the fundamental limits of quantum communication systems, such as computing quantities known as quantum channel capacities. These problems typically involve minimizing or maximizing functions composed of quantum entropies, which are complex nonlinear functions of matrices, over the set of unit-trace Hermitian matrices. Standard methods such as projected gradient descent do not work well for these problems as the gradients of quantum entropies are not well-defined on the boundary of the set of unit-trace Hermitian matrices (i.e., rank-deficient matrices). Instead, in this work, we show how non-Euclidean generalizations of gradient descent algorithms, known as Bregman proximal gradient methods, can be applied to efficiently solve many problems from quantum information theory. We show how mirror descent and primal-dual hybrid gradient are appropriate algorithms for computing many quantum information theoretic quantities. Additionally, we prove for these problems that mirror descent and primal-dual hybrid gradient are guaranteed to converge sublinearly or linearly by using convergence analysis based on relative smoothness and strong convexity properties of the objective functions. We present numerical experiments showing how these algorithms perform for computing quantum channel capacities, quantum rate-distortion functions, and relative entropies of entanglement.

1a - FA94 - 510C - **Nonconvex optimization: landscapes and dynamics II** Chair: ANDREEA MUSAT [↩](#)

Deep Learning with Nontrivial Constraints – JU SUN, UNIVERSITY OF MINNESOTA – Chuan He, Ryan Devera, Wenjie Zhang, Zhaosong Lu, Ying Cui
Despite the sweeping success and impact of deep learning in nu-

merous domains, imposing explicit constraints is relatively new but increasingly pressing in deep learning (DL), stimulated by, e.g., trustworthy AI that performs robust optimization over complicated perturbation sets and scientific and engineering applications that need to respect physical laws and constraints. In this talk, we will (1) survey DL problems with nontrivial constraints across science, engineering, and medicine, (2) highlight the NCVX computing framework we have recently built, which provides deterministic solvers to solve constrained DL problems, and (3) invite the optimization community to solve the stochastic constrained DL problems.

Adaptive Quasi-Newton and Anderson Acceleration with Global (Accelerated) Convergence Rates – DAMIEN SCIEUR, SAMSUNG SAIL MONTREAL –

Despite the impressive numerical performance of the quasi-Newton and Anderson/nonlinear acceleration methods, their global convergence rates have remained elusive for over 50 years. This study addresses this long-standing issue by introducing a framework that derives novel, adaptive quasi-Newton and nonlinear/Anderson acceleration schemes. Under mild assumptions, the proposed iterative methods exhibit explicit, non-asymptotic convergence rates that blend those of the gradient descent and Cubic Regularized Newton’s methods. The proposed approach also includes an accelerated version for convex functions. Notably, these rates are achieved adaptively without prior knowledge of the function’s parameters. The framework presented in this study is generic, and its special cases includes algorithms such as Newton’s method with random subspaces, finite-differences, or lazy Hessian. Numerical experiments demonstrated the efficiency of the proposed framework, even compared to the l-BFGS algorithm with Wolfe line-search.

The presented method can be seen as a mix between 1) (orthogonal) subspace Newton, 2) finite-difference approximation of the Hessian with successive gradients, 3) lazy Newton’s method and 4) cubic regularization of Newton’s method. The combination of those strategies ensures a fast global convergence rate, but reduces the per-iteration cost to $O(N^2d)$, where N is the number of gradient stored in memory.

Non-Convexity Two Ways: new bounds for the Paulsen Problem via Geodesic PL-condition – AKSHAY RAMACHANDRAN, CWI AND UNIVERSITY OF AMSTERDAM –

The Paulsen problem is a basic question in frame theory: given $u_1, \dots, u_n \in \mathbb{R}^d$ that ε -nearly satisfy the Parseval and equal-norm condition, what is the nearest frame $v_1, \dots, v_n \in \mathbb{R}^d$ that exactly satisfies these two conditions? For this question, the distance is measured as $\text{dist}(U, V) := \sum_{j \in [n]} \|u_j - v_j\|_2^2$, and the fundamental conjecture was whether $\text{dist}(U, V) \leq \text{poly}(d)\varepsilon$, namely bounded as a function independent of n . In a breakthrough result, [Kwok et al 17] were able to affirm this conjecture using techniques from the operator scaling framework of [Gurvits et al 15]. Subsequently, this was improved to the optimal bound $\text{dist}(U, V) \leq \varepsilon$ in [R 21] using a principled perspective of geodesic convex optimization over positive definite matrices. In this talk, we revisit some partial results that were known before the above resolution of the conjecture: in [Bodmann, Casazza 10] and [Casazza et al 12], it was shown that if n, d are relatively prime, then $\text{dist}(U, V) \leq \text{poly}(d, n)\varepsilon^2$. No-

tably, this is a beyond worst case bound as it contradicts the ε lower bound. In this work, we are able to re-derive these results via the perspective of geodesic optimization. Specifically, we show that the procedures developed in [BC10], [CFM 12] can be understood as gradient flows over certain non-Euclidean manifolds. Unlike the results from the scaling framework, this is still a non-convex optimization problem. Nevertheless, we are able to prove global convergence by appealing to a geodesic version of the Polyak-Lojasiewicz condition. We are also able to use these techniques to give another proof of a result in [Kwok, Lau, R 19] which shows distance bound ε^2 for the setting of random input.

Strict saddle avoidance for Riemannian Gradient Descent – ANDREEA MUSAT, EPFL – Nicolas Boumal

It is well-known that for a C^2 function with L -Lipschitz gradients, Euclidean Gradient Descent with a step size less than $1/L$ does not converge to a strict saddle point for almost all initializations.

The equivalent scenario in the Riemannian setting has received limited attention, with existing results only valid in the context of the metric projection retraction and subject to overly restrictive conditions on the step size.

In this talk, I will present some first steps towards generalizing these results in the case of Riemannian Gradient Descent with the exponential map or the metric projection as retractions, emphasizing the role of the curvature and the regularity of the cost function.

Joint work with Nicolas Boumal.

1a - FA260 - 516A - **Open-source software for nonlinear programming** Chair: CHARLIE VANARET 

Graph-Based Modeling and Optimization: Algorithms, Properties, and Software – DAVID COLE, UNIVERSITY OF WISCONSIN MADISON – Victor Zavala

The mathematical abstractions chosen to represent nonlinear optimization problems can influence the algorithms and software used for solving and modeling the problem. Graph theory is an abstraction that can be applied to a variety of nonlinear optimization problems, where nodes of a graph can contain objective functions, variables, and constraints and where edges contain constraints for variables on two or more nodes. Representing optimization problems as graphs enables unique algorithm implementations, allows for the identification of properties specific to the graph, and can be performed using open-source software. In this talk, we present how nonlinear optimization problems can be constructed using the open-source Julia package Plasmo.jl, we present different algorithms and solution schemes that can be applied to these graph-based problems, and we discuss properties specific to the graph-based representation. Overall, the graph-based representation can simplify the modeling of optimization problems and provides a framework for implementing algorithms that exploit the graph structure.

Uno, a next-gen solver for unifying nonlinearly constrained nonconvex optimization – CHARLIE VANARET, ARGONNE NATIONAL LAB ZIB – Sven

Leyffer

Derivative-based iterative methods for nonlinearly constrained nonconvex optimization usually share common algorithmic components, such as strategies for computing a descent direction and mechanisms that promote global convergence. Based on this observation, we introduce an abstract framework with four common ingredients that describes most derivative-based iterative methods and unifies their workflows. We then present Uno, a modular C++ solver that implements our abstract framework and allows the automatic generation of various strategy combinations with no programming effort from the user. Uno is meant to (1) organize mathematical optimization strategies into a coherent hierarchy; (2) offer a wide range of efficient and robust methods that can be compared for a given instance; (3) reduce the cost of development and maintenance of multiple optimization solvers; and (4) enable researchers to experiment with novel optimization strategies while leveraging established subproblem solvers and interfaces to modeling languages. We demonstrate that Uno is highly competitive against state-of-the-art solvers such as filterSQP, IPOPT, SNOPT, MINOS, LANCELOT, LOQO and CONOPT on a set of 429 problems from the CUTEst collection. Uno is available as open-source software under the MIT license at <https://github.com/cvanaret/Uno>.

Accelerating Optimal Power Flow with GPUs: SIMD Abstraction of Nonlinear Programs and Condensed-Space Interior-Point Methods – MIHAI ANITESCU, ARGONNE NATIONAL LABORATORY – Sungho Shin, Francois Pacaud

This paper introduces a framework for solving alternating current optimal power flow (ACOPF) problems using graphics processing units (GPUs). While GPUs have demonstrated remarkable performance in various computing domains, their application in ACOPF has been limited due to challenges associated with porting sparse automatic differentiation (AD) and sparse linear solver routines to GPUs. We address these issues with two key strategies. First, we utilize a single-instruction, multiple-data abstraction of nonlinear programs. This approach enables the specification of model equations while preserving their parallelizable structure and, in turn, facilitates the parallel AD implementation. Second, we employ a condensed-space interior-point method (IPM) with an inequality relaxation. This technique involves condensing the Karush–Kuhn–Tucker (KKT) system into a positive definite system. This strategy offers the key advantage of being able to factorize the KKT matrix without numerical pivoting, which has hampered the parallelization of the IPM algorithm. By combining these strategies, we can perform the majority of operations on GPUs while keeping the data residing in the device memory only. Comprehensive numerical benchmark results showcase the advantage of our approach. Remarkably, our implementations – MadNLP.jl and ExaModels.jl – running on NVIDIA GPUs achieve an order of magnitude speedup compared with state-of-the-art tools running on contemporary CPUs.

A Funnel SQP and Feasibility Solver for Direct Optimal Control within the acados Software Framework – DAVID KIESSLING, KU LEUVEN – Katrin Baumgärtner, Jonathan Frey, Moritz Diehl
We present our current developments of SQP-type solvers for direct optimal control within the acados software framework which is a versatile software framework that implements fast

solvers exploiting the specific optimal control problem structure. While established solvers within acados mainly focus on full-step algorithms to achieve fast execution times, this talk focuses on recent work towards a fast and reliable SQP algorithm. The presented approach combines a funnel to drive global convergence with an l1-relaxation of the constraints to avoid infeasible subproblems. The SQP solver is planned to be combined with a DDP-based feasibility solver in a novel two-phase feasibility framework. Rather than achieving feasibility with respect to all iterations, a feasible iterate is aimed to be found after a given time window. The novel algorithm first solves for an optimal solution and eventually switches to a feasibility phase ensuring a feasible iterate at the end of the time window. Simulation results on time-optimal control in motion planning will be presented.

1a - FB70 - 516D - **Nonlinear Programming for Data Science II** Chair: VIVAK PATEL

High order stochastic optimization without evaluating the function. – SERGE GRATTON, UNIVERSITY OF TOULOUSE, INP, IRIT, ANITI – Jerad Sadok, Philippe Toint

In this work, we consider the problem of finding approximate minimizers for unconstrained nonconvex optimization problems. Our proposed algorithm falls within the category of adaptive regularization methods where Taylor approximation of order p are used for the step computation, known for their optimal worst-case complexity in a deterministic framework. This work also contributes to the field of objective function-free optimization (OFFO), a concept motivated by the fact that many optimization methods demand finer accuracy in function evaluations than in gradient evaluations for effective step size control and global convergence. We present here an extension of our algorithms to a stochastic variant where, despite not evaluating the objective function, only stochastic approximations of the derivative are available. We prove that the excellent complexity bounds of regularization methods remain applicable in this stochastic context, even though significantly less information is used. More precisely the algorithm will find an ϵ -approximate first-order minimizer in at most $O(\epsilon^{-(p+1)/p})$ iterations and $O^{-(p+1)/(p-1)}$ for second order critical points using conditions in conditional expectation on the derivatives tensors errors. Practically, our algorithm requires dynamic control over the accuracy of the derivative, aligning well with machine learning problems where batch size offers the necessary control over approximate derivatives.

Randomized matrix decompositions for faster scientific computing – ROBERT WEBBER, CALIFORNIA INSTITUTE OF TECHNOLOGY –

Traditional numerical methods based on expensive matrix factorizations struggle with the scale of modern scientific applications. For example, kernel-based algorithms take a data set of size N , form the kernel matrix of size $N \times N$, and then perform an eigendecomposition or inversion at a cost of $O(N^3)$ operations. For data sets of size $N \geq 10^5$, kernel learning is too expensive, straining the limits of personal workstations and even dedicated computing clusters. Randomized iterative methods have emerged as a faster alternative to the classical approaches. These methods combine randomized exploration with information about which matrix structures are important, leading to significant speed gains.

In this talk, I will review recent developments concerning two randomized algorithms. The first is "randomized block Krylov iteration", which uses an array of random Gaussian test vectors to probe a large data matrix in order to provide a randomized principal component analysis. Remarkably, this approach works well even when the matrix of interest is not low-rank. The second algorithm is "randomly pivoted Cholesky decomposition", which iteratively samples columns from a positive semidefinite matrix using a novelty metric and reconstructs the matrix from the randomly selected columns. Ultimately, both algorithms furnish a randomized approximation of an $N \times N$ matrix with a reduced $rank \ll N$, which enables fast inversion or singular value decomposition at a cost of $O(Nk^2)$ operations. The speed-up factor from N^3 to Nk^2 operations can be 3 million. The newest algorithms achieve this speed-up factor while guaranteeing performance across a broad range of input matrices.

Adaptive Stochastic Optimization with Constraints – MLADEN KOLAR, UNIVERSITY OF CHICAGO –

We consider optimization problems with a stochastic objective and deterministic equality constraints, and design a trust-region sequential quadratic programming method with random models (TR-SQP-STORM) to find both first- and second-order stationary points. In our method, random models are constructed in each iteration, which require estimates of the objective value, gradient, and Hessian to satisfy adaptive accuracy conditions with a fixed probability. To reduce KKT residuals and increase the smallest eigenvalue of the reduced Lagrangian Hessian if it is negative, we design gradient steps and eigensteps, respectively. Gradient steps are computed using the adaptive relaxation technique proposed by Fang et al. (2022), while eigensteps are novel and computed in a parameter-free approach. To help the algorithm move away from saddle points, we perform an additional second-order correction step when the current iteration is deemed unsuccessful and the current iterate is close to feasible sets. Under reasonable assumptions, we establish both first- and second-order global convergence guarantees: the sequence of iterates converges to first-order stationary points almost surely, and a subsequence of iterates converges to second-order stationary points almost surely. We apply our method to a subset of problems in the CUTEst set and to constrained logistic regression problems, demonstrating its promising empirical performance.

Random subspace second order methods for nonconvex optimization – EDWARD TANSLEY, UNIVERSITY OF OXFORD – Coralia Cartis, Zhen Shao

To improve the scalability of nonconvex optimization algorithms, we devise and analyse a random subspace variant of cubic regularization that has optimal complexity with high probability and almost sure convergence, relying on key properties of sketching matrices, that allow efficient projections to low dimensions while preserving useful properties. We discuss the challenges of the analysis, in particular, the key difficulty of using sketching given the required lower bound on the step (needed for optimal complexity) that depends the random gradient at the next step. We obtain improved results for the case of low rank functions that only vary within a low dimensional subspace and that are a common occurrence in applications involving overparameterized models and that can serve as an insightful proxy for the training landscape in neural networks.

Numerical experiments with different sketching matrices are presented both in general and in the special case of low rank functions, using different sketching matrices (such as Gaussian or hashing matrices), with encouraging results. We discuss the strategies needed for careful algorithm development that go beyond vanilla application of sketching, into careful design of the subspace. Time permitting, we will explore local rates of convergence and other random subspace variants and their complexity.

1a - FB75 - 510C - The interplay between optimization, statistics and machine learning II

Chair: ALEX WANG

AdaBB: A Parameter-Free Gradient Method for Convex Optimization – SHIQIAN MA, RICE UNIVERSITY – Danqing Zhou, Junfeng Yang

We propose AdaBB, an adaptive gradient method based on the Barzilai-Borwein stepsize. The algorithm is line-search-free and parameter-free, and essentially provides a convergent variant of the Barzilai-Borwein method for general unconstrained convex optimization. We analyze the ergodic convergence of the objective function value and the convergence of the iterates for solving general unconstrained convex optimization. Compared with existing works along this line of research, our algorithm gives the best lower bounds on the stepsize and the average of the stepsizes. Moreover, we present an extension of the proposed algorithm for solving composite optimization where the objective function is the summation of a smooth function and a nonsmooth function. Our numerical results also demonstrate very promising potential of the proposed algorithms on some representative examples.

Kernel Learning in Ridge Regression “Automatically” Yields Exact Low Rank Solution – FENG RUAN, NORTHWESTERN UNIVERSITY –

This talk considers a variant of kernel ridge regression which simultaneously optimizes the prediction function and the reproducing kernel Hilbert space (RKHS) parameter. We identify a previously unnoticed phenomenon for this kernel learning objective where — under certain conditions — the global minimizer is exactly low rank with high probability. This phenomenon is interesting because the low rankness property is achieved without using any explicit regularization. Our theory makes correspondence between the observed phenomenon and the notion of low rank set identifiability from the optimization literature. The low rankness property of the finite sample solutions exists because the population kernel learning objective grows “sharply” when moving away from its minimizers in any direction perpendicular to the central mean subspace.

Overparameterized Tensor Regression via Riemannian Optimization – YUETIAN LUO, UNIVERSITY OF CHICAGO – Zhang Anru

We study the tensor regression, where the goal is to connect scalar or tensor responses to tensor covariates with a low Tucker rank parameter tensor/matrix without the prior knowledge of its intrinsic rank. We propose the Riemannian gradient descent (RGD) and Riemannian Gauss-Newton (RGN) methods and cope with the challenge of unknown rank by studying the effect of rank over-parameterization. We provide the first convergence guarantee for the general tensor regression by showing that RGD and RGN respectively converge linearly

and quadratically to a statistically optimal estimate in both rank correctly-parameterized and over-parameterized settings. Our theory reveals an intriguing phenomenon: Riemannian optimization methods naturally adapt to over-parameterization without modifications to their implementation. We also prove the statistical-computational gap in tensor regression by a low-degree polynomial argument. Our theory demonstrates a “blessing of statistical-computational gap” phenomenon in a wide range of scenarios in tensor regression for tensors of order three or higher. This shows moderate rank over-parameterization is essentially “cost-free” in terms of sample size in tensor regression of order three or higher. Finally, we conduct simulation studies to show the advantages of our proposed methods and to corroborate our theoretical findings.

The Burer-Monteiro SDP method can fail even above the Barvinok-Pataki bound – VAIDEHI SRINIVAS, NORTHWESTERN UNIVERSITY – The most widely used technique for solving large-scale semidefinite programs (SDPs) in practice is the non-convex Burer-Monteiro method, which explicitly maintains a low-rank SDP solution for memory efficiency. When the maximum allowed rank p of the SDP solution is above the Barvinok-Pataki bound ($p \geq \sqrt{2n}$, where a globally optimal solution of rank at most p is guaranteed to exist), a recent line of work established convergence to a global optimum for generic or smoothed instances of the problem. However, it was open whether there exists any instance in this regime where this method fails. We provide a family of instances on n vertices that have spurious local minima even when the rank $p = n/2$. This justifies the use of beyond worst-case paradigms like smoothed analysis to obtain guarantees for the Burer-Monteiro method. Based on joint work with Liam O’Carroll and Aravindan Vijayaraghavan.

1a - FB114 - 516A - **Computational advancement in nonconvex programming** Chair: ALEXIS MONTISON

EM Algorithms for Optimization Problems with Polynomial Objectives – JUN YA GOTOH, CHUO UNIVERSITY – Kensuke Asai

EM (Expectation-Maximization) algorithm is a popular algorithm frequently used in maximum likelihood estimation of statistical models, especially when they include latent random variables. Focusing on its MM (Majorization-Minimization) optimization features, this paper presents a new algorithmic scheme based on EM algorithm for solving general (not necessarily statistical) optimization problems by introducing certain probability distributions. In particular, we demonstrate that several optimization problems with polynomial objectives, which include linear program (LP) and quadratic program (QP), can be embedded in the EM scheme with an adequate choice of a probability distribution from the exponential family. Specifically, by employing a normal distribution, the EM algorithm for the unconstrained minimization of a convex quadratic objective turns out to be a natural gradient descent without line search. For minimizing a polynomial objective over a rectangle or a simplex, we show that the EM algorithm with a binomial or multinomial distribution, respectively, becomes an interior point algorithm which is also reduced to a natural gradient descent without line search. As further extensions to

the case where a natural gradient descent without line search is no longer available, new interior point algorithms are presented for polynomial optimization over a polytope, LP, and QP, on the basis of EM and generalized EM algorithms. The presented algorithms are somewhat conceptual in that their subprocedures require solving a nonlinear convex optimization, and some practical strategies are discussed.

Level constrained first-order methods for non-convex function constrained optimization – DIGVIJAY BOOB, SOUTHERN METHODIST UNIVERSITY – Qi Deng, Guanghui Lan

We present a new feasible proximal gradient method for constrained optimization where both the objective and constraint functions are given by the summation of a smooth, possibly nonconvex function and a convex simple function. The algorithm converts the original problem into a sequence of convex subproblems. Formulating those subproblems requires the evaluation of at most one gradient value of the original objective and constraint functions. Either exact or approximate subproblem solutions can be computed efficiently in many cases. An important feature of the algorithm is the constraint level parameter. By carefully increasing this level for each subproblem, we provide a simple solution to overcome the challenge of bounding the Lagrangian multipliers and show that the algorithm follows a strictly feasible solution path till convergence to the stationary point. We develop a simple, proximal gradient descent type analysis, showing that the complexity bound of this new algorithm is comparable to gradient descent for the unconstrained setting, which is new in the literature. Exploiting this new design and analysis technique, we extend our algorithms to some more challenging constrained optimization problems where 1) the objective is a stochastic or finite-sum function, and 2) structured nonsmooth functions replace smooth components of both objective and constraint functions. Complexity results for these problems also seem to be new in the literature. Finally, our method can also be applied to convex function-constrained problems where we show complexities similar to the proximal gradient method.

Methods of Difference-of-Convex Programming for Statistical Learning with Structured Sparsity – MIJU AHN, SOUTHERN METHODIST UNIVERSITY – Chengyu Ke

We consider a statistical learning problem where the model aims to implement pre-determined hierarchical relationships among the variables as an additional criterion in the training process. The required logical conditions are formulated exactly as the affine sparsity constraints (ASCs), then approximated as a difference-of-convex (DC) constrained program by employing existing sparsity-promoting functions. This presentation introduces numerical methods that are designed to solve variants of the latter DC optimization problem. We present the convergence analysis of the methods and numerically demonstrate the effectiveness of the proposed methods on selected statistical learning problems with logical conditions, including regression with interaction effects and tree-structured hierarchical selection. Our numerical results show that the ASC approximations can enforce the required logical conditions in the model while achieving desirable computational efficiency and model accuracy compared to existing methods.

GPU Implementation of Algorithm NCL –

ALEXIS MONTOISON, GERAD – Michael Saunders, Francois Pacaud, Sungho Shin

For constrained optimization problems, LANCELOT solves a short sequence of subproblems that minimize an augmented Lagrangian subject to bounds (and are immune to LICQ difficulties). Algorithm NCL solves equivalent subproblems that are suited to nonlinear interior methods. We focus on solving the associated KKT systems. The NCL transformation permits reduction to smaller systems that can be solved by Cholesky-type factorizations. Our NCL implementation is based on MadNLP.jl (a nonlinear optimization solver working on GPU) and CUDSS.jl (a Julia interface to the NVIDIA library cuDSS). We describe the reduction to smaller systems and present numerical experiments.

1a - FB154 - 511C - **First-order Methods and Large-scale Optimization III** Chair: LEXIAO LAI

←

Inexact restoration for minimization with inexact evaluation both of the objective function and the constraints – L FELIPE BUENO, FEDERAL UNIVERSITY OF SÃO PAULO – F. Larreal, M. Leite, J. M. Martinez

In a recent paper an Inexact Restoration method for solving continuous constrained optimization problems was analyzed from the point of view of worst-case functional complexity and convergence. On the other hand, the Inexact Restoration methodology was employed, in a different research, to handle minimization problems with inexact evaluation and simple constraints. These two methodologies are combined in the present report, for constrained minimization problems in which both the objective function and the constraints, as well as their derivatives, are subject to evaluation errors. Together with a complete description of the method, complexity and convergence results will be presented. Preliminary numerical results will also be discussed.

A Generalized Version of Chung's Lemma and its Applications – LI JIANG, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Xiao Li Li, Andre Milzarek, Junwen Qiu

Chung's lemma is a well-known tool for establishing asymptotic convergence rates of (stochastic) optimization methods under strong convexity-type assumptions and appropriate polynomial diminishing step sizes. In this talk, we propose a generalized version of Chung's lemma, which provides a simple non-asymptotic convergence analysis framework for a more general setting of step size rules. We demonstrate the broad applicability of the generalized Chung's lemma by deriving tight non-asymptotic convergence rates for a variety of stochastic approaches. Specifically, we obtain partially new non-asymptotic complexity results for stochastic optimization methods (e.g., SGD and Random Reshuffling) under a general (θ, μ) -Polyak-Lojasiewicz (PL) condition and various step sizes rules, including polynomial, constant, exponential, and cosine step sizes rules. Notably, we observe that the exponential step size can adapt to the objective function's geometry, achieving the optimal convergence rate without requiring exact knowledge of the underlying landscape.

On the Trajectories of SGD Without Replacement – PIERFRANCESCO BENEVENTANO, PRINCE-

TON UNIVERSITY –

This talk is about the implicit regularization effect of Stochastic Gradient Descent (SGD). We consider the case of SGD without replacement, the variant typically used to optimize large-scale neural networks. We analyze this algorithm in a more realistic regime than typically considered in theoretical works on SGD, as, e.g., we allow the product of the learning rate and Hessian to be $O(1)$ and we do not specify any model architecture, learning task, or loss (objective) function. Our core theoretical result is that optimizing with SGD without replacement is locally equivalent to making an additional step on a novel regularizer. This implies that the trajectory of SGD without replacement diverges from both noise-injected GD and SGD with replacement (in which batches are sampled i.i.d.). Indeed, the two SGDs travel flat regions of the loss landscape in distinct directions and at different speeds. In expectation, SGD without replacement travels flat areas and may escape saddles significantly faster. Moreover, we find that SGD implicitly regularizes the trace of the noise covariance in the eigendirections of small and negative Hessian eigenvalues. This coincides with penalizing a weighted trace of the Fisher Matrix and the Hessian on several vision tasks, thus encouraging sparsity in the spectrum of the Hessian of the loss in line with empirical observations from prior work. We also propose an explanation for why SGD does not train at the edge of stability (as opposed to GD).

Revisiting the Last-Iterate Convergence of Stochastic Gradient Methods – ZIJIAN LIU, NEW YORK UNIVERSITY – Zhengyuan Zhou

In the past several years, the last-iterate convergence of the Stochastic Gradient Descent (SGD) algorithm has triggered people's interest due to its good performance in practice but lack of theoretical understanding. For Lipschitz convex functions, different works have established the optimal $O(\log(1/\delta) \log T/\sqrt{T})$ or $O(\sqrt{\log(1/\delta)/T})$ high-probability convergence rates for the final iterate, where T is the time horizon and δ is the failure probability. However, to prove these bounds, all the existing works are either limited to compact domains or require almost surely bounded noises. It is natural to ask whether the last iterate of SGD can still guarantee the optimal convergence rate but without these two restrictive assumptions. Besides this important question, there are still lots of theoretical problems lacking an answer. For example, compared with the last-iterate convergence of SGD for non-smooth problems, only few results for smooth optimization have yet been developed. Additionally, the existing results are all limited to a non-composite objective and the standard Euclidean norm. It still remains unclear whether the last-iterate convergence can be provably extended to wider composite optimization and non-Euclidean norms. In this work, to address the issues mentioned above, we revisit the last-iterate convergence of stochastic gradient methods and provide the first unified way to prove the convergence rates both in expectation and in high probability to accommodate general domains, composite objectives, non-Euclidean norms, Lipschitz conditions, smoothness, and (strong) convexity simultaneously. Additionally, we extend our analysis to obtain the last-iterate convergence under heavy-tailed noises.

1b - MC14 - 513B - **Global Optimization** Chair: YU HONG DAI

←

Stability of Error Bounds for Convex Inequalities Systems – ZHOU WEI, HEBEI UNIVERSITY –

In this talk, we study the stability of error bounds for the convex inequalities system under small perturbations and prove primal characterizations of the stability in terms of directional derivatives. For a single convex inequality, it is proved that the stability of error bounds is essentially equivalent to the non-zero minimum of the directional derivative at the reference points over the unit sphere. For the semi-infinite convex system, the stability of error bounds is essentially equivalent to verifying that the optimal values of several min-max problems defined by the directional derivatives of the component functions in the system

International trading network optimization under extreme shocks – ZHILONG DONG, XI'AN JIAOTONG UNIVERSITY –

In this article, we portrays the global trading system as a network with countries as nodes and trade relations as edges. The general optimal equilibrium of the international trading network is obtained by constructing a nonlinear optimization model. Extreme shocks are simulated to illustrate the influence on the optimal equilibrium and suggestions are made to construct a more resilient trading network. Coping measures are proposed to alleviate pressures on the international trading network under the turbulent environment.

Efficient continuous piecewise linear regression for linearising univariate non-linear functions – JOHN WARWICKER, KIT – Steffen Rebennack

Due to their flexibility and ability to incorporate non-linear relationships, Mixed-Integer Non-Linear Programming (MINLP) approaches for optimization are commonly presented as a solution tool for real-world problems. Within this context, piecewise linear (PWL) approximations of non-linear continuous functions are useful, as opposed to non-linear machine learning-based approaches, since they enable the application of Mixed-Integer Linear Programming techniques in the MINLP framework, as well as retaining important features of the approximated non-linear functions, such as convexity. In this work, we extend upon fast algorithmic approaches for modeling discrete data using PWL regression by tuning them to allow the modeling of continuous functions. We show that if the input function is convex, then the convexity of the resulting PWL function is guaranteed. An analysis of the runtime of the presented algorithm shows which function characteristics affect the efficiency of the model, and which classes of functions can be modeled very quickly. Experimental results show that the presented approach is significantly faster than five existing approaches for modeling non-linear functions from the literature, at least 11 times faster on the tested functions, and up to a maximum speedup of more than 328,000. The presented approach also solves six benchmark problems for the first time.

1b - TA324 - 512H - **Global Optimization at work** Chair: SERGIY BUTENKO

GPU-Accelerated Deterministic Global Optimization – ROBERT GOTTLIEB, UNIVERSITY OF CONNECTICUT – Matthew Stuber

Deterministic global optimization is a powerful tool with unique value in parameter estimation and model validation tasks, among others, yet the long run-time of the standard

branch-and-bound (B&B) deterministic search algorithm limits its application to many practical problems of interest to engineers. One potential path to shorter run-times is the use of alternative computing hardware architectures such as graphics processing units (GPUs). Unlike CPUs, which can execute essentially any computing operation, GPUs are designed for a much greater throughput of a more limited set of mathematical operations. This design enables significantly greater overall computing speeds for parallelizable operations—which, within the context of GPUs, is most often tensor algebra—if the calculation volume is large enough to warrant the memory transfer between hardware elements. Being able to run B&B algorithms on GPU hardware would open the door for shorter run-times for everyday practitioners, as well as enable better utilization of supercomputers to solve larger global optimization problems, with great promise for increased tractability of problems with practical importance. Our research group has been, to the best of our knowledge, the pioneers in developing deterministic global optimization algorithms that utilize relaxation-based lower-bounding routines that run on GPUs. In recent years, we introduced parallelized and GPU-compatible methods of calculating McCormick-based relaxations and their subgradients, which are integral components of B&B lower-bounding routines. We then developed the first, to our knowledge, GPU-based B&B algorithm that utilized relaxation evaluations to calculate lower bounds, which offloaded the entire lower-bounding step—typically the most computationally expensive step in B&B—to the faster GPU hardware. These previous advances established the foundation of the novel contributions of this work. In this talk, we will introduce our latest developments in the creation of a custom linear programming (LP) solver designed to parallelize the solution of a large number of small LPs on a GPU, which enables the use of LP-based lower-bounding methods. This talk will cover the development of parallelized McCormick-based relaxations and relaxation subgradients, including details of their performance on GPUs (such as vis-a-vis GPU warp divergence characteristics), our implementation of the GPU-parallelized LP solver, and a numerical example comparing our algorithms to a state-of-the-art standard B&B solver. These developments represent a crucial first step toward enabling more efficient and faster deterministic global optimization algorithms that can exploit modern hardware's parallelization capabilities to obtain solutions that remain outside the reach of existing serial B&B approaches.

Centrality of shortest paths: algorithms and complexity results – DMYTRO MATSYPURA, THE UNIVERSITY OF SYDNEY – Johnson Phosavanh

The degree centrality of a node, defined as the number of nodes adjacent to it, is often used as a measure of importance of a node to the structure of a network. This metric can be extended to paths in a network, where the degree centrality of a path is defined as the number of nodes adjacent to it. In this paper, we consider the problem of finding the most degree-central shortest path in an unweighted network. We propose a polynomial algorithm with the worst-case running time of $O(|E||V|^2\Delta(G))$, where $|V|$ is the number of vertices in the network, $|E|$ is the number of edges in the network, and $\Delta(G)$ is the maximum degree of the graph. We conduct a numerical study of our algorithm on synthetic and real-world networks and compare our results to the existing literature. In addition,

we show that the same problem is NP-hard when a weighted graph is considered. We also extend the problem to other centrality measures and prove complexity results for these measures.

Simple Randomized Rounding for Max-Min Eigenvalue Augmentation – JOURDAIN LAMPERTSKI, UNIVERSITY OF PITTSBURGH – Haeseong Yang, Oleg Prokopyev

We consider the max-min eigenvalue augmentation problem, which captures both the maximum algebraic connectivity augmentation and E-optimal design problems. The goal in the maximum algebraic connectivity augmentation problem is to add k edges to a graph to maximize its algebraic connectivity, while the goal in the E-optimal design problem is to choose k design points that minimize the variance of the corresponding ordinary least squares estimator in its direction of greatest variance. There are two distinct research streams dedicated to these problems that develop approximation algorithms based on rounding optimal semidefinite programming relaxation solutions. Interestingly, both research streams empirically evaluate the performance of “simple” randomized rounding methods but do not establish theoretical approximation guarantees for them. In this talk we establish approximation guarantees for a simple randomized rounding method applied to the general max-min eigenvalue augmentation problem, and we investigate its empirical performance through computational experiments.

Convexification of optimization problems involving the Euclidean norm – NIKOLAOS V. SAHINIDIS, GEORGIA INSTITUTE OF TECHNOLOGY – Anatoliy Kuznetsov

Several important open problems in chemical physics, mathematics, and operations research can be posed as nonconvex global optimization problems involving the Euclidean norm. For example, pairwise potential energy minimization, object packing and cutting, the kissing number problem, and planar facility location problems admit a representation of this form. We present new formulations and convexification strategies for problems involving the Euclidean norm and use them to solve several instances to global optimality for the first time.

1b - TB5 - 512H - **Polynomial and Conic Optimization** Chair: IMMANUEL BOMZE

Strengthening convex relaxations of the radial optimal power flow problem via copositive optimization – MARKUS GABL, KARLSRUHE INSTITUTE OF TECHNOLOGY – Paul Niklas Kandora, Steffen Rebennack

Copositive reformulations can be used to derive convex underestimators of optimal value functions of QCQPs, which can be dualized to derive piecewise affine approximations. We apply this machinery to the so-called radial OPF problem. First, we reformulate the problem via indicator functions, which can be cast as optimal value functions. Using the copositivity-based convex underestimators we can derive a family of convex relaxations that are at least as strong as the popular positive semidefinite relaxation. We use a Benders Decomposition method for solving the resulting relaxations which require solving copositive subproblems. These are tackled via a cut-

ting plane algorithm where a copositivity test with respect to a special quadratic cone has to be performed in each iteration. Due to the structure of said quadratic cone, this test can be performed in polynomial time.

Uncertain standard quadratic optimization under distributional assumptions: a chance-constrained epigraphic approach – IMMANUEL BOMZE, UNIVERSITY OF VIENNA – Daniel De Vicente Jiang

The standard quadratic optimization problem (StQP) consists of minimizing a quadratic form over the standard simplex. Without convexity or concavity of the quadratic form, the StQP is NP-hard. This problem has many interesting applications. Sometimes, the data matrix is uncertain. We investigate models where the distribution of the data matrix is known but where both the StQP after realization of the data matrix and the here-and-now problem are indefinite.

SDPs for Moment Optimisation with Piecewise SOS-Convex Functions – QUEENIE HUANG, UNIVERSITY OF NEW SOUTH WALES –

In this talk, we present exact Semi-Definite Program (SDP) reformulations for a class of moment optimisation problems involving piecewise Sums-of-Squares (SOS) convex functions and projected spectrahedral support sets. These moment problems cover real-world applications such as the newsvendor and revenue maximisation problems with higher-order moments. Generally, solving moment problems is computationally intractable because evaluating a multi-dimensional integral for the expectations and searching through the infinite-dimensional space of probability distributions is numerically hard.

Our approach involves establishing an SOS representation for the non-negativity of a piecewise SOS-convex function on a projected spectrahedron and employing conic program duality. We show that both the optimal value and an optimal probability measure of the original moment problem can be found by solving a single SDP, which can be solved efficiently using commonly available software.

This talk is based on the articles by Huang, Q. Y., Jeyakumar, V., and Li, G. (2024). Exact Semi-Definite Programs for Piecewise SOS-Convex Moment Optimisation and Applications. arXiv preprint: <https://arxiv.org/abs/2402.07064> Huang, Q. Y., and Jeyakumar, V. (2024). A distributional Farkas’ lemma and moment optimisation problems with no-gap dual semi-definite programs. Optimisation Letters. DOI: <https://doi.org/10.1007/s11590-024-02097-x>

1b - WA905 - 512H - **Global Optimization 1** Chair: FABIO TARDELLA

An Adaptive Block Proximal ADMM for Weakly Convex, Linearly-Constrained Composite Functions – LEANDRO MAIA, TEXAS A&M UNIVERSITY – Renato Monteiro, David Gutman, Gilson N. Silva

This work presents an adaptive, cyclic block proximal version of the alternating direction method of multipliers (AP-ADMM) for solving linearly-constrained, weakly convex, composite optimization problems. This method is adaptive to all problem parameters, including smoothness and weak convexity constants. We assume that the smooth component of the

objective is weakly convex and possibly nonseparable, while the non-smooth component is convex, block-separable, and has easily computable block proximal maps. Each iteration of AP-ADMM consists of: (1) sequentially solving the sequence of all block proximal subproblems, and (2) performing full Lagrange multiplier and/or penalty parameter updates based on a novel update criterion. Without any rank assumptions on the constraint matrices, it is shown that AP-ADMM obtains an approximate first-order stationary point of the constrained problem in $O(\epsilon^{-3})$ iterations.

Manifolds of solutions in deterministic global optimization – HERMANN SCHICHL, UNIVERSITY OF VIENNA –

Continuous symmetries and submanifolds of solutions frequently pose problems for deterministic global optimization methods like branch-and-bound or branch-and-reduce. Usually, this leads to excessive box splitting in the vicinity of the solutions and exponential running times. We will develop verified computing methods for symmetries based on compact and non-compact matrix Lie groups and their application in the branch-and-bound framework. We furthermore present some verified continuation techniques that lead to inclusion/exclusion areas around solution manifolds.

Difference-of-Convex Piecewise Linear Delta-Approximation and Its Application in Global Optimization – XIANG LI, ZHEJIANG UNIVERSITY – Kody Kazda

Using piecewise linear functions to approximate nonlinear relationships is a widely used strategy in engineering practice. Piecewise linear approximation has also been shown to be effective in solving challenging optimization problems. In this work, we consider the problem of finding continuous piecewise-linear (CPWL) approximation of any multivariable continuous function over any compact set for any predefined error-tolerance δ while keeping the number of polytopes low. We introduce the difference-of-convex (DC) CPWL representation that represents any CPWL function as the difference of two convex CPWL functions. A major advantage of the DC-CPWL representation is that it allows polytopes of any shape, so it is more efficient in expressing CPWL functions than representations that are limited to rectangular and triangular polytopes. We develop a finitely convergent algorithm to find a DC-CPWL δ -approximator. The computational bottleneck of the algorithm is the solution of mixed-integer linear programming (MILP) fitting problems with growing sizes. To alleviate computational burden, we also develop novel algorithms that solve a sequence of linear programming (LP) subproblems instead of MILP subproblems. The advantages of the proposed DC-CPWL δ -approximation methods are demonstrated through computational studies.

A piecewise linear (PWL) δ -approximator can be trivially modified into a δ -underestimator, which leads to tight cutting planes that can be used to accelerate branch-and-bound based global optimization. We consider nonconvex operational optimization problems that include complicating functions. These problems are to be solved repeatedly in the real-time within limited time windows. We first generate the PWL δ -underestimators of the complicating functions “offline”. Then at the root node of a branch-and-bound based solution procedure, we generate a pool of valid cutting planes from the PWL

δ -underestimators of the complicating functions. At each subsequent node, one or multiple cutting planes in the pool may be added to the problem, depending on how effective they can tighten the convex relaxation problem. We will demonstrate the advantages of the proposed solution method through preliminary computational results.

Location of optima and vertex optimality – FABIO TARDELLA, UNIVERSITY OF FLORENCE –

We present new necessary conditions for the existence of a global minimizer of a nonlinear function on a polyhedron in the relative interior of one of its faces. As a consequence, we obtain an extension to nonlinear and to quadratic programming of the classical fundamental theorem of linear programming which guarantees the existence of a global minimum at a vertex of the feasible region. In the quadratic case we also provide a simple characterization of the class of those functions whose restriction to each face of a polyhedron attains its minimum at a vertex of such face.

1b - WB350 - 512H - Optimization on manifolds
Chair: DAVID GUTMAN

An optimized composite gradient method for minimizing the gradient mapping – MIHAI I. FLOREA, UCLOUVAIN –

The gradient mapping norm is a strong and easily verifiable stopping criterion for first-order methods on composite problems. When the objective exhibits the quadratic growth property, the gradient mapping norm minimization problem can be solved by online, parameter-free and adaptive first-order schemes with near-optimal worst-case rates. No such methods can be applied when quadratic growth is absent. In this work, we propose a quasi-online parameter-free method addressing the entire class of composite problems with state-of-the-art optimal worst-case rates. Preliminary simulation results suggest that our scheme is highly competitive in practice with the existing approaches, even with those that rely on regularization and restarting strategies.

Regularized Stein Variational Gradient Flow – YE HE, GEORGIA INSTITUTE OF TECHNOLOGY –

The Stein Variational Gradient Descent (SVGD) algorithm is a deterministic particle method for sampling. However, a mean-field analysis reveals that the gradient flow corresponding to the SVGD algorithm (i.e., the Stein Variational Gradient Flow) only provides a constant-order approximation to the Wasserstein Gradient Flow corresponding to the KL-divergence minimization. In this work, we propose the Regularized Stein Variational Gradient Flow, which interpolates between the Stein Variational Gradient Flow and the Wasserstein Gradient Flow. We establish various theoretical properties of the Regularized Stein Variational Gradient Flow (and its time-discretization) including convergence to equilibrium, existence and uniqueness of weak solutions, and stability of the solutions. We provide preliminary numerical evidence of the improved performance offered by the regularization.

Tangent Subspace Descent via Discontinuous Subspace Selections on Fixed-Rank Manifolds – DAVID GUTMAN, TEXAS A&M UNIVERSITY – Nam Ho Nguyen, Daniel Grady

The tangent subspace descent method (TSD) extends the coordinate descent algorithm to manifold domains. The key insight underlying TSD is to draw an analogy between coordinate blocks in Euclidean space and tangent subspaces of a manifold. The core principle behind ensuring convergence of TSD for smooth functions is the appropriate choice of subspace at each iteration. Previously, it was shown that it is always possible to appropriately pick such subspaces on the broad class of manifolds known as naturally reductive homogeneous spaces. In this talk, we provide the first instances of TSD for manifolds outside of this class. The main idea underlying these new instances is the use of discontinuous subspace selections. As a result of our developments we derive new and efficient methods for large-scale optimization on the fixed-rank and fixed-rank positive semidefinite matrix manifolds.

Randomized Riemannian Submanifold Subgradient Method for Optimization over Stiefel Manifold – MAN CHO SO, CHINESE UNIVERSITY OF HONG KONG –

In this talk, we present the randomized Riemannian submanifold subgradient method (RSSM), a lightweight "block-coordinate"-type algorithm for weakly convex optimization over the Stiefel manifold. We show that RSSM finds an ϵ -nearly stationary point in $O(\epsilon^{-4})$ iterations. To the best of our knowledge, this is the first convergence guarantee of a coordinate-type algorithm for tackling non-convex non-smooth optimization over the Stiefel manifold.

This is joint work with Andy Yat-Ming Cheung, Jinxin Wang, and Man-Chung Yue.

1b - ThA24 - 512H - Recent Advances in Global Optimization Chair: ANA ROCHA

On a Penalty Weighted Tchebycheff Scalarizing Function for Multiobjective Optimization

– ANA ROCHA, UNIVERSITY OF MINHO – M. Fernanda P. Costa, Edite M.G.P. Fernandes Fernandes, A. Gaspar Cunha

When solving a multiobjective optimization problem (MOP) one single optimal solution does not exist. Instead, a set of non-dominated solutions can be obtained, the well-known Pareto optimal front (POF). Decomposition-based methods decompose a MOP into a set of subproblems, each one associated to a weight vector, using a scalarizing function to combine the multiple objectives into a single objective function. Since the position of the contour lines of the scalarizing function is crucial, a penalty-based weighted Tchebycheff scalarizing function based on the normalized weight vectors is proposed in order to promote convergence to the POF and maintain diversity of solutions on the POF. The influence of the penalty parameter on the opening angle of the contour lines and on the improvement region of a solution is analyzed. Appropriate values for the penalty parameter are obtained. The proposed algorithm solves each subproblem by the Simulated Annealing to compute a candidate solution. The process is repeated randomly a certain number of times and the final non-dominated solutions are used to approximate the POF. The tests carried out to optimize a polymer single screw extrusion with two objectives show that the obtained solutions converge well to the POF and have a very good distribution.

Investigating polytopal global optimization branch and bound – ELIGIUS MT HENDRIX, UNIVERSIDAD DE MÁLAGA –

In the field of Interval Arithmetic Branch and Bound (B&B) monotonicity detection and its use for dimension reduction and eliminations of partition sets has a long tradition. Recent investigation of extending ideas to simplicial B&B has shown that the relative success depends on the dimension of the simplex, i.e. the dimension of its affine hull. Linear constraints may provide a feasible area which is a polytope in a lower dimension due to equality constraints. We investigate the question how to deal with subsets defined by polytopes. We study how to construct a converging branch and bound, how to define bounds and how to refine and how to exploit monotonicity when bounds are provided by the interval hull.

Column generation for multistage stochastic mixed-integer nonlinear programs with discrete state variables – TUSHAR RATHI, UNIVERSITY OF MINNESOTA – Benjamin P. Riley, Angela Flores Quiroz, Qi Zhang

Sequential decision-making is crucial in a wide range of industrial problems, including energy expansion planning, power systems operation, and inventory management. We often need to account for uncertainty in model parameters when determining future policies, and multistage stochastic programming provides a powerful framework for doing so. However, solving these models can be challenging, forcing decision makers to rely on heuristics that may produce suboptimal decisions and make it difficult to assess their quality. Two common reasons for the difficulty in solving these problems are discrete decisions such as the installation/expansion of new/existing production units or inventory decisions in discrete production settings, and complex nonlinear (often nonconvex) relationships between decision variables, such as nonlinear cost functions or power flow equations; this leads to mixed-integer nonlinear programming (MINLP) models, which are among the most difficult class of optimization problems to solve. In this work, we address these issues by proposing an exact method for solving multistage stochastic MINLPs with discrete state variables using a tailored column generation (CG) scheme.

Multistage stochastic MINLPs with discrete state variables have an inherently decomposable structure that allows the use of CG-based decomposition. Following a Dantzig-Wolfe reformulation, we apply CG with each pricing subproblem being an MINLP of much smaller size, making it more amenable to global solvers such as Gurobi and BARON. Furthermore, the independent subproblems can be solved in parallel, significantly reducing the solution time. The well-known convergence issues in CG, namely the heading-in and tailing-off effects, are mitigated by column sharing, an information sharing technique among the subproblems that ensures nonanticipativity and allows for the generation of many more feasible columns in each CG iteration. Lastly, we demonstrate the effectiveness of our tailored decomposition scheme through case studies on multistage blending and optimal power flow problem.

Advances in edge-concave relaxation based global optimization – FARUQUE HASAN, TEXAS A&M UNIVERSITY –

A new relaxation technique is developed for the deterministic global optimization of twice-differentiable continuous prob-

lems. Instead of using a convex underestimator, the relaxation is based on an underestimator which is edge-concave. The underestimator is constructed by subtracting a positive quadratic expression. We use the linear facets of the vertex polyhedral convex envelope of the edge-concave underestimator to obtain a linear programming (LP)-based relaxation of general nonconvex functions. The method will be presented with theoretical results and will be compared with convexification/ underestimation techniques such as α BB and its variants through test examples.

1b - ThB26 - 512H - Large-Scale and Structured Optimization Chair: NAM HO NGUYEN

A procedure for listing KKT points of a quadratic reverse convex programming problem – SYUJI YAMADA, NIIGATA UNIVERSITY –

In this talk, we propose a procedure for listing KKT(Karush-Kuhn-Tucker) points of a quadratic reverse convex programming problem (QRC) whose feasible set is expressed as the area excluded the interior of a convex set from another convex set. Several types of iterative solution methods for solving (QRC) have been proposed by many other researchers. However, such algorithms are not effective in the case where the dimension of variables is so large. For such problems, we introduce an algorithm for listing KKT points of (QRC). Moreover, by combining our algorithm into a branch and bound procedure, we obtain most of KKT points of (QRC). By utilizing our algorithm, we can calculate most of locally optimal solutions contained in the intersection of the boundaries of convex sets defining the feasible set. Moreover, by choosing a calculated locally optimal solution having the smallest value of the objective function, we can obtain an approximate solution of a globally optimal solution is obtained. Furthermore, to improve calculation efficiency of our algorithm, we propose an update method of Lagrange multipliers for convex constraint conditions. The effectiveness of the improvement has been shown by the result of the computer experiment.

Continuous Approaches to Discrete Optimization Problems – MILTIADES PARDALOS, TEXAS A&M UNIVERSITY – Sergiy Butenko

We discuss continuous formulations for several important discrete optimization problems. The proposed formulations can be used to develop analytical bounds as well as effective algorithms for some of the problems. Moreover, a novel hierarchy of nonconvex continuous reformulations of discrete optimization problems is discussed.

Viral Quasispecies Analysis via Tensor Factorization – QIAN ZHANG, SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN – Meixia Lin

RNA viruses exhibit high mutation rates, leading to the formation of new viral strains through point mutations, insertions, and deletions, collectively known as a viral quasispecies. Despite the advancements enabled by high-throughput sequencing (HTS) in quasispecies analysis, challenges arise from sequencing errors and limited read lengths. The reconstruction of viral quasispecies is particularly challenging due to the unknown number of strains, non-uniform strain frequencies, and sparse observed data with missing elements. This difficulty is further amplified when the diversity among the strains is small. Our research introduces an algorithm that employs a cluster-

ing and tensor factorization framework to analyze sparse tensor data derived from aligned sequencing reads, aiming to reconstruct viral quasispecies characterized by highly uneven component frequencies. Specifically, we utilize an iterative clustering framework to estimate the number of strains. Within each iteration, we utilize an alternating least-square method, incorporating missing data imputation, to solve the corresponding sparse tensor factorization problem and reconstruct the strains of quasispecies. Our results, obtained from simulated datasets with small distances between strains, demonstrate the capability of our method to accurately estimate frequencies and reconstruct strains characterized by highly uneven frequencies. Generally, our method outperforms state-of-the-art approaches, particularly at diversities of around 1%.

Projection-Free Methods for Solving Convex Bilevel Optimization Problems – NAM HO NGUYEN, THE UNIVERSITY OF SYDNEY – Khanh Hung Giang Tran, Dabeen Lee

When faced with multiple minima of an "inner-level" convex optimization problem, the convex bilevel optimization problem selects an optimal solution which also minimizes an auxiliary "outer-level" convex objective of interest. Bilevel optimization requires a different approach compared to single-level optimization problems since the set of minimizers for the inner-level objective is not given explicitly. Furthermore, Slater constraint qualification is never satisfied for this problem, thus strong duality is not guaranteed to hold in many cases. In this paper, we propose two new projection-free methods for convex bilevel optimization which require only a linear optimization oracle over the base domain. The first method employs the iterative regularization idea to form a blended objective with regularization parameter determined by a fixed schedule. The second method interprets the regularization parameter as a Lagrangian dual variable, and employs estimates of the inner-level optimal objective to allow for dynamic adjustment of this parameter throughout the course of the algorithm. We provide convergence guarantees for both inner- and outer-level objectives that hold under our proposed projection-free methods. In particular, we highlight how our guarantees are affected by the presence or absence of an optimal dual solution. Lastly, we conduct numerical experiments that demonstrate the performance of the proposed methods.

1b - FA27 - 512H - Global Optimal Search: Unleashing Maximum Potential Chair: FREDERIC MESSINE

Computing the conjugate of nonconvex bivariate piecewise linear-quadratic functions. – TANMAYA KARMAKAR, UBC –

Computing the minima of a convex function is much easier than computing the same for a nonconvex function. Since every nonconvex function shares the same minima as its convex envelope, we compute the conjugate of a bivariate piecewise linear-quadratic (PLQ) function as a first step toward the computation of the convex envelope.

Our algorithm starts with computing the convex envelope of each piece obtaining a rational function defined over a polyhedral subdivision. Then we compute the conjugate of each of those pieces and obtain a fractional form defined over a parabolic subdivision. The last step is to compute the maxi-

imum of all those functions to obtain the conjugate of the original piecewise linear-quadratic function as a piecewise function defined on a parabolic subdivision.

Our implementation in MATLAB uses symbolic computation and rational numbers to avoid any floating-point errors.

Norm-induced Non-Convex Relaxations for Lipschitz Constraints – ADRIAN GÖSS, UNIVERSITY OF TECHNOLOGY NUREMBERG (ÜTN) – Alexander Martin, Sebastian Pokutta, Kartikey Sharma
In this paper, we consider a finite dimensional optimization problem minimizing a continuous objective on a compact domain subject to a multi-dimensional constraint function. For the latter, we only assume the availability of a Lipschitz property. Recent literature proposes methods based on non-convex outer approximation for one dimensional equality constraints on bounded polyhedral domains, which are Lipschitz with respect to the maximum norm. To the best of our knowledge, however, there exists no formulation of such a non-convex outer approximation method for a general problem class. We introduce a meta-level solution framework to solve this class and tackle the underlying theoretical foundations. Considering the feasible domain without the constraint function as manageable, our method relaxes the multidimensional constraint and iteratively refines the feasible region by means of norm-induced cuts, relying on an oracle for the resulting sub-problems. We show the method’s correctness and investigate the problem complexity. In order to account for discussions about functionality, limits, and extensions, we present computational examples including illustrations.

On the impact of reliable boundary calculations by convex relaxation in global optimization – FREDERIC MESSINE, LAPLACE UNIVERSITY OF TOULOUSE – Gilles Trombettoni

In order to obtain reliable deterministic global optima, all computed bounds must be certified in such a way that floating-point operations cannot cause a Branch and Bound algorithm to converge to a wrong solution. Branch-and-bound codes based on interval arithmetic possess this reliability property. However, some efficient acceleration techniques, such as convex relaxation techniques, cannot be used directly in such a global optimization code without losing the reliability property. In this work, we explain how to correct these bounds in order to make convex relaxation methods reliable for solving constrained and unconstrained global optimization problems. In addition, we explain how to exploit them within an interval branch-and-bound code and show that the effort made to provide a certified and reliable solution is not prohibitive.

Advances in convex relaxation generation for dynamic global optimization – KAMIL KAMIL KHAN, MCMASTER UNIVERSITY – Yulan Zhang
Deterministic methods for global optimization proceed by generating upper and lower bounds on the unknown globally optimal value, and then refining these bounds. In this setting, lower bounds are typically computed by minimizing convex relaxations of the original problem using local solvers. This is challenging for dynamic global optimization, however, where embedded ordinary differential equations (ODEs) can complicate convex relaxation generation. In this presentation, we describe recent approaches for automatically constructing ef-

fective convex relaxations for parametric ODE solutions, for use in global dynamic optimization. These approaches make use of embedded optimization problems in new ways, and have been implemented in Julia.

1b - FB1 - 512H - Recent Advances in Global Optimization of Nonconvex Programs Chair: HARSHA NAGARAJAN 

Discretization Approaches for Non-Convex Mixed-Integer Nonlinear Programs – ROBERT BURLACU, UNIVERSITY OF TECHNOLOGY NUREMBERG – Ben Beach, Kristin Braun, Lukas Hager, Robert Hildebrand, Thomas Kleinert

Solving MINLPs is to this day a very popular and challenging topic. Over the last two decades, mixed-integer programming relaxation techniques have increasingly established themselves as a serious alternative to classical spatial branch-and-bound. The underlying idea is to relax the non-convex nonlinearities in a linear-discrete fashion while controlling the relaxation accuracy. By using integer variables to model these relaxations, the discrete approaches aim at exploiting the availability of mature solvers for mixed-integer programs. Quadratically constrained quadratic programs are probably the most important subclass of mixed-integer nonlinear programs. The talk will include both a broad overview and some of the recent developments in the field, focusing largely on the handling of quadratic monomials and bivariate products. A best practice recommendation with numerical results based on the QPLIB and the MINLPLib will complete the presentation.

Constructing Tight Quadratic Underestimators for Global Optimization – ARVIND RAGHUNATHAN, MITSUBISHI ELECTRIC RESEARCH LABORATORIES – William Strahl, Nikolaos Sahinidis, Chrysanthos Gounaris

Global optimization algorithms extensively utilize polyhedral outer approximations to determine lower bounds. Noting recent improvements in algorithms that solve quadratically constrained quadratic programming problems, we investigate introducing quadratic non-linearity into outer approximation methods. We generalize a 2nd-order Taylor series approximation to create a hierarchy of convex quadratic outer approximators and an algorithm to construct them for twice differentiable convex functions and non-convex difference of convex (d.c.) functions. We demonstrate the quality of our quadratic outer approximators on several example functions extracted from global optimization benchmark libraries. Furthermore, we show that for systematically generated d.c. optimization problems defined in greater than one dimension, the convex QCQP relaxations constructed using our methodology determine tighter root node lower bounds compared to lower bounds computed by the state-of-the-art global optimization solver BARON.

A Generalizable End-to-End Learning Approach for Global Optimization of Quadratically-Constrained Quadratic Programs – ERIN GEORGE, UCLA MATHEMATICS – Rohit Kannan, Harsha Nagarajan, Deepjyoti Deka
Quadratically-constrained quadratic programs (QCQPs) are a broad class of non-convex optimization problems which occur in applications such as network optimization and the pooling

problem, but are often slow to solve in practice. We present a neural network approach that can learn partitioning heuristics which result in significant time saves for existing global optimization solvers. We construct a novel end-to-end loss function so that the desired attributes of an output partitioning scheme can be learned directly, without reference to a pre-selected heuristic. Additionally, we describe how to use the graph scattering transform to produce a single model which works on a variety of different problem structures and sizes. We demonstrate the neural network is able to learn heuristics that speed up QCQP optimization significantly, greatly exceeding those of previous work. Furthermore, we demonstrate the model is able to learn heuristics which generalize well even to problem sizes not seen in training.

Piecewise Polyhedral Relaxations of Multilinear Optimization – MOHIT TAWARMALANI, PURDUE UNIVERSITY – Jongeun Kim, Jean Philippe P. Richard

We develop piecewise polyhedral relaxations for multilinear optimization problems over hyper-rectangular partitions. We interpret associated multipliers as multilinear terms, which allows us to derive linking constraints across relaxations of different multilinear terms. The associated cuts have been incorporated within ALPINE, an open-source software for global optimization that relies on piecewise approximations of functions. The resulting cuts speed-up the solver by an order of magnitude and enable solution of many more instances from our test-set. We also consider non-regular partitions which are particularly useful when the functions are modeled using linear regression trees. We show that resorting to regular partitioning for a problem that was natively described using non-regular partitions can lead to an exponential blowup in the number of partitions. We give the first locally ideal formulation for non-regular hyper-rectangular partitions and show that for certain tree ensemble optimization problems, the resulting formulation performs significantly better than alternative formulations.

1c - MB71 - 510B - **Inertial and stochastic algorithms** Chair: SEBASTIAN KASSING

Convergence of Momentum Stochastic Gradient Descent – SEBASTIAN KASSING, UNIVERSITY OF BIELEFELD – Benjamin Gess

In this talk, we discuss optimizing a non-convex objective function using Momentum Stochastic Gradient Descent (MSGD) in the context of supervised learning. Empirical observations motivate the long-standing conjecture that including a momentum term improves the performance of stochastic optimization algorithms. However, there are only few theoretical results proving the advantage of these methods. We present a framework that allows the theoretical analysis of MSGD. Using local Łojasiewicz-inequalities we show almost sure convergence of the MSGD scheme under classical assumptions on the noise as well as in the overparametrized regime. These results are made possible by a careful choice of a Lyapunov function and rely on the derivation of polynomial (for classical noise), respectively exponential (in the overparametrized setting), convergence rates. In the overparametrized and small learning-rate regime, we prove explicit bounds on the exponential rate of convergence for the MSGD scheme. Moreover, we analyze

the optimal friction parameter and prove that, for the optimal choice of parameters, MSGD indeed speeds up convergence in the case of flat minima. This result extends the classical work of Polyak (1964) which showed a beneficial effect of Momentum in the optimization of a strictly convex objective function using deterministic gradient descent.

Convergence of a Normal Map-Based Prox-SGD Method for Stochastic Composite Optimization – JUNWEN QIU, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Andre Milzarek

In this talk, we present a novel stochastic normal map-based algorithm (nor-SGD) for nonsmooth nonconvex composite-type optimization problems and discuss its asymptotic convergence properties. Using the time window-based strategy and a suitable merit function, we first analyze the global convergence behavior of nor-SGD and show that every accumulation point of the generated sequence of iterates is a stationary point almost surely and in an expectation sense. The obtained results hold under standard assumptions and extend the more limited convergence guarantees of the basic proximal stochastic gradient method. In addition, based on the well-known Kurdyka-Łojasiewicz (KL) analysis framework, we provide novel pointwise convergence results for the iterates generated by nor-SGD. In the meanwhile, we have derived convergence rates that depend on the KL exponent and the step size dynamics. The obtained rates are faster than related and existing convergence rates for SGD in the nonconvex setting. It is worth noting that the obtained asymptotic rates align with those in the strongly convex scenario when the KL exponent is less than or equal to $1/2$. The techniques studied in this work can be potentially applied to other families of stochastic and simulation-based algorithms.

Stochastic composition optimization in the absence of Lipschitz continuous gradient – SAM DAVANLOO TAJBAKSHI, THE OHIO STATE UNIVERSITY – Yin Liu

This talk considers the optimization of the nested composition of two functions where at least the inner function has the expectation form. In such nested structures, obtaining unbiased estimates of the gradient of the composition is challenging. Nested stochastic optimization has a range of recent applications in machine learning, particularly in reinforcement learning and meta-learning. We develop stochastic algorithms for composition optimization with convergence guarantees in the absence of the Lipschitz continuity of the gradient of the inner and/or outer functions. More specifically, the smoothness property is generalized by the notion of relative smoothness which provokes the Bregman gradient method. We propose three Stochastic Composition Bregman Gradient algorithms for the three possible relatively smooth compositional scenarios and provide their sample complexities to achieve an ϵ -approximate stationary point. For the smooth of relatively smooth composition, the first algorithm requires $O(\epsilon^{-2})$ calls to the stochastic oracles of the inner function value and gradient as well as the outer function gradient. When both functions are relatively smooth, the second algorithm requires $O(\epsilon^{-3})$ calls to the inner function value stochastic oracle and $O(\epsilon^{-2})$ calls to the inner and outer functions gradients stochastic oracles. We further improve the second algorithm by variance reduction for the setting where just the inner function is smooth. The resulting algorithm requires $O(\epsilon^{-5/2})$ calls to the inner

function value stochastic oracle, $O(\epsilon^{-3/2})$ calls to the inner function gradient, and $O(\epsilon^{-2})$ calls to the outer function gradient stochastic oracles. Some numerical studies on the performance of the proposed algorithms will be presented.

Adam-family Methods with Decoupled Weight Decay in Deep Learning – KUANGYU DING, NATIONAL UNIVERSITY OF SINGAPORE – Nachuan Xiao, Kim Chuan Toh

In this paper, we investigate the convergence properties of a wide class of Adam-family methods for minimizing quadratically regularized nonsmooth nonconvex optimization problems, especially in the context of training nonsmooth neural networks with weight decay. Motivated by the AdamW method, we propose a novel framework for Adam-family methods with decoupled weight decay. In our framework, the estimators for the first-order and second-order moments of stochastic subgradients are updated independent of the weight decay term. Under mild assumptions, we prove that our framework converges with non-diminishing stepsizes to the variables. In addition, we show that our proposed framework encompasses a wide variety of well-known Adam-family methods, hence offering convergence guarantees for these methods in the training of nonsmooth neural networks. More importantly, we prove that our proposed framework asymptotically approximates the SGD method, thereby providing an explanation for empirical observations that decoupled weight decay enhances generalization performance for Adam-family methods. As a practical application of our proposed framework, we propose a novel Adam-family method named Adam with Decoupled Weight Decay (AdamD), and establish its convergence properties under mild conditions. Extensive numerical experiments exhibit that AdamD outperforms Adam and is comparable to AdamW, in the aspects of both generalization performance and efficiency.

1c - MB92 - 511A - **Douglas–Rachford Awesome Cones, Great Job!** 1 Chair: LUIZ RAFAEL SANTOS

On tight error bounds for conic optimisation – SCOTT BOIVIN LINDSTROM, CURTIN UNIVERSITY – Ying Lin, Bruno Lourenço, Ting Kei Pong

Conic-linear programming allows researchers to tackle many important problems. Popular programs such as Mosek, CVX, and Hypatia solve problems with conic-linear methods. Such software packages provide users with so-called “backward error bounds,” which may not be useful accuracy guarantees. I will describe the framework that we built for obtaining useful “forward” bounds, and explain how we have used it to find such bounds classes of next generation cones in Mosek and Hypatia. I will also discuss how such bounds also admit convergence rate guarantees for projection methods.

On particularly awesome cones – VERA ROSHCHINA, UNSW SYDNEY –

Awesome cones were first introduced by Scott Lindstrom in 2023, with closed convex cones known to be particularly awesome.

This talk is focused on different ways to describe and capture the awesomeness (or lack thereof) of closed convex cones, including such properties as facial exposure, facial dual completeness, amenability and projectional exposure. In Euclidean

spaces of dimensions four, five and higher some of the common geometric insights that guide our understanding of the three-dimensional objects become increasingly more difficult to rely on as dimensions grow, while phenomena that are impossible in three-dimensions emerge in the facial structure of higher-dimensional cones.

These difficulties call for new tools that can help us understand, study and construct closed convex cones with peculiar facial structure in higher dimensional spaces. For instance, the facial structure of these cones can be modified in subtle ways to preserve important properties and at the same time simplify the analysis, the facial structure of four-dimensional cones can be studied using three-dimensional affine cuts, and the facial structure of five-dimensional cones can be handled using a combination of four-dimensional cuts and appropriately mapping the three-dimensional boundaries of these cuts to the three-dimensional Euclidean space.

Proximal splitting methods avoid strict saddle points of weakly convex problems – FELIPE ATENAS, THE UNIVERSITY OF MELBOURNE –

For nonconvex problems, critical points may not necessarily be local minimizers. In practice, it has been observed that commonly used methods find good-quality critical points, even close to optimality. For well-structured nonconvex problems, first-order methods and the proximal point algorithm can avoid strict saddle points with high probability. In this talk, we present how proximal splitting methods, such as Douglas-Rachford and other proximal-type splitting methods, inherit this property. More specifically, we discuss how proximal splitting methods converge to local minimizers of semialgebraic weakly convex optimization problems, when randomly initialized.

Solving large-scale basis pursuit by infeasible alternating projections – LUIZ RAFAEL SANTOS, UNIVERSIDADE FEDERAL DE SANTA CATARINA –

Roger Behling, Yunier Bello Cruz, Paulo J. S. Silva
Basis pursuit is the problem of finding a vector with smallest 1-norm among the solutions of a given linear system of equations. It is a famous convex relaxation of what the literature refers to as sparse affine feasibility (SAF) problem, in which sparse solutions to undetermined systems are sought. Problem SAF is a fundamental problem in Compressed Sensing (CS), which is a technique for recovering sparse signals from incomplete measurements. While SAF is known to be NP-hard, there are some instances where the solution of (BR) and coincide. The importance of basis pursuit led to a great deal of research devoted to the development of efficient methods for solving it, mainly for large-scale problems. There are several methods for solving (BP). One can think, for instance, in recasting it as a linear program (LP) and applying to the reformulation any standard LP solver.

In turn, we tackle the basis pursuit in its very original form with a scheme that uses alternating projections in its subproblems. These subproblems are designed to be inconsistent in the sense that they relate to two non-touching sets. Quite recently, inconsistency coming from infeasibility has been seen to work in favor of alternating projections and correspondent convergence rates. In our work, this feature is now suitably enforced in a new and numerically competitive method for solving the basis pursuit.

1c - MC44 - 510B - **Design of New Nonsmooth Optimization Algorithms** Chair: BENJAMIN GRIMMER

Primal-Dual First-Order Methods for Convex-Concave Semi-Infinite Programs – QIHANG LIN, UNIVERSITY OF IOWA – Yao Yao, Tianbao Yang

This paper explores numerical methods for solving a convex-concave semi-infinite program (SIP). SIP has been systematically discussed by optimization community for many years with many real-world applications. Most numerical algorithms for optimization with finitely many constraints can be somehow extended for SIP. The examples include the penalty method, the interior-point method, the Lagrangian method, and the trust-region method. However, most of these works only show an asymptotic convergence property or a local convergence rate. On the contrary, we establish the global non-asymptotic convergence rate for our algorithm and characterize its iteration complexity for finding an ϵ -optimal solution. We introduce a primal-dual gradient method which performs three updates iteratively: a momentum gradient ascend step to update the constraint parameters, a momentum gradient ascend step to update the dual variables, and a gradient descend step to update the primal variables. Our approach also extends to scenarios where gradients and function values are accessible solely through stochastic oracles. We show the iteration complexity under different smoothness, convexity and concavity assumptions on the functions. In addition, we explore the application of the proposed algorithms on training machine learning models with fairness constraints.

First-Order Methods for Nonsmooth Nonconvex Functional Constrained Optimization with or without Slater Points – ZHICHAO JIA, GEORGIA INSTITUTE OF TECHNOLOGY – Benjamin Grimmer

Constrained optimization problems where both the objective and constraints may be nonsmooth and nonconvex arise across many learning and data science settings. In this talk, we discuss a simple first-order method that for any Lipschitz, weakly convex objectives and constraints, it finds a feasible, ϵ -stationary point at a convergence rate of $O(\epsilon^{-4})$ without relying on compactness or Constraint Qualification (CQ). When CQ holds, this convergence is measured by approximately satisfying the Karush-Kuhn-Tucker conditions. When CQ fails, we guarantee the attainment of weaker Fritz-John conditions. As an illustrative example, our method stably converges on piecewise quadratic SCAD regularized problems despite frequent violations of constraint qualification. The considered algorithm is similar to those of "Quadratically regularized subgradient methods for weakly convex optimization with weakly convex constraints" by Ma et al. and "Stochastic first-order methods for convex and nonconvex functional constrained optimization" by Boob et al. (whose guarantees further assume compactness and CQ), iteratively taking inexact proximal steps, computed via an inner loop applying a switching subgradient method to a strongly convex constrained subproblem. Our non-Lipschitz analysis of the switching subgradient method appears to be new and may be of independent interest.

Variance Reduced Halpern Iteration for Op-

timal Finite-Sum Min-Max Optimization – AHMET ALACAOGU, UNIVERSITY OF WISCONSIN MADISON – Xufeng Cai, Jelena Diakonikolas

This talk will consider finite-sum monotone inclusion problems, which is a generalized framework for min-max optimization and variational inequalities. We will focus on guarantees on residual and operator norm. These optimality measures generalize the notion of "gradient norm" from minimization to more general problems we consider. We will cover algorithms incorporating variance reduction on Halpern iteration that achieve optimal complexity guarantees not only in terms of the desired accuracy, but also number of functions involved in the finite-sum and Lipschitz constant. We will discuss further consequences of the result and open problems.

1c - TA65 - 510B - **Nonlinear optimization problems** Chair: WEI BIAN

Convergence and error estimation of smoothing consensus-based optimization algorithm for nonsmooth nonconvex optimization – WEI BIAN, HARBIN INSTITUTE OF TECHNOLOGY –

Global optimization has received increased attention across almost every field of economy, science and artificial intelligence. Lately, a novel swarm intelligence model, namely the consensus-based optimization (CBO) algorithm, was introduced to deal with the global optimization problems. Limited by the conditions of Ito's formula, the convergence analysis of the previous CBO algorithms mainly focus on the problem with smooth objective function. With the help of smoothing method, this paper achieves a breakthrough by proposing an effective CBO algorithm for solving the global solution of a nonconvex, nonsmooth, and possible non-Lipschitz continuous minimization problem with theoretical analysis. We indicate that the proposed algorithm exhibits a global consensus with any initial data, and then give a more detailed error estimation on the objective function values from the proposed algorithm towards the global minimum. Inspired by the smoothing techniques, we give a sufficient condition on the system parameters and initial data to guarantee that the function value at global consensus state can be close enough to the global minimum. And note that the conditions on the parameters are independent of the dimension of the problem. Finally, some numerical examples are presented to illustrate the appreciable performance of the proposed method on solving the nonsmooth, nonconvex minimization problems.

On solving a rank regularized minimization problem via equivalent factorized column-sparse regularized models – WENJING LI, HARBIN INSTITUTE OF TECHNOLOGY – Wei Bian, Kim Chuan Toh

Rank regularized minimization problem is an ideal model for the low-rank matrix completion/recovery problem. The matrix factorization approach can transform the high-dimensional rank regularized problem to a low-dimensional factorized column-sparse regularized problem. The latter can greatly facilitate fast computations in applicable algorithms, but needs to overcome the simultaneous non-convexity of the loss and regularization functions. In this talk we consider the factorized column-sparse regularized model. Firstly, we optimize this model with bound constraints, and establish a certain equivalence between the optimized factorization problem and rank

regularized problem. Further, we strengthen the optimality condition for stationary points of the factorization problem and define the notion of strong stationary point. Moreover, we establish the equivalence between the factorization problem and its a nonconvex relaxation in the sense of global minimizers and strong stationary points. To solve the factorization problem, we design two types of algorithms and give an adaptive method to reduce their computation. The first algorithm is from the relaxation point of view and its iterates own some properties from global minimizers of the factorization problem after finite iterations. We give some analysis on the convergence of its iterates to the strong stationary point. The second algorithm is designed for directly solving the factorization problem. We improve the PALM algorithm introduced by Bolte et al. (Math Program Ser A 146:459-494, 2014) for the factorization problem and give its improved convergence results. Finally, we conduct numerical experiments to show the promising performance of the proposed model and algorithms for low-rank matrix completion.

A Corrected Inexact Proximal Augmented Lagrangian Method with a Relative Error Criterion for a Class of Group-quadratic Regularized Optimal Transport Problems – LEI YANG, SUN YAT SEN UNIVERSITY – Ling Liang, Hong T.M. Chu, Kim Chuan Toh

The optimal transport (OT) problem and its related problems have attracted significant attention and have been extensively studied in various applications. In this paper, we focus on a class of group-quadratic regularized OT problems which aim to find solutions with specialized structures that are advantageous in practical scenarios. To solve this class of problems, we propose a corrected inexact proximal augmented Lagrangian method (ciPALM), with the subproblems being solved by the semi-smooth Newton (Ssn) method. We establish that the proposed method exhibits appealing convergence properties under mild conditions. Moreover, our ciPALM distinguishes itself from the recently developed semismooth Newton-based inexact proximal augmented Lagrangian (Snipal) method for linear programming. Specifically, Snipal uses an absolute error criterion for the approximate minimization of the subproblem for which a summable sequence of tolerance parameters needs to be pre-specified for practical implementations. In contrast, our ciPALM adopts a relative error criterion with a single tolerance parameter, which would be more friendly to tune from computational and implementation perspectives. These favorable properties position our ciPALM as a promising candidate for tackling large-scale problems. Various numerical studies validate the effectiveness of employing a relative error criterion for the inexact proximal augmented Lagrangian method, and also demonstrate that our ciPALM is competitive for solving large-scale group-quadratic regularized OT problems.

An iteratively reweighted second-order method for nonconvex regularization – HAO WANG, SHANGHAITECH UNIVERSITY –

This paper considers a class of nonconvex sparsity-promoting regularization problems with a twice continuously differentiable loss function. We present a second-order algorithm to solve this class of nonconvex and nonsmooth problems. Most existing algorithms are first-order methods and only a hybrid of proximal gradient method and subspace regularized Newton method is proposed for lp regularization until recently. Our

new method is also a hybrid method with main features include: (i) our method is based on the iteratively reweighted method with the regularization term being iteratively approximated by a weighted ℓ_1 regularization term, so that it can be applied to various nonconvex regularization problems. (ii) Our method alternatively solves the weighted ℓ_1 approximation by a soft-thresholding step and the subspace approximate Newton step. (iii) The iterates generated by our algorithm have unchanged sign values and the nonzero components are bounded away from 0 for sufficiently large iterations and the algorithm eventually reverts to a perturbed Newton method. (iv) We prove global convergence and local quadratic convergence rate under loose assumptions for our method and demonstrate its efficiency on a large set of model prediction problems.

1c - TA112 - 511A - **Advances in Variational Analysis and Nonsmooth Optimization I**
 Chair: JOHANNES ROYSET 

Variational theory and algorithms for a class of asymptotically approachable nonconvex problems – YING CUI, UNIVERSITY OF CALIFORNIA BERKELEY –

In this talk, motivated by the inverse optimal value based nonlinear optimization problems, we introduce a class of composite nonconvex functions, where the outer function is the sum of univariate extended-real-valued convex functions and the inner function is the limit of difference-of-convex functions. A notable feature of this class is that the inner function can be merely lower semicontinuous instead of continuously differentiable. It covers a range of important yet challenging applications, including the composite value functions of nonlinear programs and the value at risk constraints for continuous random vectors. We propose an asymptotic decomposition of the composite function that guarantees epi-convergence to the original function, leading to necessary optimality conditions for the corresponding minimization problem. The proposed decomposition also enables us to design a numerical algorithm to solve the composite minimization problem such that any accumulation point of the generated sequence, if exists, satisfies the newly introduced optimality conditions. Our derived results expand on the study of so-called amenable functions introduced by Poliquin and Rockafellar in 1992, which are compositions of convex functions with smooth maps, and the prox-linear methods for their minimization. The talk is based on the joint work with Hanyang Li at the UC Berkeley.

Approximations of Rockafellians, Lagrangians, and Dual Functions – JOHANNES ROYSET, UNIVERSITY OF SOUTHERN CALIFORNIA –

Optimization problems are prone to instability, with seemingly minor approximations causing disproportional changes in minimizers, minimum values, and even level-sets. Fortunately, every optimization problem can be associated with many substitute problems potentially having more desirable properties. While focusing on the nonconvex case, we develop sufficient conditions under which Rockafellian relaxations, Lagrangian relaxations, and dual problems are well-behaved, in a precise and quantifiable sense, when subject to approximations. Examples from composite and stochastic optimization illustrate the results.

Heaviside Composite Optimization: Continuous meets Discrete – JONG SHI PANG, UNIVERSITY OF SOUTHERN CALIFORNIA –

A Heaviside function is an indicator function of an semi-infinite interval. A Heaviside composite function is a Heaviside function composed with a multivariate function that may be non-smooth. This talk presents a touch of this novel class of discontinuous optimization problems that borders on continuous and discrete optimization. Among the rich applications of such problems, we highlight one pertaining to classification and treatment with constraints solvable by an elementary yet novel progressive integer programming (PIP) method that can be applied broadly to many other problems, including indefinite quadratic programs and linear programs with linear complementarity constraints. Computational results demonstrate the excellent performance of the PIP idea and its superiority over a full integer programming approach.

Computing Wasserstein Barycenter via operator splitting: the method of averaged marginals – DANIEL MIMOUNI, MINES PARIS – Wellington De Oliveira, Paul Malisani, Jiamin Zhu

The Wasserstein barycenter (WB) is an important tool for summarizing sets of probability measures. It finds applications in applied probability, clustering, image processing, etc. When the measures' supports are finite, computing a (balanced) WB can be done by solving a linear optimization problem whose dimensions generally exceed standard solvers' capabilities. In the more general setting where measures have different total masses, we propose a convex nonsmooth optimization formulation for the so-called unbalanced WB problem. Due to their colossal dimensions, we introduce a decomposition scheme based on the Douglas-Rachford splitting method that can be applied to both balanced and unbalanced WB problem variants. Our algorithm, which has the interesting interpretation of being built upon averaging marginals, operates a series of simple (and exact) projections that can be parallelized and even randomized, making it suitable for large-scale datasets. Numerical comparisons against state-of-the-art methods on several data sets from the literature illustrate the method's performance.

1c - TB63 - 510B - Nonsmooth Stochastic Optimization and Applications Chair: HAILIN SUN

High-dimensional Limit of Stochastic Optimization Methods with Structured Data – INBAR SEROUSSI, TEL-AVIV UNIVERSITY –

Stochastic optimization methods are fundamental in large-scale machine learning. However, understanding why these methods perform so well remains a major challenge. In this talk, I will present a theory for stochastic gradient descent (SGD) and its adaptive variant in high dimensions when the number of samples and problem dimensions is large. I will show that the dynamics of SGD applied to generalized linear models and multi-index problems with data possessing a general covariance matrix become deterministic in the large sample and dimensional limit. In particular, the limiting dynamics are governed by a set of low-dimensional ordinary differential equations (ODEs). This setup encompasses many optimization problems, including linear regression, logistic regression, and two-layer neural

networks. In addition, it unveils the implicit bias inherent in SGD. For each of these problems, the deterministic equivalent of SGD enables us to derive a close approximation of the generalization error (with explicit and vanishing error bounds). Furthermore, we establish explicit conditions on the step size, ensuring the convergence and stability of SGD and a wide class of adaptive stochastic method.

Distributionally Robust Mean-Lower Partial Moment (LPM) Model and its Application in Portfolio Selection – JIE JIANG, CHONGQING UNIVERSITY –

The mean-risk model is a popular modeling paradigm, which quantifies the problem in a lucid form of only two criteria: the mean and the risk. Of particular interest of this paper, we consider the case that the risk is measured by a general lower partial moment (LPM) risk measure. When the underlying distribution of the random parameter is unknown, the distributionally robust approach is employed to handle the mean-LPM model by hedging against a range of possible distributions, which is the so-called ambiguity set. First of all, we discuss the attainability of the distributionally robust mean-LPM model regarding to both the decision variable and the worst probability distribution. After that, in the data-driven process, we investigate how the data size affects the parameters in the ambiguity set as well as the constraint endogenously when the ambiguity set is constructed by the first- and second-order moment information and the Wasserstein ball, respectively. This is not only curious in theory but also crucial in practical implementation. Finally, by defining some suitable optima, we give numerical procedures for computing these suitable optima. Some numerical results are reported by using real data in finance to validate the effectiveness and the efficiency of our method.

Data-driven Distributionally Robust Multiproduct Pricing Problems under Pure Characteristics Demand – HAILIN SUN, NANJING NORMAL UNIVERSITY –

This paper considers a multiproduct pricing problem under pure characteristics demand models when the probability distribution of the random parameter in the problem is uncertain. We formulate this problem as a distributionally robust optimization (DRO) problem based on a constructive approach to estimating pure characteristics demand models with pricing by Pang, Su and Lee. In this model, the consumers' purchase decision is to maximize their utility. We show that the DRO problem is well-defined, and the objective function is upper semicontinuous by using an equivalent hierarchical form. We also use the data-driven approach to analyze the DRO problem when the ambiguity set, i.e., a set of probability distributions that contains some exact information of the underlying probability distribution, is given by a general moment-based case. We give convergence results as the data size tends to infinity and analyze the quantitative statistical robustness in view of the possible contamination of driven data. Furthermore, we use the Lagrange duality to reformulate the DRO problem as a mathematical program with complementarity constraints, and give a numerical procedure for finding a global solution of the DRO problem under certain specific settings. Finally, we report numerical results that validate the effectiveness and scalability of our approach for the distributionally robust multiproduct pricing problem.

1c - TB207 - 511A - **Large scale continuous optimization and first-order methods and related topics** Chair: HAIHAO LU 

Level-set geometry and the complexity of restarted PDHG for conic LP – ZIKAI XIONG, MIT OPERATIONS RESEARCH CENTER – Robert Freund

The restarted primal-dual hybrid gradient method (rPDHG) – with heuristic enhancements and GPU implementation – has been very successful in solving huge-scale LP problems. Here we analyze the theoretical and practical performance of rPDHG for general (convex) conic optimization. We show a relationship between the geometry of the primal-dual sublevel sets W_{eps} and the convergence rate of rPDHG. In particular, we show that the convergence rate of rPDHG is faster when there is a primal-dual sublevel set W_{eps} for which (i) W_{eps} is close (in Hausdorff distance) to the optimal solution set, and (ii) the ratio of the diameter to the “conic radius” of W_{eps} is small. Based on these geometric condition numbers, we will present our most recent results on the complexity analysis and practical enhancements of rPDHG, such as explaining the easiness and hardness of certain important applications and improving the practical performance of rPDHG using scaling transformations.

A Practical and Optimal First-Order Method for Large-Scale Convex Quadratic Programming – HAIHAO LU, THE UNIVERSITY OF CHICAGO –

Convex quadratic programming (QP) is an important class of optimization problem with wide applications in practice. The classic QP solvers are based on either simplex or barrier method, both of which suffer from the scalability issue because their computational bottleneck is solving linear equations. In this paper, we design and analyze a first-order method called the restarted accelerated primal-dual hybrid gradient method for QP, whose computational bottleneck is matrix-vector multiplication. We show that the proposed algorithm has a linear convergence rate when solving generic QP, and the obtained linear rate is optimal among a wide class of primal-dual methods. Furthermore, we connect the linear rate with a sharpness constant of the KKT system of QP, which is a standard quantity to measure the hardness of a continuous optimization problem. Numerical experiments on a standard QP benchmark set showcase the advantage of the proposed algorithm compared to its first-order counterparts.

A feasible method for solving general convex low-rank SDP problems – TIANYUN TANG, NATIONAL UNIVERSITY OF SINGAPORE – Kim Chuan Toh

Riemannian optimization method is efficient for solving special SDP problems with block-diagonal constraints such as max-cut and orthogonal synchronization. However, for a general SDP problem, after low-rank decomposition, there might be inequality constraints and its feasible set may contain singular points where the LICQ doesn't hold, both of which imply that its feasible set is not a manifold. In this talk, we are going to discuss how to deal with these issues so that we can apply Riemannian optimization method to solve a general SDP problem

with provable convergence. Apart from this, we will also discuss how to make use of the sparsity of the constraint matrices to reduce the computational cost in computing the projection and retraction mapping. Numerical experiments will be displayed to compare our algorithm with some state of the art SDP solvers.

Practical Large-Scale Linear Programming using Primal-Dual Hybrid Gradient – DAVID AP-LEGATE, GOOGLE RESEARCH – Mateo Diaz, Oliver Hinder, Haihao Lu, Miles Lubin, Brendan O'donoghue, Warren Schudy

We present PDLP, a practical first-order method for linear programming (LP) that can solve to the high levels of accuracy that are expected in traditional LP applications. In addition, it can scale to very large problems because its core operation is matrix-vector multiplications. PDLP is derived by applying the primal-dual hybrid gradient (PDHG) method, popularized by Chambolle and Pock (2011), to a saddle-point formulation of LP. PDLP enhances PDHG for LP by combining several new techniques with older tricks from the literature; the enhancements include diagonal preconditioning, presolving, adaptive step sizes, and adaptive restarting. PDLP improves the state of the art for first-order methods applied to LP.

1c - WA285 - 510B - **Efficient algorithms for large-scale sparse optimization problems and applications** Chair: DEFENG SUN 

Convex facial reduction algorithm and strong extended dual – YING LIN, DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY – Tianxiang Liu, Bruno Lourenço

The facial reduction algorithm plays an important role in regularizing any conic convex program and establishing the extended duals that enjoy strong duality without requiring any constraint qualifications. When the cone is nice, the extended dual has a relatively simpler formula. In this work, we generalize both the facial reduction algorithm and the concept of niceness from closed convex cones to closed convex sets. Different from the classical facial reduction algorithm that only considers a linear objective function and an intersection of a closed convex cone and an affine subspace, our convex facial reduction algorithm has the flexibility to handle a nonlinear objective function and an intersection of two closed convex sets, and so is able to deal with problems that are inefficiently SDP representable. Based on the convex facial reduction algorithm we establish a class of extended duals that covers the extended duals in conic cases. Similarly, when both the two sets are nice, the formula of the extended dual is relatively simpler.

An Accelerated Preconditioned ADMM for Solving the Optimal Transport Problem – GUOJUN ZHANG, DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY –

In this paper, we present an accelerated preconditioned ADMM (pADMM) to efficiently solve large-scale optimal transport (OT) problems. Specifically, to address the challenges posed by solving large-scale linear equations in ADMM sub-problems, we propose a pADMM with a specialized proximal term that enables fast solutions to the subproblems. Furthermore, to fur-

ther improve the efficiency of pADMM, we propose an accelerated pADMM, which enjoys an $o(1/k)$ non-ergodic iteration complexity with respect to the Karush–Kuhn–Tucker (KKT) residual. Moreover, we implement the proposed accelerated pADMM on GPU. Numerical results demonstrate the outstanding performance of the accelerated pADMM algorithm in solving large-scale OT problems.

Collective Certified Robustness against Graph Injection Attacks – KAIHUANG CHEN, DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY – Lai Yuni, Yancheng Yuan, Zhou Kai, Bailin Pan

We investigate certified robustness for GNNs under graph injection attacks. Existing research only provides sample-wise certificates by verifying each node independently, leading to very limited certifying performance. In this paper, we present the first collective certificate, which certifies a set of target nodes simultaneously. To achieve it, we formulate the problem as a binary integer quadratic constrained linear programming (BQCLP). We further develop a customized linearization technique that allows us to relax the BQCLP into linear programming (LP) that can be efficiently solved. Through comprehensive experiments, we demonstrate that our collective certification scheme significantly improves certification performance with minimal computational overhead. For instance, by solving the LP within 1 minute on the Citeseer dataset, we achieve a significant increase in the certified ratio from 0.0 to 81.2 when the injected node number is 5 of the graph size. Our step marks a crucial step towards making provable defense more practical.

An efficient sieving based secant method for sparse optimization problems with least-squares constraints – DEFENG SUN, THE HONG KONG POLYTECHNIC UNIVERSITY – Qian Li, Yancheng Yuan

We propose an efficient sieving based secant method to address the computational challenges of solving sparse optimization problems with least-squares constraints. A level-set method has been introduced in [X. Li, D.F. Sun, and K.-C. Toh, SIAM J. Optim., 28 (2018), pp. 1842–1866] that solves these problems by using the bisection method to find a root of a univariate nonsmooth equation $\phi(\lambda) = \rho$ for some $\rho > 0$, where $\phi(\cdot)$ is the value function computed by a solution of the corresponding regularized least-squares optimization problem. When the objective function in the constrained problem is a polyhedral gauge function, we prove that (i) for any positive integer k , $\phi(\cdot)$ is piecewise C^k in an open interval containing the solution λ^* to the equation $\phi(\lambda) = \rho$; (ii) the Clarke Jacobian of $\phi(\cdot)$ is always positive. These results allow us to establish the essential ingredients of the fast convergence rates of the secant method. Moreover, an adaptive sieving technique is incorporated into the secant method to effectively reduce the dimension of the level-set subproblems for computing the value of $\phi(\cdot)$. The high efficiency of the proposed algorithm is demonstrated by extensive numerical results.

1c - WA333 - 511A - Recent Advances in Regularized and Composite Optimization Chair:

ALEXANDER HSU

Epi-convergent approximations of discontinu-

ous generalized eigenvalue functions with application to topology optimization – AKATSUKI NISHIOKA, THE UNIVERSITY OF TOKYO – Yoshihiro Kanno

Optimization of the maximum generalized eigenvalue of symmetric-matrix-valued functions is an important class of structural topology optimization, which is related to robustness, vibration, and buckling. The maximum generalized eigenvalue can be ill-defined (unbounded or discontinuous) where the matrices become singular (where a topological change of the structural design occurs). In our presentation, we redefine the maximum generalized eigenvalue as an extended real-valued function. Then, we propose a continuous approximation that epi-converges to the original extended real-valued function. Epi-convergence is a useful tool in variational analysis, and is used to show that solutions of approximated optimization problems converge, in some sense, to solutions of the original optimization problems. We conduct simple numerical experiments on truss topology optimization of the minimum eigenfrequency, although most of the theoretical results also hold for other types of topology optimization problems.

A Proximal Modified Quasi-Newton Method for Nonsmooth Regularized Optimization – MOHAMED LAGHDAF HABIBOULLAH, POLYTECHNIQUE MONTRÉAL – Youssef Diouane, Dominique Orban

We develop R2N, a modified quasi-Newton method for minimizing the sum of a continuously differentiable function f and lower semi-continuous prox-bounded h subject to bound constraints. Both f and h may be nonconvex. At each iteration, our method computes a step by minimizing the sum of a convex quadratic model of f , a model of h , and an adaptive quadratic regularization term. A step may be computed using the R2 method (Aravkin, Baraldi and Orban, 2020), or the variant R2DH, in which the Hessian model of f is diagonal, when h is separable. We establish convergence of a first-order stationarity measure to zero for both bounded and unbounded Hessian approximations. Under a Lipschitz assumption, we derive a worst-case evaluation complexity bound that matches the best-known bound for bounded Hessian approximations, and that deteriorates with the growth of the approximations when the latter are unbounded, as in (Leconte and Orban, 2023). We describe our implementation in Julia and report numerical experiences on inverse problems. Additionally, we discuss our findings on a minimum-rank matrix completion problem.

Feature Selection for Linear Fixed Effects Models – JAMES BURKE, UNIVERSITY OF WASHINGTON, SEATTLE –

Linear mixed-effects (LME) models are used to analyze nested or combined data across a range of groups or clusters. These models use covariates to separate the total population variability (the fixed effects) from the group variability (the random effects). LMEs borrow strength across groups to estimate key statistics in cases where the data within groups may be sparse or highly variable, and play a fundamental role in population health sciences, meta-analysis, life sciences, and in many others domains. In this talk we formally introduce a mathematical description of the LME model and its feature selection variant. A naive proximal gradient descent (PGD) algorithm for its solution is described and its deficiencies are explained. A novel solution strategy is proposed that is based on a relax-

ation strategy that decouples the smooth from the nonsmooth components of the maximum likelihood objective. An optimal value function is obtained by partially optimizing the smooth component of the decoupled problem. We show that the resulting optimal value function has a locally Lipschitz gradient and so a PGD algorithm can be applied to a feature selecting regularization of the optimal value function.

An efficient continuation algorithm for regularized optimization through the Moreau envelope – ALEXANDER HSU, UNIVERSITY OF WASHINGTON – Dominique Orban, Robert J Baraldi, Aleksander Aravkin

We develop a fast algorithm for solving sequences of regularized optimization problems by applying continuation to the Moreau envelope of the nonsmooth portion. We additionally show that our objective can be derived by replacing any of three terms with their Moreau envelope, and discuss efficient first order algorithms and linear algebra for solving the associated subproblems.

1c - WB331 - 511A - **Advances in First Order Methods** Chair: MARC TEBoulLE 

A Bestiary of Counterexamples in Smooth Convex Optimization – JEROME BOLTE, TOULOUSE SCHOOL OF ECONOMICS –

Counterexamples to some longstanding optimization problems in the smooth convex coercive setting are provided. Block-coordinate, steepest descent with exact search, or Bregman descent methods do not generally converge. Other failures of various desirable features are established: directional convergence of Cauchy's gradient curves, convergence of Newton's flow, finite length of Tikhonov path, convergence of central paths, or smooth Kurdyka-Lojasiewicz inequality. All examples are planar. These examples rely on a new convex interpolation result (and some variants): given a decreasing sequence of positively curved k differentiable smooth convex compact sets in the plane, we provide a k -times differentiable convex function having these sets as level sets. The integer k is arbitrarily large, and in addition, the Hessian is positive definite out of the solution set.

First Order Methods for Convex Bi-Level Optimization Problems – SHOHAM SABACH, TECHNION ISRAEL INSTITUTE OF ISRAEL –

Recently, bilevel optimization problems have received growing attention within the optimization community, because of their numerous important applications in machine learning and data science. We will discuss some of the recent developments in bilevel optimization problems in terms of algorithms and their theoretical convergence/complexity analysis, as well as some of the attendant difficulties and open challenges.

Survey Descent and Its Initialization – X.Y. HAN, CHICAGO BOOTH – Adrian Lewis

For strongly convex objectives that are smooth, the classical theory of gradient descent ensures linear convergence relative to the number of gradient evaluations. An analogous nonsmooth theory is challenging. Even when the objective is smooth at every iterate, the corresponding local models are unstable and the number of cutting planes invoked by traditional remedies is difficult to bound, leading to convergences guarantees that are sublinear relative to the cumulative number of gradient

evaluations. We instead propose a multipoint generalization of the gradient descent iteration for local optimization call Survey Descent. In this method, one first leverages a one-time initialization procedure to gather a "survey" of points. Then, during each iterate of the method, the survey points are updated in parallel using a simple, four-line procedure inspired by gradient descent. The survey points will then collectively move towards the objective minimizer. We prove linear convergence when the objective is itself max-of-smooth, and experiments suggest a more general phenomenon. While designed with general objectives in mind, we are motivated by a "max-of-smooth" model that captures the subdifferential dimension at optimality. I will review the Survey Descent method, what constitutes a good survey initialization, and discuss a potential heuristic for this initialization.

A Unified Framework for Convergence Analysis of Proximal Gradient Methods in Nonsmooth Nonconvex Minimax Problems – MARC TEBoulLE, TEL AVIV UNIVERSITY –

Nonconvex minimax problems are prevalent in modern applications. We focus on nonsmooth nonconvex minimax, thus departing from the more common weakly convex/concave and smooth models assumed in the recent literature. We present a unified framework for the convergence analysis of proximal gradient methods. In contrast to the current literature which addresses the complexity of obtaining only –nearly– approximate stationary solutions, here we derive pointwise global convergence and refined complexity results. Furthermore, our approach allows to expand the scope of minimax problems that can be addressed through Non Euclidean proximal gradient schemes. Based on joint work with E. Cohen.

1c - WB906 - 510B - **Nonsmooth Optimization**
1 Chair: ISABELLA KÄMING 

Commutation principles for nonsmooth variational problems on Euclidean Jordan algebras – JUYOUNG JEONG, CHANGWON NATIONAL UNIVERSITY – David Sossa

The commutation principle proved by Ramirez, Seeger, and Sossa (SIAM J Optim 23:687–694, 2013) in the setting of Euclidean Jordan algebras says that for a Frechet differentiable function Θ and a spectral function F , any local minimizer or maximizer a of $\Theta + F$ over a spectral set E operator commutes with the gradient of Θ at a . In this paper, we improve this commutation principle by allowing Θ to be nonsmooth with mild regularity assumptions over it. For example, for the case of local minimizer, we show that an operator commutes with some element of the limiting (Mordukhovich) subdifferential of Θ at a provided that Θ is subdifferentially regular at a satisfying a qualification condition. For the case of local maximizer, we prove that an operator commutes with each element of the (Fenchel) subdifferential of Θ at a whenever this subdifferential is nonempty. As an application, we characterize the local optimizers of shifted strictly convex spectral functions and norms over automorphism invariant sets.

An efficient active-set method with applications to sparse approximations and risk minimization – POUKAKIOTIS SPYRIDON, UNIVERSITY OF DUNDEE – Jacek Gondzio, Dionysis Kalogerias
In this talk we present an efficient active-set method for the so-

lution of convex quadratic programming problems with general piecewise-linear terms in the objective, with applications to sparse approximations and risk-minimization. The algorithm is derived by combining a proximal method of multipliers (PMM) with a standard semismooth Newton method (SSN), and is shown to be globally convergent under minimal assumptions. Further local linear (and potentially superlinear) convergence is shown under standard additional conditions. The major computational bottleneck of the proposed approach arises from the solution of the associated SSN linear systems. These are solved using a Krylov-subspace method, accelerated by certain novel general-purpose preconditioners which are shown to be optimal with respect to the proximal penalty parameters. The preconditioners are easy to store and invert, since they exploit the structure of the nonsmooth terms appearing in the problem's objective to significantly reduce their memory requirements. We showcase the efficiency, robustness, and scalability of the proposed solver on a variety of problems arising in risk-averse portfolio selection, L^1 -regularized partial differential equation constrained optimization, quantile regression, and binary classification via linear support vector machines. We provide computational evidence to demonstrate, on real-world datasets, the ability of the solver to efficiently and competitively handle a diverse set of medium- and large-scale instances.

Douglas-Rachford is the best projection method – HONGZHI LIAO, THE UNIVERSITY OF NEW SOUTH WALES – Minh Dao, Mareike Dressler, Vera Roshchina

We prove that the Douglas-Rachford method applied to two closed convex cones in the Euclidean plane converges in finitely many steps if and only if the set of fixed points of the Douglas-Rachford operator is nontrivial. We analyze this special case using circle dynamics.

We also construct explicit examples for a broad family of projection methods for which the set of fixed points of the relevant projection method operator is nontrivial, but the convergence is not finite. This three-parametric family is well-known in the projection method literature and includes both the Douglas-Rachford method and the classic method of alternating projections.

Even though our setting is fairly elementary, this work contributes in a new way to the body of theoretical research justifying the superior performance of the Douglas-Rachford method compared to other techniques. Moreover, our result leads to a neat sufficient condition for finite convergence of the Douglas-Rachford method in the locally polyhedral case on the plane, unifying and expanding several special cases available in the literature.

A new qualification condition for Lipschitzian optimization problems – ISABELLA KÄMING, TECHNISCHE UNIVERSITÄT DRESDEN – Andreas Fischer, Alain Zemkoho

We consider an inequality constrained optimization problem defined by possibly nonsmooth but Lipschitz continuous functions. Often, standard constraint qualifications may fail at local minimizers of the problem. It is known that corresponding approximate KKT conditions still hold in such situations without requiring a constraint qualification to be fulfilled. Therefore, it is of interest to find weak conditions which guarantee that a point satisfying the approximate KKT conditions fulfills

the usual KKT conditions. In this talk, we present a condition, which is not a constraint qualification, that possesses the aforementioned property and can be satisfied by a wide range of Lipschitzian optimization problems. A comparison with existing constraint qualifications reveals that the new condition is implied by quasinormality in local minimizers. Relations to further constraint qualifications will be discussed as well.

1c - WC45 - 510B - **Large-scale Iterative Optimization** Chair: MATEO DIAZ

Improving the Geometry of (Conic) Linear Optimization Problems for the Primal-Dual Hybrid Gradient Method (PDHG) – ROBERT FREUND, MIT SLOAN SCHOOL OF MANAGEMENT – Zikai Xiong

The primal-dual hybrid gradient method (PDHG) (with restarts and heuristic enhancements) has shown significant success in solving huge-scale LP problems, and our recent research has shown that this success depends on certain geometric condition measures of the LP related to error ratios and sharpness. However, depending on the problem instance, these condition measures may take on extreme values and lead to poor performance both in theory and in practice. In such instances the LP problem is perhaps better understood as a special case of a more general conic (nonlinear) optimization problem. This begs the following two questions: (1) can the geometric condition measures for LP be improved by applying suitable transformations of the given LP instance?, and (2) moving beyond LP to conic optimization, to what extent can we generalize the methods, analysis, and computational performance (theory and practice) from LP to conic optimization? Regarding these questions, we show how column rescaling (for LP) and barrier-based rescaling (for conic problems) can improve the condition measures in theory and lead to improved performance of PDHG in theory and in practice. Our theoretical development leads to guidelines for practical implementation of re-scaling based on ideas from analytic centers. Also, our experiments on LP relaxations of the MIPLIB 2017 dataset demonstrate the impact of rescaling on the actual performance of PDHG.

The role of sparsity in differentially-private optimization – CRISTÓBAL GUZMÁN, PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE –

With the increasing use of personally-sensitive data in machine learning applications, privacy has become a central concern. In this context, differential privacy (DP) offers a rigorous and quantifiable control of the privacy risk of a machine learning model. One of the main problems of interest in differential privacy is stochastic optimization, where we are interested in computing a model that approximately minimizes the empirical or population excess risk, while satisfying differential privacy with respect to the data used for training the model.

In this talk, we will present various settings where one can obtain nearly dimension independent accuracy rates in differentially private (stochastic) optimization and saddle-point problems. We start with results involving stochastic optimization under polyhedral constraints, popular in sparsity-oriented machine learning, which are based on the variance-reduced stochastic Frank-Wolfe method. Next, we move to stochastic saddle-point problems, where we study the use of stochas-

tic mirror-descent methods and vertex sampling; these results are applied to problems including DP and group-fair machine learning, and DP synthetic data generation. Finally, I will present results on the "dual" counterpart of the above problems: stochastic optimization with sparse gradients, a setting of high relevance in large embedding models. Here, we provide new matching upper and lower bounds both in the convex and nonconvex settings.

Density Separation with Tensor Factorization – NAOMI GRAHAM, UNIVERSITY OF BRITISH COLUMBIA – Michael Friedlander

A nonnegative tensor factorization model is proposed for separating mixtures of probably densities. The approach uses kernel density estimation to transform raw samples into compressed and discretized densities. The factorization is computed using a specialized block-coordinate descent algorithm that enforces the required simplex constraints and is guaranteed to converge to Nash equilibria. Numerical experiments on real geological and spatial transcriptomics data, based on our open source Julia implementation, illustrate the model's effectiveness at separating the density mixtures.

1c - ThA30 - 510B - **Nonsmooth nonconvex optimization with geometric structure** Chair: XIN LIU ↩

Away-step Frank-Wolfe algorithm for multi-objective optimization on polytopes – DOUGLAS SOARES GONCALVES, UNIVERSIDADE FEDERAL DE SANTA CATARINA –

An away-step Frank-Wolfe algorithm designed for solving multi-objective optimization problems over polytopes is discussed. We prove that each limit point of the sequence generated by the algorithm is a weak Pareto optimal solution and, under additional conditions, we show linear convergence of the whole sequence to a Pareto optimal solution. At each iteration a non-smooth subproblem needs to be solved. Some solution strategies for these subproblems are proposed for specific polytopes.

A Near-optimal Method for Quadratic Constrained Composite Non-convex Non-smooth Problems – WEI LIU, RENSSELAER POLYTECHNIC INSTITUTE –

This paper studies first-order methods (FOMs) for solving *composite non-convex non-smooth* optimization with linear constraints. Recently, the lower complexity bounds of FOMs on finding a (near) ϵ -stationary point of the considered problem have been established. However, optimization algorithms that can achieve this lower bound have never been developed. This paper presents a primal-dual recovery method for solving the targeted problem. The oracle complexity of the proposed method to find a (near) ϵ -stationary point of the considered problem matches the lower bounds up to a logarithmic factor. Therefore, our proposed method is almost non-improvable. We illustrate the advantages of our proposed algorithm over existing methods in the numerical experiments.

Global and local convergence guarantees for bounded-rank optimization with a desingularization geometry – QUENTIN REBJOCK, EPFL – Nicolas Boumal

Low-rank optimization is difficult because the set of bounded-rank matrices is a non-smooth and non-convex algebraic variety. Existing paradigms include fixed-rank optimization and the LR parameterization. They lack either fast convergence guarantees to lesser rank points, or the ability to accumulate only at critical points. We seek optimization algorithms that overcome these flaws.

Khrulkov and Oseledets (2018) proposed to smoothly parameterize the low-rank variety via a desingularization. This allows to recast the optimization problem onto a smooth manifold. Building on their ideas, we propose to endow this desingularization with a metric to take advantage of existing Riemannian optimization algorithms. We study its geometry and show that this parameterization has multiple theoretical benefits. It notably preserves the compactness of sublevel sets and Polyak–Łojasiewicz-type conditions. We deduce global and local convergence guarantees for a range of algorithms, even at non-maximal rank points. Finally, we experimentally compare our algorithms to other existing techniques on matrix completion problems.

1c - ThA56 - 511A - **Optimization with Projections** Chair: XIAOQI YANG ↩

Projectional coderivatives with applications – XIAOQI YANG, THE HONG KONG POLYTECHNIC UNIVERSITY – Kaiwen Meng, Minghua Li, Wenfang Yao In this talk we introduce a projectional coderivative of set-valued mappings and present its calculations in some special cases. We apply this coderivative to obtain a complete characterization for a set-valued mapping to have the Lipschitz-property relative to a closed and convex set. For an extended real-valued function, we apply the obtained results to investigate its Lipschitz continuity relative to a closed and convex set and the Lipschitz-like property of a level-set mapping relative to a half line. We apply our results to study the Lipschitz-like property of the solution mapping of a parametric affine variational inequality problem.

Sorting Functions for Sparse Projection with Applications to Sparse Optimization – JINGLAI SHEN, UNIVERSITY OF MARYLAND BALTIMORE COUNTY –

Motivated by the symmetric sparse projection results developed by Beck, Eldar, and Hallmark, we introduce a general notion of sorting function for sparse projection without imposing full permutation symmetry on an underlying set. If a sorting function is defined by the monotone order of a real-valued univariate function, then it is called a simple sorting function. For example, when an underlying set is nonnegative and fully permutation symmetric, it is known that its simple sorting function can be defined by t . Fundamental properties of simple sorting functions on a possibly non-fully permutation symmetric set are studied. Specific results include closed form expressions of a simple sorting function, and necessary/sufficient conditions for an underlying set to attain such a simple sorting function. In particular, we derive necessary and/or sufficient conditions on a set under which $t, |t|, t_+$ (respectively t_-) defines a simple sorting function, and show that a set satisfying certain partially permutation symmetric property may attain a simple sorting function. These results broaden sorting based

sparse projection to a larger class of sets and lay a foundation for developing efficient sparse projection algorithms with various applications to sparse optimization.

A projection algorithm for nonlocal low-rank tensor models with orthogonal constraints –

XIAOXIA LIU, HONG KONG POLYTECHNIC UNIVERSITY – Shijie Yu, Jian Lu, Xiaojun Chen

Hyperspectral images (HSIs) are often contaminated by mixed noises such as Gaussian noise, dead lines, stripes and so on. In this talk, we will present an optimization model for HSI denoising using a tensor $l_{2,p}$ group sparsity measure and nonlocal low-rank regularization. In the proposed nonlocal low-rank tensor model, the low-rank constraints are imposed on the nonlocal similar tensors collected via block matching according to the spatial nonlocal self-similarity and spectral correlation of HSIs. The low-rank regularization term is formulated based on independent 3-D higher-order singular value decomposition with sparsity enhancement on its core tensor to prompt more low-rankness. The resulting denoising model is a nonconvex nonsmooth optimization problem with orthogonal constraints. We propose a proximal block coordinate descent algorithm to solve the model, in which we utilize the projection onto Stiefel manifold to solve the equivalent subproblems. We can show any accumulation point of the sequence generated by the proposed algorithm converges to a stationary point, which is defined using three equalities of substationarity, symmetry, and feasibility for orthogonal constraints. Experiments on HSI denoising and destriping demonstrate the great performance of our proposed method over state-of-the-art methods in terms of metrics such as mean peak signal-to-noise ratio as well as visual quality.

1c - ThB54 - 511A - **Projection-free first-order methods for convex optimization** Chair: JAVIER PENA 

Saddle-point Frank-Wolfe methods for Linear Programming – STEPHEN VAVASIS, UNIVERSITY OF WATERLOO – Matthew Hough

We propose two saddle-point Frank-Wolfe methods called FWLP and FWLP-P to solve general LP instances. Convergence of FWLP-P is proved, and a convergence rate of $O(k^{-1/2})$ is established. FWLP is simpler than FWLP-P and more efficient in practice, but no convergence proof is known for FWLP. The methods have an advantage over other first-order methods for LP in that they do not need the entire constraint matrix on each iteration.

The Generalized Multiplicative Gradient Method for a Class of Convex Symmetric Cone Optimization Problem – RENBO ZHAO, UNIVERSITY OF IOWA –

We consider a class of convex symmetric cone optimization problem, where the objective function is a convex log-homogeneous function that satisfies the "log-gradient convexity". Applications of this class of problems include inference of multivariate Hawkes processes, D-optimal design, quantum state tomography, and Nesterov's reformulated convex relaxation of Boolean quadratic problems. We develop a generalized multiplicative gradient method for this class of problems, and show that it converges with rate $O(1/k)$. We demonstrate that for all the aforementioned applications, the computational

complexity of our method is superior or comparable to other applicable first-order methods.

Adaptive Open-Loop Step-Sizes for Accelerated Frank-Wolfe Convergence – ELIAS WIRTH, TU BERLIN –

We extend acceleration results from fixed- ℓ open-loop step-sizes, $\ell \in \mathbb{N}$, characterized by $\eta_t = \frac{\ell}{t+\ell}$, to adaptive open-loop step-sizes $\eta_t = \frac{g(t)}{t+g(t)}$, with $g: \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ a nondecreasing function that fulfills several well-founded assumptions. This development inspires the adoption of the log-adaptive step-size, derived by setting $g(t) = 2 + \log(t+1)$, which achieves convergence rates comparable to or faster than those of any fixed- ℓ step-size, up to polylogarithmic factors. In the prevalent strong growth setting, the log-adaptive step-size outperforms any fixed- ℓ step-size and can admit arbitrarily fast sublinear convergence rates.

Affine-invariant convergence of Frank-Wolfe algorithms on polytopes – JAVIER PENA, CARNEGIE MELLON UNIVERSITY –

We provide a unified analysis of the following popular versions of the Frank-Wolfe method on polytopes: vanilla Frank-Wolfe, Frank-Wolfe with away steps, Frank-Wolfe with blended pairwise steps, and Frank-Wolfe with in-face directions. For each of these algorithms we establish convergence rates ranging from sublinear to linear. Our approach is entirely affine-invariant as it relies on some suitable error bound and curvature properties that depend solely on the objective function and the polytope but not on any affine-dependent objects such as norms. We highlight the interesting geometric interpretation of the error bound properties underlying our main developments.

1c - ThB907 - 510B - **Nonsmooth Optimization**
2 Chair: ZEHUI LIU 

An abstract convergence framework with application to proximal Heavy-ball line-search algorithms – SIMONE REBEGOLDI, UNIVERSITY OF MODENA AND REGGIO EMILIA – Bonettini Silvia, Peter Ochs, Marco Prato

In this talk, we analyze the convergence of an abstract descent algorithm to a stationary point of a proper and lower semicontinuous function. The proposed algorithm is defined through a specific set of properties that are common to several first-order methods for nonsmooth optimization. Such properties guarantee the convergence of the whole sequence of iterates to a stationary point, if the objective function satisfies the Kurdyka-Lojasiewicz property. Unlike similar prior works, we relax two of these properties, the so-called sufficient decrease condition and the relative error condition, allowing for the design of non-monotone algorithms with an inexact iteration update step that is implementable in practice. We propose three novel proximal Heavy-ball-type algorithms that are special instances of our abstract framework. The first algorithm - "i2Piano" - adjusts the steplength by backtracking until a local approximation of the Lipschitz constant is found. The two other algorithms - "iPila" and "Phila" - enforce a descent condition on a merit function rather than the objective function itself by performing a linesearch along the descent direction. We apply the proposed algorithms on a series of image deblurring problems, assessing their efficiency and superiority with respect to other competitors available in the literature, espe-

cially when the inertial parameter is selected by mimicking the Conjugate Gradient updating rules.

On the limit form of the Su-Boyd-Candès dynamic version of Nesterov’s accelerated gradient method when the viscous parameter becomes large – SAMIR ADLY, UNIVERSITY OF LIMOGES –

In a Hilbert setting, our study concerns the dynamical system introduced by Su-Boyd-Candès as a low resolution ODE of Nesterov’s accelerated gradient method and of the Ravine method. This inertial system $(AVD)_\alpha$ is driven by the gradient of the function f to be minimized, and is damped by a viscous friction whose coefficient is equal to α/t , and therefore tends to zero when t tends to infinity. Taking α sufficiently large plays a crucial role in the asymptotic convergence properties of the trajectories generated by this system. In particular, for a general convex function f , the condition $\alpha > 3$ guarantees the asymptotic convergence rate of the values $o(1/t^2)$, as well as the convergence of the trajectories towards optimal solutions. On the basis of temporal scaling techniques, we obtain, for α large, equivalent simple continuous dynamics which are autonomous and of the first order in time. Precisely, we show that a judicious time rescaling of $(AVD)_\alpha$ results in trajectories close to those of the continuous steepest descent associated with f . This explains the passage from $1/t$ to $1/t^2$ of the convergence rate of values between the steepest descent method and the Nesterov accelerated gradient method (as well as the Ravine accelerated gradient method). At the same time, new results are obtained concerning the complexity analysis of the Nesterov method and of the Ravine method. Some numerical experiments are provided to verify the theoretical results.

Inertial Accelerated Stochastic Mirror Descent for Large-Scale Generalized Tensor CP Decomposition – ZEHUI LIU, BEIHANG UNIVERSITY – Qingsong Wang, Chunfeng Cui, Yong Xia

The majority of classic tensor CP decomposition models are designed for squared loss, employing Euclidean distance as a local proximal term. However, the Euclidean distance is unsuitable for the generalized loss function applicable to various types of real-world data, such as integer and binary data. Consequently, algorithms developed under the squared loss are not easily adaptable to handle these generalized losses, partially due to the lack of the gradient Lipschitz continuity. This paper considers the generalized tensor CP decomposition. We use the Bregman distance as the proximal term and propose an inertial accelerated block randomized stochastic mirror descent algorithm (iTableSMD). Within a broader multi-block variance reduction and inertial acceleration framework, we demonstrate the sublinear convergence rate for the subsequential sequence produced by the iTableSMD algorithm. We further show that iTableSMD requires at most $O(\varepsilon^{-2})$ iterations in expectation to attain an ε -stationary point and establish the global convergence of the sequence. Numerical experiments on real datasets demonstrate that our proposed algorithm is efficient and achieves better performance than the existing state-of-the-art methods.

Global stability of first-order methods for coercive tame functions – LEXIAO LAI, THE UNIVERSITY OF HONG KONG – Cedric Josz

We consider first-order methods with constant step size for

minimizing locally Lipschitz coercive functions that are tame in an $\mathbf{o}$ -minimal structure on the real field. We prove that if the method is approximated by subgradient trajectories, then the iterates eventually remain in a neighborhood of a connected component of the set of critical points. Under suitable method-dependent regularity assumptions, this result applies to the subgradient method with momentum, the stochastic subgradient method with random reshuffling and momentum, and the random-permutations cyclic coordinate descent method.

In this talk, we provide global stability guarantees for first-order methods with constant step size for objective functions that are locally Lipschitz, coercive, and tame in an $\mathbf{o}$ -minimal structure on the real field. In order to do so, we show that the function values and the iterates of an iterative method eventually stabilize around some critical value and the set of critical points, respectively, given that the method is approximated by subgradient trajectories of the objective function. As it turns out, all of the aforementioned first-order methods are approximated by subgradient trajectories of locally Lipschitz functions under method-dependent regularity assumptions. Therefore, these methods fit into our framework, and their stability is guaranteed. To the best of our knowledge, these methods have not been studied before at such generality as in this paper. In particular, we do not require the objective function to be convex and we do not require it to be differentiable with a locally Lipschitz gradient.

1c - ThC38 - 510B - **Nonsmooth calculus and optimization in machine learning** Chair: JALAL FADILI 

Differentiating Nonsmooth Solutions to Parametric Monotone Inclusion Problems – ANTONIO SILVETI FALLS, CENTRALESUPÉLEC –

Understanding the differentiability and regularity of the solution to a monotone inclusion problem as a function of the problem data is an important question with consequences for convex optimization, machine learning, signal processing, and beyond. Past attempts have been made either under very restrictive assumptions that ensure the solution is continuously differentiable or using mathematical tools that are incompatible with automatic differentiation. In this talk, we discuss how to leverage path differentiability and a recent result on nonsmooth implicit differentiation calculus to give sufficient conditions ensuring that the solution to a monotone inclusion problem will be path differentiable. We also provide formulas for computing its generalized gradient that can be used in practice, notably for applications in deep learning. Our approach is fully compatible with automatic differentiation and the back-propagation algorithm and comes with assumptions which are easy to check. In fact, we show that semialgebraicity (or even just definability in some $\mathbf{o}$ -minimal structure) and strong monotonicity of the monotone inclusion problem are sufficient to guarantee the path differentiability of the solution. We illustrate the scope of our results by considering three fundamental composite problem settings that are often encountered in practice: strongly convex problems, dual solutions to convex minimization problems and primal-dual solutions to min-max problems.

The gradient’s limit of a definable family of functions is a conservative set-valued field –

SHOLOM SCHECHTMAN, TÉLÉCOM SUDPARIS –

It is well-known that the convergence of a family of smooth functions does not imply the convergence of its gradients. In this talk, we will show that if the family is definable in an o-minimal structure (for instance semialgebraic, subanalytic, or any composition of the previous with $\exp$, $\log$), then the gradient's limit is actually a conservative set-valued field in the sense introduced by Bolte and Pauwels. This result will be presented in a more general setting, where the functions in the original family might be merely locally Lipschitz continuous, vector-valued and the gradients are replaced by the Clarke Jacobians (or an arbitrary conservative mapping).

In particular, while a conservative set-valued field of a function is not necessarily equal to its gradient, it is equal to its gradient almost everywhere, implying the convergence of the gradients on a set of full measure. Moreover, the general geometric structure of a conservative field and its links with the usual subgradients are well-understood and will be presented in the talk.

Finally, we will discuss consequences of this result for convergence guarantees of smoothing methods and potential generalizations to the case where the functions are not locally Lipschitz continuous.

Riemannian Stochastic Approximation for Minimizing Tame Nonsmooth Objective Functions – VYACHESLAV KUNGURTSEV, CZECH TECHNICAL UNIVERSITY – Johannes Aspmann, Reza Roohi Seraji

In many learning applications, the parameters in a model are structurally constrained in a way that can be modeled as them lying on a Riemannian manifold. Riemannian optimization, wherein procedures to enforce an iterative minimizing sequence to be constrained to the manifold, is used to train such models. At the same time, tame geometry has become a significant topological description of nonsmooth functions that appear in the landscapes of training neural networks and other important models with structural compositions of continuous nonlinear functions with nonsmooth maps. In this paper, we study the properties of such stratifiable functions on a manifold and the behavior of retracted stochastic gradient descent, with diminishing stepsizes, for minimizing such functions.

Functions whose graph satisfies this structure, called tame, have been the subject of significant interest in recent years. This is in light of the fact that the optimization problem defining the training of the weights of a deep neural network is expressed as the minimization of a tame function. The function is a nested composition of component-wise maximums (rectified linear units) and nonlinear activation functions (softmax, hyperbolic tangent, etc.), resulting in a nonsmooth and non-convex objective landscape, however one that is also locally Lipschitz and with a well-defined set of points at which it is nondifferentiable with a hierarchical structure.

1c - FA25 - 510B - **Some recent advances in zeroth and first-order optimization methods**
Chair: TING KEI PONG

Structure Utilization for Reducing Dimensional Dependency for Nonsmooth Derivative-Free Optimization – CHING PEI LEE, INSTITUTE

OF STATISTICAL MATHEMATICS –

Deterministic methods for derivative-free optimization that use finite difference to approximately conduct first-order updates usually have an evaluation cost per iteration proportional to the problem dimension, making these methods prohibitively expensive for high-dimensional problem. In this work, we consider a regularized setting of minimizing the sum of a smooth function and a nonsmooth regularization term, where we only have access to the function value of the smooth term but assume full knowledge of the regularization. By identifying the structure at the point of convergence induced by the regularization using an inexact proximal gradient method constructed through finite difference, we propose an acceleration method whose per-iteration evaluation cost locally depends on the rank of the structure instead of the original problem dimension, and thus making the algorithm much more efficient without harming the convergence guarantees. We discuss how the structure can be found under a general error bound condition, and cover related convergence properties. We further consider a progressive identification strategy that gradually decreases the estimated rank of the structure through an approximate Riemannian proximal gradient method by confining to the manifold that represents the structure at the current point, and discuss how it could potentially improve the convergence speed. This is joint work with Geovani Nunes Grapiglia.

Adaptable proximal algorithms: improved stepsize bounds and weaker assumptions – ANDREAS THEMELIS, KYUSHU UNIVERSITY – Puya Latafat, Lorenzo Stella, Patrinos Panagiotis

The choice of stepsize parameters has a significant impact on the performance of gradient-based optimization methods. Linesearches are arguably the most successful established strategies, for they enact an online tuning of the stepsize that reflects the local landscape of the function, thus striking a significant advantage over constant regimes which are instead only concerned with worst-case scenarios. On the other hand, the extra numerical effort involved in the necessary trial-and-error inner loops can constitute a severe drawback, especially in those cases in which functions are expensive to evaluate.

An answer to this dilemma has been recently found in (self-)adaptive strategies, in which stepsize parameters are updated based on local information yet with no need of posterior validation. Building upon recent works on these linesearch-free adaptive first-order methods, this talk proposes a framework that both unifies and tightens existing results by providing larger stepsize selection policies and improved lower bounds. While retaining a robust and linesearch-free nature, the employment of time-varying algorithmic meta-parameters provides an extra layer of adaptivity that can further boost performance in practice. Primal, dual, and primal-dual algorithms are investigated, and their effectiveness showcased with numerical simulations. Finally, we comment on recent developments demonstrating their effectiveness even under mere local Hölder continuity of the gradient.

MGProx: A nonsmooth multigrid proximal gradient method with adaptive restriction for strongly convex optimization – ANDERSEN (MAN SHUN) ANG, UNIVERSITY OF SOUTHAMPTON – Hans De Sterck, Stephen Vavasis

We study the combination of proximal gradient descent with

multigrid for solving a class of possibly nonsmooth strongly convex optimization problems. We propose a multigrid proximal gradient method called MGProx, which accelerates the proximal gradient method by multigrid, based on utilizing hierarchical information of the optimization problem.

MGProx applies a newly introduced adaptive restriction operator to simplify the Minkowski sum of subdifferentials of the nondifferentiable objective function across different levels. We provide a theoretical characterization of MGProx. First we show that variables at all levels exhibit a fixed-point property at convergence. Next, we show that the coarse correction is a descent direction for the fine variable in the general nonsmooth case. Lastly, under some mild assumptions we provide the convergence rate for the algorithm, such as the classical sub-linear $O(1/k)$ rate and also the linear rate.

By treating the multigrid proximal gradient iteration as a black-box, we also proposed a fast MGProx with Nesterov's acceleration, together with the classical $O(1/k^2)$ rate.

In the numerical experiments, we show that MGProx has a significantly faster convergence speed than proximal gradient descent and proximal gradient descent with Nesterov's acceleration on nonsmooth convex optimization problems such as the Elastic Obstacle Problem, which the restriction operator is well known.

Lastly, if time permits, we also discuss the possible extension of MGProx to other first-order methods in primal-dual settings and splitting settings, and also the possible extension to second-order method such as non-smooth trust region methods.

Understanding Neural Network Binarization with Forward and Backward Proximal Quantizers – YAOLIANG YU, UNIVERSITY OF WATERLOO –

In neural network binarization, BinaryConnect (BC) and its variants are considered the standard. These methods apply the sign function in their forward pass and their respective gradients are backpropagated to update the weights. However, the derivative of the sign function is zero whenever defined, which consequently freezes training. Therefore, implementations of BC (e.g., BNN) usually replace the derivative of sign in the backward computation with identity or other approximate gradient alternatives. Although such practice works well empirically, it is largely a heuristic or “training trick.” We aim at shedding some light on these training tricks from the optimization perspective. Building from existing theory on ProxConnect (PC, a generalization of BC), we (1) equip PC with different forward-backward quantizers and obtain ProxConnect++ (PC++) that includes existing binarization techniques as special cases; (2) derive a principled way to synthesize forward-backward quantizers with automatic theoretical guarantees; (3) illustrate our theory by proposing an enhanced binarization algorithm BNN++; (4) conduct image classification experiments on CNNs and vision transformers, and empirically verify that BNN++ generally achieves competitive results on binarizing these models.

An asynchronous parallel bundle method – FRANK FISCHER, JOHANNES GUTENBERG UNIVERSITY MAINZ –

In this talk we present a fully asynchronous proximal bundle method for solving non-smooth, convex optimization problems. The algorithm can be used as a drop-in replacement for classic bundle methods, i.e., the function must be given by a first-order oracle for computing function values and subgradients. The algorithm allows for an arbitrary number of master problem processes computing new candidate points and oracle processes evaluating functions at those candidate points. These processes share information by communication with a single supervisor process that resembles the main loop of a classic bundle method. All processes run in parallel and no explicit synchronization step is required. Instead, the asynchronous and possibly outdated results of the oracle computations can be seen as an inexact function oracle. Hence, we show the convergence of our method under weak assumptions very similar to inexact and incremental bundle methods. In particular, we show how the algorithm learns important structural properties of the functions to control the inaccuracy induced by the asynchronicity automatically such that overall convergence can be guaranteed. Our tests on relaxations of some multi-commodity flow problems show the advantageousness of our approach if there are some slower processes.

Geometric characterizations of Lipschitz stability for convex optimization problems – NGHIA TRAN, OAKLAND UNIVERSITY –

In this paper, we mainly study tilt stability and Lipschitz stability of convex optimization problems. Our characterizations are geometric and fully computable in many important cases. As a result, we apply our theory to the group Lasso problem and the nuclear norm minimization problem and reveal that the Lipschitz stability of the solution mapping in these problems is automatic whenever the solution mapping is single-valued.

First-Order Methods for Constrained Optimization Problems Under Distributional Uncertainties: Empirical Evaluations and Theory – MEISAM RAZAVIYAYN, UNIVERSITY OF SOUTHERN CALIFORNIA –

Nonconvex models, like deep learning, have excelled in recent machine learning tasks. Practical training of such models involves meeting multiple requirements, including fairness, robustness, and accuracy-at-top. To address these requirements, training can be performed by solving a constrained optimization problem, with each requirement represented as a constraint. Moreover, distributional shifts from the training phase to the deployment (test) phase might occur. Practitioners commonly tackle this constrained optimization problem by solving the Lagrangian Relaxation and fine-tuning the Lagrange multipliers. Although easily implemented in training frameworks by altering the loss function, Lagrangian Relaxation for non-convex problems may fail to converge, even with close initialization to the optimal solution. In this work, we investigate existing alternative approaches for training large models under the presence of constraints (and possible distribution shifts). We specifically assess the performance of Exact Penalty (EP), Dual Ascent (DA), and the Method of Multipliers (MM). Like Lagrangian relaxation, these methods seamlessly integrate into existing frameworks, such as PyTorch. We investigate convergence to stationarity, stability around optimal solutions, and

hyperparameter tuning for these approaches both in theory and with empirical evaluations. Our extensive numerical experiments demonstrate that, across a diverse range of scenarios, these alternative approaches consistently outperform Lagrangian Relaxation which is commonly used by practitioners.

Proximal Algorithms for Gaussian variational inference via JKO in the Bures-Wasserstein Space – KRISHNAKUMAR BALASUBRAMANIAN, UNIVERSITY OF CALIFORNIA, DAVIS – Variational inference (VI) has emerged as a tractable alternative to computationally demanding Monte Carlo Markov Chain (MCMC) methods. VI seeks to approximate a target distribution by an element of a tractable family of distributions. Of key interest in statistics and machine learning is Gaussian VI, which approximates the target by minimizing the Kullback-Leibler (KL) divergence to over the space of Gaussians. In this work, we develop the (Stochastic) Forward-Backward Gaussian Variational Inference (FB-GVI) algorithm to solve Gaussian VI. Our approach exploits the composite structure of the KL divergence, which can be written as the sum of a smooth term (the potential) and a non-smooth term (the entropy) over the Bures-Wasserstein (BW) space of Gaussians endowed with the Wasserstein distance. FB-GVI hence leverages celebrated ideas from the literature on composite and non-smooth optimization. A key insight in this work is that the JKO operator for the entropy, when restricted to the BW space, admits a closed form, and hence leads to an implementable (stochastic) forward-backward (or proximal gradient) algorithm for Gaussian VI. In turn, it yields new computational guarantees for Gaussian VI under a variety of standard assumptions. Specifically, we obtain state-of-the-art convergence guarantees when is log-smooth and log-concave, as well as the first convergence guarantees to first-order stationary solutions when is only log-smooth.

1c - FB257 - 510B - **Complexity of Stationarity Concepts in Non-Smooth Optimization** Chair: MAN CHO SO

Testing Approximate Stationarity Concepts for Piecewise Affine Functions – LAI TIAN, THE CHINESE UNIVERSITY OF HONG KONG – Man Cho So

We study various aspects of the basic computational problem of detecting approximate stationary points for continuous piecewise affine (PA) functions, including computational complexity, regularity conditions, and robustness in implementation. Specifically, for a PA function, we show that testing first-order approximate stationarity concepts in terms of commonly used subdifferential constructions is computationally intractable unless $P=NP$. To facilitate computability, we establish the first simultaneously necessary and sufficient condition for the validity of an equality-type (Clarke) subdifferential sum rule for a certain representation of arbitrary PA functions. Our main tools are nonsmooth analysis and polytope theory. Moreover, to address an important implementation issue, we introduce the first oracle-polynomial-time algorithm to detect near-approximate stationary points for PA functions. We complement our results with extensions to other subdifferentials and applications to a series of structured piecewise smooth functions, including rho-margin-loss SVM, piecewise affine regression, and neural

networks with nonsmooth activation functions.

Strong Second-order Sufficient Condition of Decomposable Functions – WENQING OUYANG, THE CHINESE UNIVERSITY OF HONGKONG (SHENZHEN) –

Local superlinear convergence of the semismooth Newton method usually requires the uniform invertibility of the generalized Jacobian matrix, e.g. BD-regularity or CD-regularity. For several types of nonlinear programming and composite-type optimization problems – for which the generalized Jacobian of the stationary equation can be calculated explicitly – this is characterized by the strong second-order sufficient condition. However, general characterizations are still not well understood. In this paper, we propose a strong second-order sufficient condition (SSOSC) for composite problems whose nonsmooth part has a generalized conic-quadratic second subderivative. We then discuss the relationship between the SSOSC and another second order-type condition that involves the generalized Jacobians of the normal map. In particular, these two conditions are equivalent under certain structural assumptions on the generalized Jacobian matrix of the proximity operator. Next, we verify these structural assumptions for C^2 -strictly decomposable functions via analyzing their second-order variational properties under additional geometric assumptions on the support set of the decomposition pair. Finally, we show that the SSOSC is further equivalent to the strong metric regularity condition of the subdifferential, the normal map, and the natural residual. Counterexamples illustrate the necessity of our assumptions.

Nonsmooth Nonconvex-Nonconcave Minimax Optimization: Primal-Dual Balancing and Iteration Complexity Analysis – LINGLINGZHI ZHU, THE CHINESE UNIVERSITY OF HONG KONG – Jiajin Li, Man Cho So

In this talk, we consider a class of nonconvex-(non)concave minimax optimization problems with a nonsmooth composite structure. We propose a smoothed proximal linear descent ascent (smoothed PLDA) method for tackling this class of problems and develop a novel analysis framework for studying the iteration complexity of the method. Central to this framework is the Kurdyka-Lojasiewicz (KL) property of the dual function, which provides insights into the interplay between the primal decrease and the dual increase. Consequently, we are able to obtain or match the best known complexity results for various minimax optimization problems. To further demonstrate the effectiveness of our analysis framework, we show that certain max-structured problem possesses the KL property with exponent zero under mild assumptions. We also establish algorithm-independent quantitative relationships among various stationarity concepts.

1d - MB6 - 510A - **Convex Approaches for Quadratically Constrained Quadratic Programs** Chair: SAMUEL BURER

Reformulation-Perspectification Technique for Global Optimization – JIANZHE ZHEN, SWISSQUANT – Dimitris Bertsimas, Danique De Moor, Dick Den Hertog, Thororis Koukouvinos

We propose a new global optimization approach for solv-

ing nonconvex optimization problems in which the nonconvex components are sums of products of convex functions. A broad class of nonconvex problems can be written in this way, such as concave minimization problems, difference of convex problems, and fractional optimization problems. Our approach exploits two techniques: first, we introduce a new technique, called the Reformulation-Perspectification Technique (RPT), to obtain a convex approximation of the considered nonconvex continuous optimization problem. Next, we develop a spatial Branch and Bound scheme, leveraging RPT, to obtain a global optimal solution. Numerical experiments on four different convex maximization problems, a quadratic constrained quadratic optimization problem, and a dike height optimization problem demonstrate the effectiveness of the proposed approach. In particular, our approach solves more instances to global optimality for the considered problems than BARON and SCIP. Moreover, for problem instances of larger dimension our approach outperforms both BARON and SCIP on computation time for most problem instances, while for smaller dimension BARON overall performs better on computation time.

A Slightly Lifted Convex Relaxation for Nonconvex Quadratic Programming with Ball Constraints – SAMUEL BURER, UNIVERSITY OF IOWA –

Globally optimizing a nonconvex quadratic over the intersection of m balls in n -dimensional space is known to be polynomial-time solvable for fixed m . Moreover, when $m=1$, the standard semidefinite relaxation is exact. When $m=2$, it has been shown recently that an exact relaxation can be constructed using a disjunctive semidefinite formulation based essentially on two copies of the $m=1$ case. However, there is no known explicit, tractable, exact convex representation for m bigger than 2. In this paper, we construct a new, polynomially sized semidefinite relaxation for all m , which does not employ a disjunctive approach. We show that our relaxation is exact for $m=2$. Then, for larger m , we demonstrate empirically that it is fast and strong compared to existing relaxations. The key idea of the relaxation is a simple lifting of the original problem into dimension $n+1$. Extending this construction: (i) we show that nonconvex quadratic programming over a ball and a second-order cone, which enforces a linear bound on the norm of the n -dimensional variable, has an exact semidefinite representation; and (ii) we construct a new relaxation for quadratic programming over the intersection of two ellipsoids, which globally solves all instances of a benchmark collection from the literature.

Semidefinite representable reformulations for two variants of the trust-region subproblem – BOSHI YANG, CLEMSON UNIVERSITY – Sarah Kelly, Yuyuan Ouyang

Motivated by encouraging numerical results in Eltvéd and Burer (2023), we consider two specific variants of the trust-region subproblem and provide exact semidefinite representable reformulations. The first is over the intersection of two balls; the second is over the intersection of a ball and a special second-order conic representable set. The reformulations are based on partitions of the feasible regions into sub-regions with known lifted convex hulls.

Hidden convexity, optimization and algorithms on rotation matrices – ALEX WANG,

PURDUE UNIVERSITY –

This talk studies hidden convexity properties associated with constrained optimization problems over the set of rotation matrices $SO(n)$. Such problems are nonconvex due to the constraint X in $SO(n)$. Nonetheless, we show that certain linear images of $SO(n)$ are convex, opening up the possibility for convex optimization algorithms with provable guarantees for these problems. Our main technical contributions show that any two-dimensional image of $SO(n)$ is convex and that the projection of $SO(n)$ onto its strict upper triangular entries is convex. These results allow us to construct exact convex reformulations for constrained optimization problems over $SO(n)$ with a single constraint or with constraints defined by low-rank matrices. Both of these results are optimal in a formal sense.

1d - MB29 - 510D - Recent advances in conic optimization - Part 1 Chair: MAKOTO YAMASHITA



Procrustean Data Collaboration for Cross-Silo Privacy-Preserving Machine Learning – KEIYU NOSAKA, GRADUATE SCHOOL OF SCIENCE AND TECHNOLOGY UNIVERSITY OF TSUKUBA – Akiko Yoshise

Machine learning (ML) technologies have revolutionized various sectors by enabling advanced data analysis and decision-making processes. The effectiveness of these technologies heavily depends on the quality and diversity of their training datasets. While integrating data from multiple sources boosts ML models' predictive power and generalizability, it also introduces significant challenges, most notably the increased risk of user privacy violations. This concern is further compounded by stringent global privacy regulations, including the European Union's General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), and Japan's amended Act on the Protection of Personal Information (APPI), which enforce strict standards on data privacy and user consent.

The Privacy-Preserving Machine Learning (PPML) domain has evolved in response to these challenges, with Federated Learning (FL) emerging as a notable methodology. FL enables collaborative training of either a global model or enhanced individual models across multiple entities while keeping the data localized, thus securely enhancing model performance. A key algorithm within FL, Federated Averaging (FedAvg), exemplifies this by enabling a central server to distribute a model to clients for local updates, as demonstrated in its application in the Google Keyboard for privacy-preserving query suggestions.

However, FL's reliance on iterative communication poses limitations, particularly in cross-silo settings such as healthcare, where data privacy is critical. Traditional FL approaches, which necessitate frequent model exchanges, often prove impractical in these settings.

Data Collaboration (DC) analysis presents a robust alternative. Distinct from typical FL frameworks, DC centralizes secure, dimensionally reduced intermediate representations of raw data rather than sharing models, thus eliminating the need for iterative updates. Each entity independently generates these representations using proprietary dimension reduction techniques akin to data preprocessing methods like differential privacy or k -anonymity. However, DC uniquely preserves

the utility of reduced models by aligning these representations into a unified collaborative representation with minimal distortion.

Our research introduces an optimized collaborative representation critical for the efficacy of the final ML model within DC frameworks. We define and efficiently resolve a collaborative representation optimization problem using the intermediate representations of a shared anchor dataset uniformly distributed to all participating entities. Our proposed technique, Procrustean Data Collaboration (PDC), utilizes orthogonal matrices to minimize distortion in linear transformations, effectively preserving distances and angles between data samples. This approach is analogous to the orthogonal Procrustes problem, which has an established analytical solution.

Empirical evaluations across multiple public datasets demonstrate that our approach adheres to strict privacy standards and significantly improves model performance compared to existing DC frameworks without compromising communication or computational efficiency. We hope the theoretically robust PDC framework represents a significant innovation in PPML, promising to facilitate the broader adoption of ML technologies in privacy-sensitive industries. This is particularly crucial in an era where data privacy concerns are at the forefront of public and regulatory scrutiny.

Riemannian Interior Point Methods for Constrained Optimization on Manifolds – AKIKO YOSHISE, UNIVERSITY OF TSUKUBA – Zhijian Lai

Riemannian Optimization (RO) is a vibrant research area in the field of optimization, which focuses on optimizing real-valued functions over Riemannian manifolds. The manifolds are characterized as smooth curved spaces that generalize Euclidean spaces. Despite the short history of RO, both its theoretical research and practical applications have experienced rapid growth in the past 20 years. The main trend in RO's theoretical research is to extend classical algorithms (e.g., gradient descent, Newton) to manifolds. In practice, manifolds are often represented as the feasible region of problems, while being a subset of a Euclidean space. By utilizing the Riemannian geometry of manifolds, Riemannian algorithms tend to be more efficient than Euclidean algorithms.

Existing research on RO focuses on optimization problems with a smooth objective function over a single manifold. However, this type of optimization problem sometimes does not satisfy diverse demands well. This talk investigates one variant of RO, called Riemannian Constrained Optimization Problem (RCOP), addresses optimization problems involving additional constraints that cannot form a manifold. This variant can address new challenges in practical applications.

In this talk, for the RCOP problem, we will extend the primal-dual interior point algorithms from the Euclidean setting to the Riemannian setting. We call this extension the Riemannian Interior Point Method (RIPM). Our contributions are summarized as follows:

(1) To our knowledge, this is the first study to apply the primal-dual interior point method to the nonconvex constrained optimization problem on Riemannian manifolds. One significant contribution is that we establish many essential foundational concepts for the general interior point method in the Riemannian context, such as the KKT vector field and

its covariant derivative. In addition, we build the first framework for the Riemannian version of the interior point method. These contributions will have uses in the future, especially in developing more advanced interior point methods.

(2) We give a detailed theoretical analysis to ensure local and global convergence of RIPM. Considering that many practical problems involve minimizing a nonconvex function on Riemannian manifolds, the theoretical counterparts of our method are the early interior point methods for nonlinear nonconvex programming first proposed by El-Bakry et al. in 1996.

(3) Our numerical experiments demonstrate the great potential of RIPM. The method meets the challenges presented in these experiments with better stability and higher accuracy compared with the current Riemannian exact penalty, augmented Lagrangian, and sequential quadratic programming methods. The code is freely available at <https://github.com/GALVINLAI/RIPM>

Projection Onto Convex Sets Approach in Solving Homogeneous Self-Dual Model of Quadratically Linear Programming Problems (Case Study: Portfolio Optimization) – SENA SAFARINA, INSTITUT TEKNOLOGI SEPULUH NOPEMBER – Alvian Alif Hidayatullah

The quadratically constrained linear programming (QCLP) problem is a convex optimization problem where a quadratic function serves as a constraint alongside an objective function represented as a linear function. One method for solving convex optimization problems is through the homogeneous self-dual (HSD) model. This model combines the primal and dual problems derived from conic programming problems. It is based on determining a nonzero point from the intersection of two convex sets: a convex cone and a subspace. The HSD model finds application in various problem domains, including QCLP. This research contributes to solving the HSD model of the QCLP problem in the case of portfolio optimization. We proposed an approach based on Dykstra's projection that can be employed to assess the existence of solutions in the HSD model and to approximate the optimal solution of the portfolio optimization problem.

Dual Spectral Projected Gradient Method for Generalized Log-det Semidefinite Programming – CHARLES NAMCHASIRI, TOKYO INSTITUTE OF TECHNOLOGY – Makoto Yamashita

The log-det semidefinite programming is a problem that appears in the Gaussian Graphic Models. There are many methods proposed for solving this problem, and one of them is the Dual Spectral Projected Gradient Method (DSPG) which was proposed by Nakagaki et al. in 2020. The DSPG method is designed to solve the dual problem of the log-det semidefinite programming by combining the non-monotone line search projected gradient method with the step adjustment. In this presentation, we propose the generalized log-det semidefinite programming by adding the extra term into the objective function to support the extra structure of Gaussian Graphical Models and propose the application of the DSPG method to the proposed problem. We show the global convergence of the DSPG on the proposed problem and give some numerical results from the experiments on synthesis data.

1d - MC908 - 510A - **Semi-Definite Programming 1** Chair: ELLEN HIDEMI FUKUDA 

On randomly signed sums of vectors – NATHAN KELLER, BAR ILAN UNIVERSITY – Ohad Klein

A randomly signed sum of vectors is a random variable of the form $X = \sum_{i=1}^n a_i x_i$, where $\{a_i\}$ are vectors in R^d with $\sum_{i=1}^n \|a_i\|_2^2 = 1$, and $\{x_i\}$ are independent 'random signs', distributed as $Pr[x_i = 1] = Pr[x_i = -1] = 1/2$. Questions regarding tail probabilities of the form $Pr[|X| \leq a]$ and $Pr[|X| \geq a]$ appear naturally in problems in optimization theory (e.g., when dealing with robust solutions of uncertain quadratic problems), as well as in various other fields, including probability theory, geometric analysis, statistics, and theoretical computer science. A well-known specific example is Tomaszewski's conjecture from 1986, which asserts that in the one-dimensional case (i.e., where $\{a_i\}$ are real numbers), we have $Pr[|X| \leq 1] \geq 1/2$. The conjecture was recently proved by the authors (Advances in Mathematics, 2022). This talk will survey the tools used in the proof of the conjecture, several very recent developments (including the sharp 'counterpart' bound $Pr[|X| \geq 1] \leq 7/32$ obtained by Hollom and Portier), and various questions which remain open.

Reducing Sum-of-Squares constraints over a domain – BENOÎT LEGAT, KU LEUVEN – Bart De Moor

Numerous applications require verifying the nonnegativity of a polynomial over a set defined by polynomial equalities and inequalities. A classical approach is to search for multipliers that would constitute a Sum-of-Squares certificate of the nonnegativity. While this corresponds to a hierarchy of convex programs, the size of these programs grows rapidly with the total degree of the multipliers. If a Gröbner basis for the ideal generated by the equalities is known, it would allow to drastically reduce the size of the program to solve. Such basis is however NP-hard to compute. Using information from the null space of a Macaulay matrix, we show how to reduce the size of the programs for any degree. While the program size converges to the size that would be obtained by the Gröbner basis approach, our approach does not need Gröbner basis to be computed and is therefore not limited to small size problems. Furthermore, we show how to also compute reduced gram bases for the multipliers associated to the inequalities by exploiting both properties of a Macaulay null space and Newton polytopes. We show applications of these results to the analysis of descriptor systems.

Approximate Karush-Kuhn-Tucker 2 conditions for nonlinear second-order cone optimization problems – ELLEN HIDEMI FUKUDA, KYOTO UNIVERSITY – Kosuke Okabe

The sequential optimality conditions, which improve algorithm convergence assumptions, have been studied over the past two decades. In particular, by considering the second-order information, the so-called approximate Karush-Kuhn-Tucker 2 (AKKT2) conditions for nonlinear programming problems were also proposed. While studies on a more general conic problem have been conducted, an explicit definition for AKKT2 has not been proposed in the literature. This work addresses this gap by providing a precise definition of AKKT2 conditions for second-order cone programming problems. We prove that

the proposed AKKT2 conditions hold at local optimal points, without any constraint qualifications. Moreover, we show two algorithms based on augmented Lagrangian and sequential quadratic programming approaches that generate points satisfying these conditions.

1d - TA138 - 510A - **Moment, Polynomial and Tensor Optimization - Part 1** Chair: SUHAN ZHONG 

Convergence of the momentum method for semialgebraic functions with locally Lipschitz gradients – CEDRIC JOSZ, COLUMBIA UNIVERSITY – Lexiao Lai, Xiaopeng Li

We propose a new length formula that governs the iterates of the momentum method when minimizing differentiable semialgebraic functions with locally Lipschitz gradients. It enables us to establish local convergence, global convergence, and convergence to local minimizers without assuming global Lipschitz continuity of the gradient, coercivity, and a global growth condition, as is done in the literature. As a result, we provide the first convergence guarantee of the momentum method starting from arbitrary initial points when applied to matrix factorization, matrix sensing, and linear neural networks.

Our contributions are as follows. We consider objective functions that are semialgebraic and differentiable with locally Lipschitz gradients. We show that the length of the iterates generated by the momentum method is upper bounded by an expression depending on the objective function variation, the step size, and a desingularizing function. This length formula enables us to show that global Lipschitz continuity of the gradient and the global growth condition are superfluous when establishing local convergence. It also enables us to establish global convergence under the assumption that the continuous-time gradient trajectories of the objective function are bounded. As a result, we bypass the need for coercivity and globally Lipschitz gradients. Finally, the length formula enables us to guarantee convergence to local minimizers almost surely, under second-order differentiability and the strict saddle property.

Certifying nonnegative functions on finite sets via Fourier SOS – KE YE, CHINESE ACADEMY OF SCIENCES –

This talk consists of two parts. In the first part, we introduce a framework of certifying the non-negativity of a function on a finite set, which is equivalent to certifying the nonnegativity of tensors. Via the Fourier sum of squares (FSOS), we are able to discuss both the theoretical and algorithmic aspects of this problem. The second part is concerned with applications of our framework to MAX-SAT. We will discuss how to find low degree FSOS certificates and estimate the lower bound for MAX-SAT problems.

Generalized Nash equilibrium problems with Quasi-linear constraints – JIYOUNG CHOI, UNIVERSITY OF CALIFORNIA SAN DIEGO – Jiawang Nie, Xindong Tang, Suhan Zhong

We discuss generalized Nash equilibrium problems (GNEPs) that are given by polynomial functions. We consider cases each player's constraints are linear in their own strategy vectors. For this kind of GNEPs, the KKT set can be conveniently represented as a union of simpler sets, by using partial Lagrange

multiplier expressions. This representation produces a set of branch polynomial optimization problems, each of which can be conveniently solved by Moment-SOS relaxations. We give a method to compute all generalized Nash equilibria, under some genericity assumptions. Extensive numerical experiments are provided to demonstrate the efficiency of the proposed method.

Solving robotic control problems with safety constraints using moment and SOS relaxations – XINDONG TANG, HONG KONG BAPTIST UNIVERSITY –

We consider the robotic control problems with safety constraints, which can be represented as the nonnegativity on a semi-algebraic set of polynomials parameterized by robotic control variables. We apply sum-of-squares (SOS) relaxations to characterize safe controls, and we use moment relaxations to obtain semidefinite relaxation problems. Under mild conditions, we can get optimal safe control for the robot. Numerical simulation and real-robot experiments are made to show the efficiency of our approach.

1d - TB66 - 510A - **SDP and Combinatorial Optimization** Chair: NATHAN KRISLOCK 

Eigenvalue cuts for the Traveling Salesman Problem – MICHAEL SMITH, LEWIS UNIVERSITY/NORTHERN ILLINOIS UNIVERSITY –

Solutions to the Traveling Salesman Problem (TSP) are known to come from permutation matrices (in the asymmetric case) and adjacency matrices (in the symmetric case). With this knowledge, we propose an alternative formulation of the sub-tour elimination constraints that is based on eigenvalues and eigenvectors of permutation and adjacency matrices. Both SDP and non-SDP formulations of this approach will be discussed. Comparisons to existing methods will be explored, and potential paths for future improvement will be identified.

Computational study of rounded psd inequalities for the Max-Cut and k-cluster problem – TIMOTEJ HRGA, UNIVERSITY OF KLAGENFURT –

We present a computational study of SDP-based bounds for Max-Cut and k-cluster problems that use rounded psd inequalities as cutting planes to strengthen the basic relaxation. These are a subset of the gap inequalities introduced by Laurent and Poljak. We propose a fast optimization algorithm to efficiently separate violated rounded psd inequalities and use the augmented Lagrangian method as a bounding routine. Two simple and effective heuristics based on the greedy algorithm and a variation of the alternating direction method of multipliers (ADMM) are presented. Our numerical experiments show that our approach is very efficient in finding cutting planes with large violation and provide strong upper bounds in practice.

Edge Expansion: Exploring SDP-Based Computational Strategies – ANGELIKA WIEGELE, UNIVERSITY OF KLAGENFURT – Akshay Gupte, Melanie Siebenhofer

Computing the edge expansion of a graph is a famously hard combinatorial problem for which there have been many approximation studies. We present two variants of exact algorithms using semidefinite programming (SDP) to compute this

constant for any graph. The first variant uses the SDP relaxation first to reduce the search space considerably. One implementation of this variant applies then an SDP-based branch-and-bound algorithm, along with heuristic search. The second implementation transforms the problem into an instance of a max-cut problem and solves this using the state-of-the-art solver BiqBin.

Our second variant to compute the edge expansion uses Dinkelbach's algorithm for fractional programming, again, using tools from semidefinite programming.

To the best of our knowledge, these are the first SDP-based solvers for computing the edge expansion of a graph.

Facial Reduction of Semidefinite Relaxations of Binary Quadratic Problems – IAN SUGRUE, NORTHERN ILLINOIS UNIVERSITY –

A standard technique for bounding a combinatorial optimization problem is to form a semidefinite relaxation of the problem by “lifting” the problem into a higher dimensional space of symmetric matrices. For some combinatorial problems such as maximum k-cluster, the resulting semidefinite relaxation is not strictly feasible. In the context of the maximum k-cluster problem, we present an equivalent, strictly feasible semidefinite relaxation by way of facial reduction. We provide a recipe to form this semidefinite relaxation directly, bypassing the lifting and facial reduction steps entirely. We then present our work on generalizing this result, and provide a recipe for constructing a strictly feasible semidefinite relaxation of binary quadratic problems with a single linear constraint with non-zero right-hand side. In addition, we also present numerical results that support the inclusion of facial reduction in the binary quadratic solver BiqCrunch, as well as results supporting the use of facial reduction when solving a semidefinite relaxation of the maximum k-cluster problem.

1d - TB87 - 510D - **Semidefinite Programming and Convex Algebraic Geometry** Chair: REKHA THOMAS 

Closing nonzero duality gaps by perturbation in SDP and its extension – TAKASHI TSUCHIYA, NATIONAL GRADUATE INSTITUTE FOR POLICY STUDIES – Bruno Lourenço, Masakazu Muramatsu, Takayuki Okuno

A semidefinite program (SDP) is called singular if primal (P) and dual (D) are weakly feasible or weakly infeasible. Singular SDPs may have nonzero duality gaps and are hard to solve. In this talk, we analyze singular SDPs with finite nonzero duality gaps. A possible approach to solve such SDPs by the interior-point method is to perturb primal and dual to recover strong feasibility shifting the positive semidefinite cones of primal and dual by ηI_P and εI_D , respectively, where I_P and I_D are symmetric positive definite matrices and η and ε are positive numbers. Let us denote by $v(\varepsilon, \eta)$ the common optimal value of $v(\varepsilon, \eta)$ (if it exists). It is known that $v(\varepsilon, \eta)$ is finite and well-defined for $(0, 0) \neq (\varepsilon, \eta) \geq 0$. In particular, if (P) and (D) admit nonzero duality gap, $v(0, 0)$ is not defined, and its behavior in the neighborhood of $(0, 0)$ was not clear. Recently, the authors showed that $v(\varepsilon, \eta)$ has a directional limit $v_a(\theta)$ when approaching to $(0, 0)$ from a unique direction where θ is direction of approach, and $v_a(\theta)$ and the limiting value converges to “a value between the optimal value of (P) and (D)”.

In this talk, we show that if the singularity degree of (P) and (D) is one, v_a is a bijective mapping from the domain $[0, \pi/2]$ to $[v(D), v(P)]$, where $v(D)$ and $v(P)$ are the optimal values of (D) and (P), respectively, closing completely the duality gap. We also discuss possible extensions of some properties of $v(\varepsilon, \eta)$ explained above to general conic linear programs.

Any-dimensional learning – EITAN LEVIN, CALTECH – Venkat Chandrasekaran

Mappings arising in various application domains are defined for inputs of different sizes. Examples include standard norms defined for vectors and matrices of any size, graph parameters defined for graphs of any size, and physics quantities defined for systems of any number of particles. In contrast, standard supervised learning deployed to learn these mappings operates with training data of a fixed size, and outputs a mapping only defined for inputs of that same size. How can we learn mappings defined for inputs of infinitely-many sizes from a finite dataset? How do we describe such mappings, and how do we parametrize them to enable such learning from data?

In this talk, we tackle these questions for mappings representable as optimal values or solutions of conic programs. Describing such mappings reduces to describing convex sets. These, in turn, are often “freely” described, meaning that their description makes instantiation in every dimension obvious. We show that such free descriptions of convex sets arise from two ingredients: group invariance and the recently-identified phenomenon of representation stability; and embeddings and projections relating different-sized problem instances. We combine these ingredients to obtain parametrized families of infinitely instantiable convex sets. To extend a set learned from data in a fixed dimension to higher ones, we identify compatibility conditions relating sets in different dimensions that are satisfied in a variety of applications, and obtain parametrizations respecting these conditions. Our parametrizations can be obtained computationally.

Finally, we take a statistical learning perspective on the problem and bound the error any estimator incurs when extended to higher dimensions.

Polynomial Lower Approximation for Two-Stage Stochastic Optimization – SUHAN ZHONG, TEXAS A&M UNIVERSITY – Ying Cui, Jiawang Nie

This talk discusses the challenging problem of finding global optimal solutions for two-stage stochastic programs with continuous decision variables and nonconvex recourse functions. We introduce a two-phase approach. The first phase involves the construction of a polynomial lower bound for the recourse function through a linear optimization problem over a nonnegative polynomial cone. Given the complex structure of this cone, we employ semidefinite relaxations with quadratic modules to facilitate our computations. In the second phase, we solve a surrogate first-stage problem by substituting the original recourse function with the polynomial lower approximation obtained in the first phase. Our method is particularly advantageous for two reasons: it not only generates global lower bounds for the nonconvex stochastic program, aiding in the verification of global optimality for prospective solutions like stationary solutions computed from other methods, but it also simplifies the computation of the expected value of the recourse function by using moments of random vectors. This makes our overall algorithm particularly suitable for the case where the

random vector follows a continuous distribution or when dealing with a large number of scenarios. Numerical experiments are given to demonstrate the effectiveness of our proposed approach.

Cubic interpolation to certify nonnegativity – MITCHELL HARRIS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Pablo Parrilo

A number of problems can be framed as mathematical programs with the special constraint that a function is nonnegative. For many years, custom methods were devised for these so-called semi-infinite optimization problems. Recently, sum of squares programming was introduced to allow many of the constraints to be expressible via semidefinite programming. Unfortunately, sum of squares can only be applied directly when the constrained function is a polynomial. Furthermore, the scaling of the resulting semidefinite programs makes this approach impractical for high degree even in low dimension.

We propose a method for semi-infinite optimization problems harnessing the combined power of low degree sum of squares and the ability of cubic splines to approximate smooth functions very well. We replace function nonnegativity constraints by nonnegative constraints of the functions’ cubic interpolants. The cubic nonnegativity constraints are second order cone constraints – even simpler than general semidefinite constraints. A simple local method that relies only on linear interpolation (and results in a linear program) requires finer discretizations (and overall longer computation times) than are needed when using cubic approximation. We present theoretical guarantees for our method, theoretical comparisons to the linear programming competitor, and applications demonstrating the empirical benefits of using cubic interpolation.

1d - TC23 - 510A - **Semidefinite and Polynomial Optimization Chair: FRANK VALLENTIN** 

Semidefinite Programming Bounds for the Kissing Number – WILLEM DE MUINCK KEIZER, TU DELFT –

In discrete geometry, linear programming bounds are important for bounding the quality of geometric configurations. These bounds have been used to show that constructions are optimal, for example the sphere packings coming from the E8 and Leech lattices. However, for many instances the linear programming bounds are not sharp and research is being done into semidefinite programming bounds. For many problems, these higher-order bounds lead to significant improvements, including new optimality proofs via sharp bounds. The kissing number is the maximum number of non-overlapping unit spheres in n-dimensional Euclidean space that can simultaneously touch a central unit sphere. Determining the kissing number is an open problem except for dimensions 1, 2, 3, 4, 8 and 24. Bachoc and Vallentin gave improved upper bounds for the kissing number problem by computing a three-point semidefinite programming bound. We compute higher steps of the Lasserre hierarchy for this problem, which is also a semidefinite programming bound and which may be viewed as a four-point bound. To calculate this bound in practice, it is crucial to use symmetry reduction. This requires calculations related to the representation theory of the orthogonal group and invariants thereof. The resulting semidefinite program is challenging to implement and solve.

Our results give improvements in upper bounds for the kissing number.

Non-commutative polynomial optimization with polynomial constraints for quantum causality – DAVID GROSS, UNIVERSITY OF COLOGNE – Laurens Ligthart, Mariami Gachechiladze

The "practical" problem we are interested in is this: Given data obtained from a physics experiment, can one decide whether they are compatible with a quantum-mechanical model subject to causality constraints? A "causality constraint" here demands that certain subsystems be statistically independent. Mathematically, this problem is related to non-commutative (NC) polynomial optimization: Given generators and relations, optimize over all states of the universal C^* algebra they generate. The causal conditions give rise to polynomial constraints in the states, which until recently could not be handled in this setting. I'll report on results both on the mathematical and the physical problem.

Joint work with Laurens Ligthart and Mariami Gachechiladze. [2110.14659 (CMP), 2212.11299 (JMP)]

On some properties and extensions of the Christoffel function – JEAN B. LASSERRE, LAAS CNRS UNIV. TOULOUSE, FRANCE –

We provide a brief description on some (surprising) links of the Christoffel function with seemingly unrelated fields like algebraic positivity certificates, equilibrium measures of compact sets, and a generalized polynomial Pell's equation.

1d - TC299 - 510D - Semidefinite Programming for Geometric Processing Chair: AMIT SINGER ←

Benign nonconvexity of Burer-Monteiro SDP factorization for group synchronization – ANDREW MCRAE, EPFL – Nicolas Boumal

I will present and analyze nonconvex low-rank Burer-Monteiro factored approaches to Max-Cut-like SDPs for orthogonal group synchronization. In this problem, we want to estimate orthogonal group elements from (potentially noisy) pairwise relative measurements. A nonconvex least-squares estimator over the (product) orthogonal group can be lifted and relaxed to an SDP with block-diagonal constraints. However, this lifting (approximately) squares the number of variables. To reduce the dimension, we then consider a low-rank nonconvex factored formulation, which can be viewed as a partial relaxation of the original group problem. We show that, due to the problem structure, increasing the number of variables by a small constant factor results in a benign landscape in the sense that every second-order critical point yields the SDP solution; furthermore, our analysis shows that the SDP relaxation is tight (i.e., yields a solution to the original group synchronization problem), and therefore so is the nonconvex partial relaxation. Key novelties in our results are applicability to general measurement graphs, reaching the optimal relaxation dimension, and extensions to clustering problems. Joint work with Nicolas Boumal, published (in part) in <https://doi.org/10.1137/23M1584642>.

Non-unique games over compact groups and orientation estimation in cryo-EM – AMIT SINGER, PRINCETON UNIVERSITY –

Let G be a compact group and let f_{ij} be bandlimited functions over G . We define the non-unique games (NUG) problem as finding a set of group elements $g_1, \dots, g_n$ of G that minimize $\sum_{i,j=1}^n f_{ij}(g_i g_j^{-1})$. We introduce a convex relaxation of the NUG problem to a semidefinite program (SDP) by taking the Fourier transform of f_{ij} over G . The NUG framework can be seen as a generalization of the little Grothendieck problem over the orthogonal group and the unique games problem and includes many practically relevant problems, such as the maximum likelihood estimator to registering bandlimited functions over the unit sphere in d -dimensions and orientation estimation of noisy cryo-electron microscopy (cryo-EM) projection images. We implement an SDP solver for the NUG cryo-EM problem using the alternating direction method of multipliers (ADMM). Numerical study with synthetic datasets indicate that while our ADMM solver is slower than existing methods, it can estimate the rotations more accurately, especially at low signal-to-noise ratio (SNR). If time permits we will also discuss generalization to orbit estimation in the case of non-trivial point group symmetry. Joint work with Afonso Bandeira, Yutong Chen, Roy Lederman, and Ruiyi Yang.

SDP Relaxation and Topological Defects in Meshing – DAVID PALMER, HARVARD UNIVERSITY –

Constructing a hexahedral mesh of a domain is an essential prerequisite in many techniques for simulating physical systems, but automatic hexahedral meshing remains a challenge. A promising approach is to compute a frame field as a proxy for the overall geometry of the mesh. While this decouples geometric optimization from the combinatorics of the meshing problem, frame field optimization is still challenging due to the presence of singularities, i.e., topological defects. I will outline our work on alleviating these "topological barriers" using SDP-based projection onto spaces of symmetric frames. Along the way, we will explore the geometry of frame fields and their singularities. Time permitting, I will discuss our recent work on lifting and current relaxation of frame field optimization and offer some speculation on global SDP relaxation of field optimization problems.

1d - WA97 - 510A - Semidefinite programming, sums of squares, and quantum information Chair: HAMZA FAWZI ←

Bell nonlocality and Grothendieck constants via Frank-Wolfe algorithms – SÉBASTIEN DESIGNOLLE, ZUSE INSTITUTE BERLIN – Mathieu Besançon, Tamás Vértesi, Gabriele Iommazzo, Sebastian Knebel, Patrick Gelß, Sebastian Pokutta

In Bell scenarios with two outcomes per party, we algorithmically consider the two sides of the membership problem for the local polytope: constructing local models and deriving separating hyperplanes, that is, Bell inequalities. We take advantage of the recent developments in so-called Frank-Wolfe algorithms to significantly increase the convergence rate of existing techniques. With this method, we study the threshold value for the nonlocality of two-qubit Werner states under projective measurements. As already noticed by Tsirelson, this quantity is the inverse of the Grothendieck constant of order three, for which we refine the existing bounds. We further apply the method to improve the bounds on Grothendieck constants of

higher orders. These constants are approximation factors of the rank-constrained SDP relaxations of the Max-Cut problem, a connection that we discuss in more details throughout the presentation. We also study multipartite Bell scenarios, especially a symmetric case involving the Green-Horne-Zeilinger state. For this instance, we exploit the symmetry to drastically accelerate the computation of (i) a Bell inequality via Frank-Wolfe algorithms, and (ii) the corresponding local bound. We compute the detection efficiency of our inequalities and show that some give rise to activation of nonlocality in star networks, a property that was only shown before with an infinite number of measurements.

A simple approximation algorithm for Quantum Max Cut via Maximum Matching – OJAS PAREKH, SANDIA NATIONAL LABORATORIES – Eunou Lee

Finding a high (or low) energy state of quantum local Hamiltonians is fundamental physical problem and a potential area to gain a provable and practical quantum advantage. A recent line of work has focused on developing approximation algorithms for Quantum Max Cut, a quantum local Hamiltonian analogue of Max Cut, using non-commutative sum of squares (ncSoS) hierarchies. I will describe a simple ncSoS-based classical approximation algorithm for Quantum Max Cut achieving an approximation ratio of 0.595, outperforming the previous best algorithms of Lee (0.562, generic input graph) and King (0.582, triangle-free input graph). The algorithm is based on finding a maximum weight matching of the input graph and outputs a product of at most 2-qubit states, which is simpler than the fully entangled output states produced by the previous best algorithms. The analysis leverages monogamy of entanglement inequalities satisfied by ncSoS solutions as well as connections to the matching polytope.

Analysis of ncSoS relaxations for the quantum rotor model – SUJIT RAO, MIT –

The noncommutative sum-of-squares (ncSoS) hierarchy was introduced by Navascués–Pironio–Acín as a sequence of semidefinite programming relaxations for approximating values of “noncommutative polynomial optimization problems,” which were originally intended to generalize quantum values of nonlocal games. Recent work has started to analyze the hierarchy for approximating ground energies of local Hamiltonians, initially through rounding algorithms which output product states for degree-2 ncSoS. Some rounding methods are known which output entangled states, but they use degree-4 ncSoS. Based on this, Hwang–Neeman–Parekh–Thompson–Wright conjectured that degree-2 ncSoS cannot beat product state approximations for quantum max-cut and gave a partial proof relying on a conjectural generalization of Borrell’s inequality. In this talk we will describe the ncSoS hierarchy and a family of Hamiltonians (called the quantum rotor model in condensed matter literature or lattice $O(n)$ vector model in QFT) over an infinite-dimensional local Hilbert space, and sketch a proof that a degree-2 ncSoS relaxation approximates the ground energy better than any product state.

Relaxations and Exact Solutions to Quantum Max Cut via the Algebraic Structure of Swap Operators – ADAM BENE WATTS, UNIVERSITY OF WATERLOO – Anirban Chowdhury, Aidan Epperly, Igor Klep, Bill Helton

The Quantum Max Cut (QMC) problem has emerged as a test-problem for designing approximation algorithms for local Hamiltonian problems. In this talk I’ll describe an approaches to solving this problem (approximately and exactly) using the relationship between the quantum max cut Hamiltonian and the representation theory of the symmetric group. The first major result I’ll describe is an extension of non-commutative Sum of Squares (ncSoS) optimization techniques to give a new hierarchy of relaxations to Quantum Max Cut. This hierarchy is based on optimizations over polynomials in the qubit swap operators. This is in contrast to the “standard” quantum Lasserre Hierarchy, which is based on polynomials expressed in terms of the Pauli matrices. The second major result covered will be a polynomial-time algorithm that exactly computes the maximum eigenvalue of the QMC Hamiltonian for certain graphs, including graphs that can be “decomposed” as a signed combination of cliques. A special case of the latter are complete bipartite graphs with uniform edge-weights, for which exact solutions are known from the work of Lieb and Mattis [LM62]. Our methods, which use representation theory of the symmetric group, can be seen as a generalization of the Lieb-Mattis result.

1d - WA208 - 510D - **Low-rank semidefinite optimization, algorithms, software and applications** Chair: MICHAL KOCVARA

Accelerating nuclear-norm regularized low-rank matrix optimization through Burer-Monteiro decomposition – TIANYUN TANG, NATIONAL UNIVERSITY OF SINGAPORE – Ching Pei Lee, Ling Liang, Kim Chuan Toh

This work proposes a rapid algorithm, BM-Global, for nuclear-norm-regularized convex and low-rank matrix optimization problems. BM-Global efficiently decreases the objective value via low-cost steps leveraging the nonconvex but smooth Burer-Monteiro (BM) decomposition, while effectively escapes saddle points and spurious local minima ubiquitous in the BM form to obtain guarantees of fast convergence rates to the global optima of the original nuclear-norm-regularized problem through aperiodic inexact proximal gradient steps on it. The proposed approach adaptively adjusts the rank for the BM decomposition and can provably identify an optimal rank for the BM decomposition problem automatically in the course of optimization through tools of manifold identification. BM-Global hence also spends significantly less time on parameter tuning than existing matrix-factorization methods, which require an exhaustive search for finding this optimal rank. Extensive experiments on real-world large-scale problems of recommendation systems, regularized kernel estimation, and molecular conformation confirm that BM-Global can indeed effectively escapes spurious local minima at which existing BM approaches are stuck, and is a magnitude faster than state-of-the-art algorithms for low-rank matrix optimization problems involving a nuclear-norm regularizer.

Well-conditioned primal-dual interior-point method for accurate low-rank semidefinite programming – HONG MING CHIU, UNIVERSITY OF ILLINOIS URBANA CHAMPAIGN – Richard Zhang
We describe how the low-rank structure in an SDP can be

exploited to reduce the per-iteration cost of a convex primal-dual interior-point method down to cubic time and quadratic memory, even at very high accuracies. A traditional difficulty is the dense Newton subproblem at each iteration, which becomes progressively ill-conditioned as progress is made towards the solution. Preconditioners have previously been proposed to improve conditioning, but these can be expensive to set up, and become ineffective as the preconditioner itself becomes increasingly ill-conditioned at high accuracies. Instead, we present a well-conditioned reformulation of the Newton subproblem that is cheap to set up, and whose condition number is guaranteed to remain bounded over all iterations. In theory, applying an inner iterative method to the reformulation reduces the per-iteration cost of the outer interior-point method to cubic time and quadratic memory. We also present a well-conditioned preconditioner that greatly improves the convergence of the inner iterations.

Loraine news: From a low-rank to a high-precision SDP solver – MICHAL KOCVARA, UNIVERSITY OF BIRMINGHAM –

Loraine.jl is a Julia SDP solver, originally developed for large-scale problems with low-rank solutions. We will present a second generation of the code, focused on high-accuracy solution of small but difficult SDP problems. The code uses extended-precision arithmetic provided by MultiFloats.jl. We will demonstrate that the solver can deliver reliable solutions to difficult (highly ill-conditioned or almost infeasible) problems arising in Steady-State Power Network Optimization or in global polynomial optimization. In particular, we will give numerical confirmation of theoretical convergence estimates of Lasserre hierarchy for certain polynomial optimization problems studied in the literature. For instance, we will show that Loraine can give a reliable solution to problems with relaxation order 35 and more, much higher than standard SDP solvers.

An Augmented Lagrangian Primal-Dual Semismooth Newton Method for Multi-Block Composite Optimization – ZHANWANG DENG, PEKING UNIVERSITY – Zaiwen Wen

In this talk, we present a novel primal-dual semismooth Newton method for solving linearly constrained multi-block convex composite optimization problems. First, a differentiable augmented Lagrangian (AL) function is constructed by utilizing the Moreau envelopes of the nonsmooth functions. Then we derive a semismooth system of nonlinear equations and develop a semismooth Newton method ALPDSN. Through a connection to the inexact first-order steps when the regularization parameter is sufficiently large, the global convergence of ALPDSN is established. Under the regularity conditions, partial smoothness, the local error bound, and the strict complementarity, we show that both the primal and the dual iteration sequences possess a superlinear convergence rate. Numerical results demonstrate the high efficiency and robustness of our ALPDSN.

1d - WB18 - 510A - **Semidefinite Programming Performance Estimation Chair: ETIENNE DE KLERK**

Accelerating Gradient Descent – BENJAMIN GRIMMER, JOHNS HOPKINS UNIVERSITY – Kevin Shu, Alex Wang

This talk will discuss several recent works establishing provably faster convergence rates for gradient descent in smooth convex optimization via semidefinite programming and computer-assisted analysis techniques (i.e., performance estimation problems). First, we will discuss recent advances from allowing non-constant stepsize policies with frequent long steps potentially violating descent. We manage this in analysis by considering the overall effect of many iterations at once rather than the typical one-iteration inductions used in most first-order method analyses. Doing so, we prove a $O(1/T^{1.0564})$ convergence rate, beating the classic $O(1/T)$ rate simply by periodically including longer steps (no momentum needed). Secondly, additional results that exactly characterize the benefits of iterate averaging and extrapolation for gradient descent will be presented. We will show iterate averaging provably offers no benefit in terms of worst case guarantees. In contrast, we show a trivially simple extrapolation step at the end of gradient descent (costing a single scalar-vector multiply) can improve the worst-case rate of convergence by the same amount as $O(\sqrt{T}/\log(T))$ additional gradient descent iterations. Numerical results supporting the benefit of this simple, cheap extrapolation trick for a range of gradient methods will be presented as well.

Performance estimation for SDP hierarchies for polynomial optimization on the hypercube – ETIENNE DE KLERK, TILBURG UNIVERSITY – Juan Vera Lizcano

We present a way to estimate the worst-case performance of an SDP hierarchy for polynomial optimization on the hypercube. The SDP hierarchy in question is based on the Schmuedgen Positivstellensatz. Some error bounds for this hierarchy are known from recent work by Laurent and Slot [Optimization Letters 17:515-530, 2023]. We show how to obtain similar error bounds in a novel way, but also how to obtain error bounds that may be computed a priori via SDP.

The exact worst-case convergence rate of the alternating direction method of multipliers – HADI ABBASZADEHPEIVASTI, TILBURG UNIVERSITY –

Recently, semidefinite programming performance estimation has been employed as a strong tool for the worst-case performance analysis of first order methods. In this talk, I derive new non-ergodic convergence rates for the alternating direction method of multipliers (ADMM) by using performance estimation. I also study the linear and R-linear convergence of ADMM in terms of dual objective. I establish that ADMM enjoys a global linear convergence rate if and only if the dual objective satisfies the Polyak–Łojasiewicz (PL) inequality in the presence of strong convexity. In addition, I give an explicit formula for the linear convergence rate factor. Moreover, I study the R-linear convergence of ADMM under two scenarios.

Acceleration by Stepsize Hedging – JASON ALTSCHULER, UNIVERSITY OF PENNSYLVANIA –

Can we accelerate the convergence of gradient descent without changing the algorithm — just by optimizing stepsizes? Surprisingly, we show that the answer is yes. Our proposed Silver Stepsize Schedule optimizes strongly convex functions in $k^{\log_p 2} = k^{0.7864}$ iterations, where $p = 1 + \sqrt{2}$ is the silver ratio and k is the condition number. This is intermediate between

the textbook unaccelerated rate k and the accelerated rate $\sqrt{k}$ due to Nesterov in 1983. The non-strongly convex setting is conceptually identical and leads to an analogously accelerated rate $\epsilon^{-\log_p 2} = \epsilon^{-0.7864}$. We conjecture and provide partial evidence that these rates are optimal among all possible stepsize schedules.

The Silver Stepsize Schedule is an explicit non-monotonic fractal. Why should such stepsizes help? The core intuition is “hedging” between individually suboptimal strategies — short steps and long steps — since bad cases for the former are good cases for the latter, and vice versa. Properly combining these stepsizes yields faster convergence due to the misalignment of worst-case functions. This talk is based on a line of work with Pablo Parrilo that originates from my 2018 Master’s Thesis — which established for the first time that judiciously chosen stepsizes can enable accelerated convex optimization. Prior to this thesis, the only such result was for the special case of quadratics, due to Young in 1953.

1d - WB222 - 510D - **Algorithms and Applications of Semidefinite Programming** Chair: CIFUENTES DIEGO

Solving Exact and Noisy Rank-one Tensor Completion with Semidefinite Programming – ZHUORUI LI, GEORGIA INSTITUTE OF TECHNOLOGY – Cifuentes Diego

Consider the problem of retrieving a d -th order rank-one tensor given (exact or noisy) values of some of its entries. We address this problem via semidefinite programming (SDP). Concretely, we propose SDP relaxations and we study deterministic conditions under which the SDPs solve the completion problems exactly. We show that our SDPs solve the exact and noisy rank-one tensor completion problems when the observations satisfy a combinatorial condition, which we call square-propagation. We prove that the square-propagation condition is satisfied with high probability when the number of uniformly random observations is greater than or equal to some quantity. Moreover, when d equals 2 (matrix completion), the square-propagation is satisfied precisely when the completion problem admits a unique solution. Preliminary computational experiments show that our methods allow to solve tensor completion problems using significantly less observations than alternative methods.

Random Projections for Semidefinite Programming and Polynomial Optimization – MONSE GUEDES AYALA, THE UNIVERSITY OF EDINBURGH – Pierre Louis Poirion, Lars Schewe, Akiko Takeda

Random projections, a dimensionality reduction technique, have been found useful in recent years for reducing the size of optimization problems. In this presentation, we explore using random projections to approximate semidefinite programming (SDP) problems by reducing the size of matrix variables, thereby solving the original problem with less computational effort. We provide some theoretical guarantees on the quality of the projection. We also investigate the performance of the approach on semidefinite relaxations appearing in polynomial optimization, with a focus on combinatorial optimization problems.

HALLaR: A New Solver for Solving Huge SDPs – ARNESH SUJANANI, GEORGIA INSTITUTE OF TECHNOLOGY – Renato Monteiro, Cifuentes Diego

This talk discusses the computational aspects of a new first-order method for solving large-scale semidefinite programs (SDPs) with bounded domain. It is an inexact augmented Lagrangian (AL) method where the AL subproblems are solved by a hybrid low-rank method. In contrast to the classical low-rank method, the new method finds a near-optimal solution (with provable complexity bounds) of SDP instances satisfying strong duality. Computational results comparing the new method to state-of-the-art solvers on several large SDP instances show that the former finds higher accurate solutions in substantially less CPU time than the latter ones. For example, in less than 20 minutes, our method can solve a maximum stable set SDP with 1 million vertices and 10 million edges within 1e-5 relative accuracy.

Sum of squares submodularity – GEORGINA HALL, INSEAD –

Submodularity is a key property of set-valued functions, appearing in many different areas such as game theory, machine learning, and optimization. In this talk, we consider set-valued functions over $\{0,1\}^n$ with a multilinear extension of degree less than or equal to some k . For these functions, it is known that testing whether they are submodular is hard for k greater or equal to 4. We thus introduce a sufficient condition for submodularity based on sum of squares polynomials, which we refer to as sum of squares (sos) submodularity. We investigate the gap between submodularity and sos submodularity and show how to use semidefinite programming to optimize over this set. In particular, we use the techniques proposed to learn a submodular function from data, among other applications.

1d - WC909 - 510A - **Semi-Definite Programming 2** Chair: DAVID PAPP

Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem –

MANUEL RUIZ, RTE FRANCE – Patrick Panciatici, Jean B. Lasserre, Victor Magron, Adrien Le Franc

AC-OPF (Alternative Current - Optimal Power Flow) aims at minimizing the operating costs of an AC power grid. It is well-known to be a difficult optimization problem in general, as it reformulates as a nonconvex QCQP (Quadratically Constrained Quadratic Program).

The moment-SOS (Sums-Of-Squares) hierarchy has proved relevant to solve AC-OPF instances to global optimality. However, obtaining the convergence of the hierarchy may require to go beyond the first step of the involved sequence of SDP (Semidefinite Programming) relaxations and thus to solve semidefinite programs whose size grows drastically at each step of the hierarchy. Thus, the application of such relaxations to large scale AC-OPF problems (with thousands of variables) remains a challenging computing task.

Large polynomial optimization problems can be tackled efficiently if they are sufficiently sparse. In this talk, we present a new sparsity pattern, that we call minimal sparsity, inspired by the specific structure of the AC-OPF problem.

We show that minimal sparsity enables the computation of second order moment-SOS relaxations on large scale AC-OPF instances with far less computing resources — i.e. RAM and time — than the standard correlative sparsity pattern. Experimentally, we observe that it also provides tight lower bounds to certify the global optimality of AC-OPF solutions.

Rational Dual Certificates for Weighted Sums-of-Squares Polynomials: Algorithms and Complexity Bounds – DAVID PAPP, NORTH CAROLINA STATE UNIVERSITY –

Sum-of-squares (SOS) polynomials with rational coefficients in the interior of the SOS cones are sums of rational squares, meaning that they can be written as a sum of squares of rational polynomials, allowing for easy certification of their non-negativity. The complexity of these certificates is measured by their bit sizes but are challenging to bound even in the univariate case. We study this problem in the context of dual certificates, which are rational vectors from the dual-SOS cone that can be interpreted as efficiently verifiable nonnegativity certificates.

The talk will outline a very simple algorithm to compute a sequence of SOS lower bounds converging linearly to the optimal SOS lower bound and corresponding "small" rational dual certificates, whose bit sizes can be bounded as a function of interpretable parameters such as the degree and the number of variables of the polynomials, or their distance from the boundary of the cone. After providing a general complexity bound, we explore a few frequently studied special cases.

1d - ThA139 - 510A - **Moment, Polynomial and Tensor Optimization - Part 2** Chair: XINDONG TANG

Symmetric Tensor Decompositions and Gaussian Mixture Models – ZI YANG, UNIVERSITY AT ALBANY, SUNY – Jiawang Nie, Bingni Guo

This talk discusses how to recover parameters in diagonal Gaussian mixture models using tensor decompositions. High-order moments of the Gaussian mixture model are estimated from samples. They form incomplete symmetric tensors generated by hidden parameters in the model. We propose to use generating polynomials to compute incomplete symmetric tensor approximations. The obtained decomposition is utilized to recover parameters in the model. We prove that our recovered parameters are accurate when the estimated moments are accurate. Using high-order moments enables our algorithm to learn Gaussian mixtures with more components. For a given model dimension and order, we provide an upper bound of the number of components in the Gaussian mixture model that our algorithm can compute.

Moment Estimation of Nonparametric Mixtures Through Implicit Tensor Decomposition – JOE KILEEL, UT AUSTIN –

I will present computational methods to estimate certain statistical models from data samples. The models are conditionally-independent mixture models, and cover a wide range of applications. The optimization methods are based on tensor decomposition and completion, but crucially they avoid the explicit formation of tensors which enables scalability to high dimensions. Time permitting, I'll also discuss the algebraic varieties

underlying the problem. This is based on joint works with Yifan Zhang, with Joao Pereira and Tammy Kolda, and with Yulia Alexandr and Bernd Sturmfels.

Spectral Methods for Polynomial Optimization – ELVIRA MORENO, CALTECH – Venkat Chandrasekaran

We describe computationally tractable relaxations based on spectral methods for polynomial optimization. Concretely, our relaxations yield bounds based on minimum eigenvalues of matrices that are obtained as linear functions of the problem data. Spectral methods are computationally much less expensive than semidefinite programs, which makes them an attractive candidate for obtaining bounds on large-scale polynomial optimization problems. We identify the algebraic structure underlying a polynomial optimization problem that makes it amenable to spectral methods, and we demonstrate the efficiency and applicability of our framework with numerical examples.

1d - ThA214 - 510D - **On SDP relaxations for combinatorial optimization problems** Chair: MANISH KRISHAN LAL

Exploiting low-rank SDP methods for solving Max-Cut – JAN SCHWIDDESSEN, UNIVERSITY OF KLAGENFURT – Valentin Durante

We present an exact solver for the Max-Cut problem, exploiting recent advances in solving diagonally constrained semidefinite programs. The solver uses a branch-and-cut paradigm and strengthens the basic semidefinite programming relaxation by adding clique inequalities. The cutting planes are handled via Lagrangian relaxation and the inner semidefinite programming problem is approximately solved by the Mixing method, an efficient first-order and low-rank algorithm.

We propose a new branching rule that exploits both primal and dual information, resulting in fewer explored subproblems. Moreover, we reduce the number of cutting-plane iterations by effectively handing down the cuts in the search tree. We demonstrate our results against state-of-the-art algorithms by means of an in-depth study on numerous dense Max-Cut instances. It turns out that our solver is not only faster than other existing approaches, but also scales better for larger instances.

Using the ALPS framework to run the BiqCrunch binary quadratic solver on a computer cluster – NATHAN KRISLOCK, NORTHERN ILLINOIS UNIVERSITY – James Haas

BiqCrunch is a branch-and-bound solver using semidefinite optimization to compute high-quality bounds for combinatorial optimization problems that can be modeled as binary quadratic problems, such as MaxCut, Max-k-Cluster, Maximum-Independent-Set, Exact Quadratic Knapsack, and the Quadratic Assignment Problem.

BiqCrunch does not use an interior-point method for computing its bounds. Instead, an eigenvalue computation and a gradient-based method are used to compute the semidefinite bounds.

We will discuss our work on using the ALPS framework to run BiqCrunch with MPI on the HPC computer cluster at Northern Illinois University, and present numerical results.

Combinatoric Derivations and Sidorenko's Conjecture – DANIEL BROSCHE, UNIVERSITY OF KLAGENFURT –

Sidorenko's conjecture can be formulated as "Let H be a bipartite graph, and $\rho \in [0, 1]$. Of all the graphs with edge density ρ , the graph(-limit) obtained by picking edges uniformly at random minimizes the homomorphism density of H ."

This conjecture, first formulated in 1991 by Sidorenko, has received considerable attention over the last decades, and yet remains open in the general case. It was shown recently [Blekherman, Raymond, Singh, Thomas, 2020] that sums-of-squares in Razborov's flag algebra are not strong enough to prove even small, known cases of the conjecture. To circumvent this, we introduce a novel kind of (Lie-)derivation of flags. Due to their combinatoric nature, we can use them to systematically gain knowledge on global minimizers of problems in extremal graph theory. We combine them with the flag algebra method to find new proofs for various cases of Sidorenko's conjecture.

Sums of squares and Sidorenko's conjecture – ANNIE RAYMOND, UNIVERSITY OF MASSACHUSETTS, AMHERST – Pranav Garg, Amanda Redlich

A famous problem in extremal graph theory is Sidorenko's conjecture which states that for a fixed bipartite graph H , the density of H in a graph of a given edge density is asymptotically minimized by a random graph. In this talk, we discuss some strengths and limitations of the sums of squares method to prove this conjecture for various families of bipartite graphs. We also present some computational results.

1d - ThB123 - 510D - **Online optimization and learning in quantum games** Chair: ANTONIOS VARVITSIOTIS

A Semi-Definite Programming Approach for the Single-Row Facility Layout Problem with Multiple Objectives – CHRISTOF BRANDSTETTER, JOHANNES KEPLER UNIVERSITY – Arezoo Amiri, Elisabeth Gaar, Sophie Parragh, Markus Sinnl

In this work, we combine ideas from the fields of Semi-Definite Programming (SDP) and Multi-Objective Programming (MOP) to compute the Pareto frontier of combinatorial optimization (CO) problems. Our focus is on the Single-Row Facility Layout Problem (SRFLP). The SRFLP is a well-known NP-hard CO problem in the field of industrial engineering. The problem consists of finding the optimal arrangement of facilities along a single row, such that the total cost is minimized. The SRFLP is frequently extended to a multi-objective optimization problem, to simultaneously consider conflicting objectives, such as flow distance, CO2 emissions or total material handling cost. We propose a novel approach that leverages the power of SDP to solve the SRFLP with multiple objectives. Specifically, our focus is to combine objective space techniques, such as the weighted sum method and the epsilon-constraint method, with SDP. Solution algorithms based on these techniques are used in a computational study to compare our novel approach with objective space methods using integer programming formulations. For this comparison, we consider an adapted set of benchmark instances from the literature.

Semidefinite Network Games – CONSTANTIN ICKSTADT, GOETHE UNIVERSITY FRANKFURT, GERMANY – Thorsten Theobald, Elias Tsigaridas, Antonios Varvitsiotis

Network games provide a powerful framework for modeling agent interactions in networked systems, where players are represented by nodes in a graph, and their payoffs depend on the actions taken by their neighbors. Extending the framework of network games, in this work we introduce and study semidefinite network games. In this model, each player selects a positive semidefinite matrix with trace equal to one, known as a density matrix, to engage in a two-player game with every neighboring node. The player's payoff is the cumulative payoff acquired from these edge games. Network semidefinite games are of interest because they provide a simplified framework for representing quantum strategic interactions. Initially, we focus on the zero-sum setting, where the sum of all players' payoffs is equal to zero. We establish that in this class of games, Nash equilibria can be characterized as the projection of a spectrahedron. Furthermore, we demonstrate that determining whether a game is a semidefinite network game is equivalent to deciding if the value of a semidefinite program is zero. Beyond the zero-sum case, we characterize Nash equilibria as the solutions of a semidefinite linear complementarity problem.

Learning in Quantum Potential Games – WAYNE LIN, SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN –

Quantum potential games and quantum common-interest games model quantum agents acting cooperatively to maximize a common utility function. In this presentation we shall introduce the setting of quantum common-interest games where players have density matrices as strategies and seek to maximize a common multi-linear utility function, which is a setting that generalizes classical common-interest games. We shall then study various decoupled learning dynamics for learning in such games and explore their properties. Decoupled learning dynamics for quantum common-interest games can be used as algorithms for linear optimization over the set of separable states, which is otherwise known as the Best Separable State problem.

A Quadratic Speedup in Finding Nash Equilibria of Quantum Zero-Sum Games – EMMANOUIL VASILEIOS VLATAKIS GKARAGKOUNIS, UC BERKELEY – Francisca Vasconcelos

Recent developments in domains such as non-local games, quantum interactive proofs, and quantum generative adversarial networks have renewed interest in quantum game theory and, specifically, quantum zero-sum games. Central to classical game theory is the efficient algorithmic computation of Nash equilibria, which represent optimal strategies for both players. In 2008, Jain and Watrous proposed the first classical algorithm for computing equilibria in quantum zero-sum games using the Matrix Multiplicative Weight Updates (MMWU) method to achieve a convergence rate of $O(d/\epsilon^2)$ iterations to ϵ -Nash equilibria in the 4^d -dimensional spectraplex. In this work, we propose a hierarchy of quantum optimization algorithms that generalize MMWU via an extra-gradient mechanism. Notably, within this proposed hierarchy, we introduce the Optimistic Matrix Multiplicative Weights Update (OMMWU) algorithm and establish its average-iterate convergence complex-

ity as $O(d/\epsilon)$ iterations to ϵ -Nash equilibria. This quadratic speed-up relative to Jain and Watrous' original algorithm sets a new benchmark for computing ϵ -Nash equilibria in quantum zero-sum games.

1d - ThB910 - 510A - **Semi-Definite Programming 3** Chair: TYLER WEAMES 

Superlinear convergence of an interior point algorithm on linear semi-definite feasibility problems – CHEE KHIAN SIM, UNIVERSITY OF PORTSMOUTH –

Linear semi-definite feasibility problems (LSDFPs) are semi-definite programs where either the primal objective function or the dual objective function is zero. Solving linear semi-definite feasibility problems can be applied to find solutions to linear matrix inequalities that arise in various application areas in engineering for example. In this talk, I will focus on an implementable interior point algorithm that solves an LSDFP using known search directions, namely, HKM and NT search directions, found in existing interior point solvers on semi-definite programs, such as, SeDuMi. I consider the local superlinear convergence behavior of the interior point algorithm on LSDFPs, and present a new result that advances what is known in the interior point method literature on this topic. The key that makes this result possible is to consider the interior point algorithm on the homogeneous feasibility model of the LSDFP, instead of the LSDFP itself.

From SDPs to polynomial equations: validating numerical solutions for semidefinite programming. – JEFERSON ZAPATA, INSTITUTE OF SCIENCE AND TECHNOLOGY OF AUSTRIA – Vladimir Kolmogorov, Simone Naldi

We investigate the following question: Can we verify the feasibility of a given SDP by utilizing an approximate solution that is sufficiently close to an exact solution? Traditional approaches often rely on the existence of rational feasible solutions, employing techniques like rounding and lattice reduction algorithms. However, we propose an alternative method that circumvents this assumption. Specifically, we demonstrate how to construct a system of polynomial equations, guaranteeing that its set of real solutions contains an isolated correct solution (assuming the target exact solution is maximum-rank). This approach enables the utilization of algorithms from real algebraic geometry to solve polynomial equation systems, resulting in a hybrid (or symbolic-numerical) method for SDPs. Notably, this hybrid method has proven effective in certifying the feasibility of numerous SDPs in the instances we tested. We posit that our approach may find applications beyond SDP feasibility certification, such as refining an approximate solution through the application of Newton's method to a system of polynomial equations.

Term-sparse polynomial optimization for the design of frame structures – MAROUAN HANDA, POSTDOC –

This work investigates efficient solution to two fundamental problems in topology optimization of frame structures. The first one involves minimizing structural compliance under linear-elastic equilibrium and weight constraint, while the second one minimizes the weight under compliance constraints.

These problems are non-convex and generally challenging to solve globally, with the non-convexity concentrated in a polynomial matrix inequality. Recently, these problems were tackled using the moment-sum-of-squares (mSOS) hierarchy. However, only smaller instances can be solved globally. Here, we aim to improve the scalability of solution to these problems by using the mSOS hierarchy supplemented with the Term Sparsity Pattern technique (TSP). Due to the unique polynomial structure of our problems in which the objective and constraint functions are separable polynomials, we further improve scalability by adopting a reduced monomial basis containing non-mixed terms only. From extensive numerical testing, we conclude that these techniques allow for a global solution to larger instances and accelerate the solution of the problems significantly.

Nonsmooth Newton Methods for Solving the Best Approximation Problem; with Applications to Linear Programming – TYLER WEAMES, UNIVERSITY OF WATERLOO –

In this presentation, we study the effects of applying a modified Levenberg-Marquardt regularization to a nonsmooth Newton method. We expand this application to exact and inexact nonsmooth Newton methods and apply it to the best approximation constrained to a polyhedral set problem. We also demonstrate that linear programs can be represented as a best approximation problem, extending the application of nonsmooth Newton methods to linear programming. This application provides us with insight into an external path following algorithm that, like the simplex method, takes a finite number of steps on the boundary of the polyhedral set. However, unlike the simplex method, these steps do not use basic feasible solutions.

1d - FA17 - 510A - **Recent advances in conic optimization - Part 2** Chair: BRUNO LOURENÇO 

General Spectral Constrained Matrix Optimization – CASEY GARNER, UNIVERSITY OF MINNESOTA – Shuzhong Zhang, Gilad Lerman

Matrix optimization has proven to be extremely important in applications ranging across the human experience. Two sides exist to matrices: the coordinate side and the spectral side. In this talk, we shall present our latest work expanding the ability to solve optimization problems with general constraints on both the eigenvalues or singular values of a matrix and its coordinates.

Non-facial Exposedness of Copositive Cones over Symmetric Cones – MITSUHIRO NISHIJIMA, TOKYO INSTITUTE OF TECHNOLOGY / THE INSTITUTE OF STATISTICAL MATHEMATICS – Bruno Lourenço

In this talk, we consider the cone of linear mappings whose associated quadratic forms are nonnegative over symmetric cones, referred to as copositive cones over symmetric cones. We show that copositive cones are never facially exposed when the underlying symmetric cone has dimension at least 2. To achieve this, we explicitly exhibit a non-exposed extreme ray. Our result extends the known fact that the cone of copositive matrices in the conventional sense is generally not facially exposed.

Efficient Quantum Relative Entropy Programming with SDP Preprocessing – MEHDI KARIMI,

ILLINOIS STATE UNIVERSITY – Levent Tuncel

Optimization over the quantum relative entropy (QRE) cone has many applications in quantum information processing, such as calculating the key rates for quantum key distribution (QKD) protocols. Recently, a new version of our convex optimization software package Domain-Driven Solver (DDS) was released with modified numerical approaches for solving QRE programming problems, potentially combined with many other convex function/set constraints. In this talk, we propose a preprocessing and two-phase approach for QRE programming to reduce the size and improve the conditioning of a given instance. Phase-I of this two-phase approach is a well-behaved convex optimization problem, such as minimizing a self-concordant barrier or Semi-Definite Programming. The reformulated QRE programming problem can be solved more efficiently and faster by DDS or other optimization algorithms. We finish the talk by presenting numerical results of using DDS to solve many classes of problems, including calculating the QKD rate for different protocols.

Convergence rate analysis of a Dykstra-type projection algorithm – XIAOZHOU WANG, THE HONG KONG POLYTECHNIC UNIVERSITY – Ting Kei Pong

Given closed convex sets C_i , $i = 1, \dots, \ell$, and some nonzero linear maps A_i , $i = 1, \dots, \ell$, of suitable dimensions, the multi-set split feasibility problem aims at finding a point in $\bigcap_{i=1}^{\ell} A_i^{-1}C_i$ based on computing projections onto C_i and multiplications by A_i and A_i^T . In this paper, we consider the associated best approximation problem, i.e., the problem of computing projections onto $\bigcap_{i=1}^{\ell} A_i^{-1}C_i$; we refer to this problem as the best approximation problem in multi-set split feasibility settings (BA-MSF). We adapt the Dykstra's projection algorithm, which is classical for solving the BA-MSF in the special case when all $A_i = I$, to solve the general BA-MSF. Our Dykstra-type projection algorithm is derived by applying (proximal) coordinate gradient descent to the Lagrange dual problem, and it only requires computing projections onto C_i and multiplications by A_i and A_i^T in each iteration. Under a standard relative interior condition and a genericity assumption on the point we need to project, we show that the dual objective satisfies the Kurdyka-Łojasiewicz property with an explicitly computable exponent on a neighborhood of the (typically unbounded) dual solution set when each C_i is $C^{1,\alpha}$ -cone reducible for some $\alpha \in (0, 1]$: this class of sets covers the class of C^2 -cone reducible sets, which include all polyhedrons, second-order cone, and the cone of positive semidefinite matrices as special cases. Using this, explicit convergence rate (linear or sublinear) of the sequence generated by the Dykstra-type projection algorithm is derived. Concrete examples are constructed to illustrate the necessity of some of our assumptions.

1d - FA89 - 510D - **Recent advances in quadratic and semidefinite optimization**
Chair: FRANK PERMENTER

Rounding semidefinite relaxations of concave functions of quadratic forms – CHENYANG YUAN, TOYOTA RESEARCH INSTITUTE –

We study semidefinite relaxations of the problem of maximizing a concave function of quadratic forms on the unit sphere. When this function is a fractional power $0 < p < 1$, we obtain

tight approximation factors by analyzing a randomized rounding procedure. When $p=1/2$, this problem and its relaxation is equivalent to Max-Cut, and the approximation guarantees we obtain matches the best known factors, and can be interpreted as obtaining a generalization of Nesterov's $2/\pi$ approximation result. When the solution of the semidefinite program has a low rank, our analysis obtains further improved bounds. We extend this analysis to other concave functions such as logarithms and entropy. Finally, we show that this rounding procedure and analysis can also be generalized to optimization on the Stiefel manifold.

Complexity of Chordal Conversion for Sparse Semidefinite Programs with Small Treewidth – RICHARD ZHANG, UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN –

If a sparse semidefinite program (SDP), specified over $n \times n$ matrices and subject to m linear constraints, has an aggregate sparsity graph G with small treewidth, then chordal conversion will frequently allow an interior-point method to solve the SDP in just $O(m+n)$ time per-iteration, which is a significant speedup over the $\Omega(n^3)$ time per-iteration for a direct application of the interior-point method. Unfortunately, the speedup is not guaranteed by a small treewidth in G , as a diagonal SDP would have treewidth zero but can still necessitate up to $\Omega(n^3)$ time per-iteration. Instead, we construct an extended aggregate sparsity graph $\bar{G} \supseteq G$ by forcing each constraint matrix A_i to be its own clique in G . We prove that a small treewidth in $\bar{G}$ does indeed guarantee that chordal conversion will solve the SDP in $O(m+n)$ time per-iteration, to ϵ -accuracy in at most $O(\sqrt{m+n} \log(1/\epsilon))$ iterations. This sufficient condition covers many successful applications of chordal conversion, including the MAX- k -CUT relaxation, the Lovász Theta problem, sensor network localization, polynomial optimization, and the AC optimal power flow relaxation, thus allowing theory to match practical experience.

A Semidefinite Hierarchy for Expected Independence Numbers of Random Graphs – KEVIN SHU, GEORGIA INSTITUTE OF TECHNOLOGY – Cifuentes Diego, Alejandro Toriello

We introduce convex optimization methods to find upper bounds on the expected independence number of a random graph. These upper bounds are inspired by the well known Lovász theta function, which bounds for the independence number of a deterministic graph using a semidefinite program. Specifically, we propose a hierarchy of semidefinite programs (SDPs) which depend on the probabilities that given subsets of vertices in the random graph are independent. These semidefinite programs have values which each upper bound the expected independence number of the random graph. We discuss cases in which this hierarchy offers a close approximation to the true expected independence number and cases in which it is far away from the true expected independence number. These semidefinite programming bounds are related to the usual Lasserre hierarchy for general boolean quadratic optimization. For symmetric random graphs, the last level of the hierarchy is equivalent to a linear program whose optimal value can often be calculated or approximated in closed form and which have close connections to combinatorics. We show that our methods provide good upper bounds in a number of examples, including Erdos-Rényi graphs, uniformly random spanning trees, and geometric random graphs. We also discuss

empirical results.

1d - FB36 - 510A - SDP and Exact Relaxations of Hard Problems Chair: YONG XIA

Forward-Backward Extended DMD with an Asymptotic Stability Constraint – LOUIS LORTIE, MCGILL UNIVERSITY –

This presentation introduces a data-driven method to identify an asymptotically stable Koopman system from noisy data. The Koopman operator allows a nonlinear system to be modeled in a linear fashion from data. However, systems identified with noisy data are modeled as biased and potentially unstable systems. The presented approach poses the forward- and backward-in-time Koopman operator approximation problem as a convex optimization problem with linear matrix inequality (LMI) constraints to enforce asymptotic stability and reduce the bias in the identified Koopman system. The presented method is validated against state-of-the-art methods using a simulated Duffing oscillator dataset and an experimental soft robot arm dataset.

Convex conic reformulation of geometric non-convex conic optimization problems and exact solutions of QCQPs by SDP relaxations – SUNYOUNG KIM, EWA W. UNIV. – Masakazu Kojima

A geometric nonconvex conic optimization problem was recently proposed as a unified framework for convex conic reformulation of a class of quadratic optimization problems and polynomial optimization problems. The nonconvex conic optimization problem minimizes a linear function over the intersection of a nonconvex cone K , a convex subcone J of the convex hull of K , $\text{co}K$, and an affine hyperplane with a normal vector. Under the assumption $\text{co}(K \cap J) = J$, the original nonconvex conic optimization problem was shown to be equivalently formulated as a convex conic program by replacing the constraint set with the intersection of J and the affine hyperplane. In this talk, we further investigate some remaining issues such as the key assumption $\text{co}(K \cap J) = J$ in the framework and provide three sets of necessary-sufficient conditions for the assumption. As an application, we propose a new wide class of quadratically constrained quadratic programs with multiple nonconvex equality and inequality constraints that can be solved exactly by their semidefinite programming (SDP) relaxation. The proposed class of quadratically constrained quadratic programs can have any finite number of nonconvex constraints, extending the previous results on the exact SDP relaxations. The condition necessary for the exactness is quite general.

On the tightness of SDP relaxation for quadratic assignment problem – SHUYANG LING, NYU SHANGHAI –

Quadratic assignment is a fundamental problem in combinatorial optimization and finds numerous applications in operation research, computer vision, etc. However, it is in general NP-hard problem to find the global minimizer to the quadratic assignment problem (QAP). In this work, we study the semidefinite relaxation (SDR) of the QAP and investigate when the SDR recovers the global minimizer. In particular, we consider the two input matrices satisfy certain signal-plus-noise model, and show that when the noise is sufficiently smaller than the signal, then the SDR is tight. It is worth noting that this sufficient condition is purely algebraic and does not depend on any

statistical assumption of the input data.

High-rank Solution of Sum-of-Squares Relaxations for Exact Matrix Completion – GODAI AZUMA, AOYAMA GAKUIN UNIVERSITY – Sunyoung Kim, Makoto Yamashita

We study rank-one matrix completion problems which estimate all the elements of a matrix from a partially given matrix. Although these problems can be formulated as quadratically constrained quadratic problems, their standard semidefinite relaxations generally are not tight. Recently, the Lasserre's relaxation of order 2 have been shown to output the exact completion when the completion is unique and the objective function is the trace of the moment matrix. In this talk, we show that when a certain matrix is used as the coefficients of the objective function instead of the trace, the exact completion can also be obtained from the dual of the Lasserre's relaxation, called the sum of squares (SOS) relaxation. This theoretical result is based on the existence of an optimal solution such that the Gram matrix in the SOS relaxation is of high-rank due to the sparsity structure of the coefficient matrix. We also propose an algorithm for finding this coefficient matrix and illustrate its performance with computer experiments. Furthermore, additional results including the polynomial optimization problem with a hidden matrix completion problems are discussed.

1d - FB275 - 510D - Recent Advancements in the Theory and Applications of Semidefinite Programming Chair: HAO HU

Navigating hidden convexity in nonconvex projection problems – MANISH KRISHAN LAL, UBC –

We study polynomials upto degree six arising in nonconvex projection problems. When finding nearest points onto certain sets of bilinear constraints, we discover these problems have hidden convexity. We establish this fact using SDP duality and exploit it further to develop convex relaxations to solve certain bilinear combinatorial optimization problems.

The Maximum Singularity Degree for Linear and Semidefinite Programming – HAO HU, CLEMSON UNIVERSITY –

The singularity degree plays a crucial role in understanding linear and semidefinite programming, providing a theoretical framework for analyzing these problems. It is defined as the minimum number of facial reduction (FR) steps needed to reach strict feasibility for a convex set. On the other hand, the maximum singularity degree (MSD) is the maximum number of steps required. Recent progress in the applications of MSD has motivated us to explore its fundamental properties.

For semidefinite programming, we establish a necessary condition for an FR sequence to be the longest. Additionally, we propose an upper bound for MSD, which can be computed more easily. By leveraging these findings, we prove that computing MSD is NP-hard. This complexity result complements the existing algorithms for computing the singularity degree found in the literature. For linear programming, we provide a characterization for the longest FR sequences, which also serves as a polynomial-time algorithm for constructing such a sequence. In addition, we introduce two operations that ensure the longest FR sequences remain the longest. Lastly, we prove

that MSD is equivalent to a novel parameter called the implicit problem singularity.

Ramana's dual redux – GABOR PATAKI, UNC CHAPEL HILL –

Ramana's dual is a fundamental result in semidefinite programming. In contrast to the usual SDP dual, 1) it requires no Slater like constraint qualification, 2) it always attains its optimal value, and 3) has no duality gap with the primal. It is mainly of interest due to its complexity implications, which – 25 years after its publication – are still the best known results on the complexity of SDP.

The correctness proofs of Ramana's dual tend to be somewhat technical. A dual with similar properties, but based on arguments from algebraic geometry has also been proposed. Here we give a nontechnical and very short proof, that is based almost purely on elementary linear algebra. In addition we prove a result that appears to be new: we completely characterize the set of feasible solutions of Ramana's dual – a result that may be useful for computational purposes.

A Peaceman-Rachford Splitting Method for the Protein Side-Chain Positioning Problem and other Hard Problems – HENRY WOLKOWICZ, UNIVERSITY OF WATERLOO – Forbes Burkowski, Haesol Im

This paper considers the NP-hard protein side-chain positioning problem, SCP, an important final task of protein structure prediction as a integer quadratic program, IQP. We derive its doubly nonnegative, DNN, (convex) relaxation. Strict feasibility fails for this relaxation and we apply facial reduction, FR as a natural splitting of variables. This allows for a variation of the application of Peaceman-Rachford splitting method, PRSM. The resulting relaxation and rounding procedures provide strong approximate solutions. Empirical evidence shows that almost all our instances of this NP-hard SCP problem, taken from the Protein Data Bank, are solved to provable optimality. We provide further applications to the Wasserstein Barycenter problem.

1e - MB312 - 513D - **Convex analysis and beyond** Chair: ADRIAN LEWIS

Polar convexity in finite dimensional Euclidean spaces – SHUBHANKAR BHATT, UNIVERSITY OF WESTERN ONTARIO – Hristo Sendov

Let $\hat{\mathbb{R}}^n$ be the one point compactification of $\mathbb{R}^n$ obtained by adding a point at infinity. We say that a subset $A \subseteq \hat{\mathbb{R}}^n$ is $\mathbf{u}$ -convex if for every pair of points $\mathbf{z}_1, \mathbf{z}_2 \in A$, the arc of the unique circle through $\mathbf{u}, \mathbf{z}_1$ and $\mathbf{z}_2$, from $\mathbf{z}_1$ to $\mathbf{z}_2$ and not containing $\mathbf{u}$, is contained in A . In this case, we call $\mathbf{u}$ a pole of A . When the pole $\mathbf{u}$ approaches infinity, $\mathbf{u}$ -convex sets become convex in the classical sense.

The notion of polar convexity in the complex plane has been used to analyze the behavior of critical points of polynomials. In this paper, we extend the notion to finite dimensional Euclidean spaces. The goal of this paper is to start building the theory of polar convexity and to show that the introduction of a pole creates a richer theory. For example, polar convexity enjoys a beautiful duality that does not exist in classical convexity. We formulate polar analogues of several classical results

of the alternatives, such as Gordon and Farkas' lemmas. Finally, we give a full description of the convex hull of finitely many points with respect to finitely many poles.

The uniqueness of Lyapunov rank among symmetric cones – MICHAEL ORLITZKY, UMBC –

The Lyapunov rank of a cone is the dimension of the Lie algebra of its automorphism group. It is invariant under linear isomorphism and in general not unique – two or more non-isomorphic cones can share the same Lyapunov rank. It is therefore not possible to use Lyapunov rank to identify cones in general. But perhaps that's not the full story. "Loewnerian" cones are defined as being isomorphic to the (symmetric) real positive-semidefinite cone and have applications in optimization if we can recognize them. Motivated by this, we ask a more limited question: among *symmetric* cones, are there any that can be uniquely identified (up to isomorphism) by their Lyapunov ranks?

Constructions of c-splitting potentials for multi-marginal monotone sets – SEDI BARTZ, UNIVERSITY OF MASSACHUSETTS LOWELL –

In the theory of optimal transport, Rockafellar type constructions of antiderivatives are known to be solutions of the dual problem. In multi-marginal optimal transport theory, it is still unknown how to construct explicit c-splitting potentials in general. We present the problem and discuss cases where solutions are known.

Exploring averages of finite sets in nonpositively curved metric spaces – ADRIAN LEWIS, CORNELL UNIVERSITY – Ariel Goodwin, Genaro Lopez, Adriana Nicolae

In various disciplines, data points live in nonpositively curved metric spaces - those in which triangles are thinner than their Euclidean analogues. In such spaces - also known as CAT(0) spaces, or as Hadamard spaces (when complete) - convex optimization techniques are particularly attractive, because, given any point in the space, the squared-distance to that point is a strongly convex function along geodesics. A common example of a nonpositively curved space is the space of positive-definite matrices with its standard affine-invariant metric - a Hadamard manifold. However, spaces that are not manifolds are also common, in models ranging from robotics to phylogenetics: an important example is the tree space of Billera-Holmes-Vogtmann (2001), which is a complex of Euclidean orthants. Nonpositively curved spaces support various natural notions of "averaging" for finite sets of points in the space. Unfortunately, among these, the convex hull operation is poorly understood (even for three points). Weighted means (or "barycenters"), on the other hand, are in principle computable, but although they have served as the motivation for extensive phylogenetic computation, existing general-purpose algorithms for weighted means outside the case of manifolds are very slow. In the light of these challenges, we discuss the tractability of weighted mean problems - both computing them and recognizing them - in some important nonpositively curved spaces.

1e - TA313 - 514A - **Algorithms for optimization, monotone inclusions and variational inequalities 1** Chair: RADU IOAN BOT

On a result by Baillon, Bruck, and Reich – HEINZ BAUSCHKE, UBC OKANAGAN –

It is well known that the iterates of an averaged nonexpansive mapping may only converge weakly to fixed point. A celebrated result by Baillon, Bruck, and Reich from 1978 yields strong convergence in the presence of linearity. In this paper, we extend this result to allow for flexible relaxation parameters. Examples are also provided to illustrate the results.

Based on joint work with Yuan Gao (UBC Okanagan).

On the Bredies-Chenchene-Lorenz-Naldi algorithm – SHAMBHAVI SINGH, UNIVERSITY OF BRITISH COLUMBIA OKANAGAN –

Monotone inclusion problems occur in many areas of optimization and variational analysis. Splitting methods, which utilize resolvents or proximal mappings of the underlying operators, are often applied to solve these problems. In 2022, Bredies, Chenchene, Lorenz, and Naldi introduced a new elegant algorithmic framework that encompasses various well known algorithms including Douglas-Rachford and Chambolle-Pock. They obtained powerful weak and strong convergence results, where the latter type relies on additional strong monotonicity assumptions. In this paper, we complement the analysis by Bredies et al. by relating the projections of the fixed point sets of the underlying operators that generate the (reduced and original) preconditioned proximal point sequences. We also obtain strong convergence results in the case of linear relations. Various examples are provided to illustrate the applicability of our results.

The proximal point algorithm without monotonicity – RUSSELL LUKE, GEORG AUGUST UNIVERSITY OF GOETTINGEN –

We study the proximal point algorithm in the setting in which the operator of interest is metrically subregular and satisfies a submonotonicity property. The latter can be viewed as a quantified weakening of the standard definition of a monotone operator. Our main result gives a condition under which, locally, the proximal point algorithm generates at least one sequence which is linearly convergent to a zero of the underlying operator. General properties of our notion of submonotonicity are also explored as well as connections to other concepts in the literature. This is joint work with Matthew Tam.

Fast continuous time methods for monotone equations – RADU IOAN BOT, UNIVERSITY OF VIENNA –

In this talk, we discuss continuous in time dynamics for the problem of approaching the set of zeros of a single-valued monotone and continuous operator V . Such problems are motivated by minimax convex-concave and, in particular, by convex optimization problems with linear constraints. We introduce a second-order dynamical system that combines a vanishing damping term with the time derivative of V along the trajectory, which can be seen as an analogous of the Hessian-driven damping in case the operator is originating from a potential. We show that these methods exhibit fast convergence rates for $\|V(z(t))\|$ as $t \rightarrow +\infty$, where $z(\cdot)$ denotes the generated trajectory, and for the restricted gap function, and that $z(\cdot)$ converges to a zero of the operator V . For the corresponding implicit and explicit discrete time models with Nesterov's momentum, we prove that they share the asymptotic features of the continuous dynamics.

In addition, we discuss the connection between the second-order dynamical system and a Tikhonov regularized first-order dynamical system, exhibiting fast convergence rates and strong convergence of the trajectory.

1e - TB311 - 514A - New insights into first- and second-order variational analysis Chair: TIM HOHEISEL

Quasidensity and certain classes of multifunctions – STEPHEN SIMONS, UC SANTA BARBARA –

Let E be a nonzero real Banach space. *Quasidensity* is a concept that can be applied to subsets of $E \times E^*$ (or equivalently to multifunctions from E into E^*). Every closed quasidense monotone set is *maximally monotone*, but there exist maximally monotone sets that are not quasidense. The subdifferential of a proper, convex lower semicontinuous function on E is quasidense. The subdifferentials of certain *nonconvex* functions are also quasidense. (This follows from joint work with Xianfu Wang.) The closed monotone quasidense sets satisfy a *sum theorem* and a *dual sum theorem*. A maximally monotone multifunction is quasidense if, and only if, it is of type (NI). We will explore the relationship with the Fitzpatrick extension and other subclasses of the maximally monotone multifunctions as time permits.

Kurdyka-Lojasiewicz exponent for a class of Hadamard-difference-parameterized models – TING KEI PONG, HONG KONG POLYTECHNIC UNIVERSITY –

Wenqing Ouyang, Yuncheng Liu, Hao Wang
In this talk, we consider a class of L1-regularized optimization problems and the associated smooth “over-parameterized” optimization problems built upon the Hadamard difference parametrization (HDP). We show that second-order stationary points of the HDP-based model correspond to some stationary points of the corresponding L1-regularized model. More importantly, we show that the Kurdyka-Lojasiewicz (KL) exponent of the HDP-based model at a second-order stationary point can be inferred from that of the corresponding L1-regularized model under suitable assumptions. Our assumptions are general enough to cover a wide variety of loss functions commonly used in L1-regularizations, such as the least squares loss function and the logistic loss function. We also discuss how these KL exponents can help deduce the local convergence rate of a standard gradient method for minimizing the HDP-based models.

On second-order variational analysis of prox-regular functions – HELMUT GFRERER, JOHANNES KEPLER UNIVERSITY LINZ –

In order to characterize certain properties of prox-regular functions, the usual generalized derivatives of variational analysis for the subdifferential mapping can be applied only in the presence of subdifferential continuity. In this talk we will demonstrate how the definitions of these derivatives can be modified in order to be also useful for the characterization of (strong) variational convexity and tilt stability without subdifferential continuity. Further, we present a new characterization of strictly twice epi-differentiable functions shedding more light on the local behavior of the subdifferential mapping of such functions. As a consequence of this characterization we are able to show that for strictly twice epi-differentiable functions the property of metric regularity is equivalent to the one

of strong metric regularity. Our approach is mainly based on first-order theory of the subdifferential but we will also present the relation to Rockafellar's quadratic bundle consisting of the epigraphical limit of second subderivatives. Finally, for twice epi-differentiable functions, we give a relation between function values and subgradients which constitutes some special kind of second-order approximation. This result differs from other ones in the literature in so far as it gives some second-order expansion for the function values itself and not for the Moreau envelope.

Stability of nonsmooth optimization problems

– TIM HOHEISEL, MCGILL UNIVERSITY – Simone Brugiaglia, Aaron Berk

We present some applications of implicit function theorems from variational analysis to stability analysis of prominent optimization problems with a focus on regularized least-squares, including the LASSO or nuclear norm minimization. These contain results on well-posedness, Lipschitz stability and smoothness of the optimal solution function. The quantitative results are showcased in numerical computations.

1e - TB314 - 514B - **Algorithms for optimization, monotone inclusions and variational inequalities 2** Chair: ZI XU

A projected semismooth Newton method for a class of nonconvex composite programs with strong prox-regularity

– JIAYUAN WU, PEKING UNIVERSITY –

This paper aims to develop a Newton-type method to solve a class of nonconvex composite programs. In particular, the nonsmooth part is possibly nonconvex. To tackle the nonconvexity, we develop a notion of strong prox-regularity which is related to the singleton property and Lipschitz continuity of the associated proximal operator, and we verify it in various classes of functions, including weakly convex functions, indicator functions of proximally smooth sets, and two specific sphere-related nonconvex nonsmooth functions. In this case, the problem class we are concerned with covers smooth optimization problems on manifold and certain composite optimization problems on manifold. For the latter, the proposed algorithm is the first second-order type method. Combining with the semismoothness of the proximal operator, we design a projected semismooth Newton method to find a root of the natural residual induced by the proximal gradient method. Due to the possible nonconvexity of the feasible domain, an extra projection is added to the usual semismooth Newton step and new criteria are proposed for the switching between the projected semismooth Newton step and the proximal step. The global convergence is then established under the strong prox-regularity. Based on the BD regularity condition, we establish local superlinear convergence. Numerical experiments demonstrate the effectiveness of our proposed method compared with state-of-the-art ones.

Accelerated gradient descent: A guaranteed bound for a heuristic restart strategy

– VIKTOR PAVLOVIC, UNIVERSITY OF WATERLOO – Stephen Vavasis, Walaa Moursi

The $O(1/k^2)$ convergence rate in function value of accelerated gradient descent is optimal, but there are many modi-

fications that have been used to speed up convergence in practice. Among these modifications are restarts, that is, starting the algorithm with the current iteration being considered as the initial point. We focus on the adaptive restart techniques introduced by O'Donoghue and Candes, specifically their gradient restart strategy. While the gradient restart strategy is a heuristic in general, we prove that applying gradient restarts preserves and in fact improves the $O(1/k^2)$ bound, hence establishing function value convergence, for one-dimensional functions. Applications of our results to separable and nearly separable functions are presented.

Forward-Backward algorithms for weakly convex problems

– EWA BEDNARCZUK, SYSTEMS RESEARCH INSTITUTE, POLISH ACADEMY OF SCIENCES –

We investigate convergence properties of exact and inexact forward-backward algorithms for the minimisation of the sum of weakly convex functions defined on a Hilbert space, where one has a Lipschitz-continuous gradient. We show that the exact forward-backward algorithm converges strongly to a global solution, provided that the objective function satisfies a sharpness condition.

For the inexact forward-backward algorithm, the same assumptions ensure that the distance from the iterates to the solution set approaches a positive threshold depending on the accuracy level of the proximal computations.

As an application of the considered setting, we provide numerical experiments related to feasibility problems, source localisation and discrete tomography. The talk is based on a joint work with Giovanni Brucola, Gabriele Scrivanti and The Hung Tran

Zeroth order and first order primal dual alternating proximal gradient algorithms for nonsmooth nonconvex minimax problems with coupled linear constraints

– ZI XU, SHANGHAI UNIVERSITY –

Nonconvex minimax problems have attracted wide attention in machine learning, signal processing and many other fields in recent years. In this talk, we propose a zeroth-order and a first order primal dual alternating proximal gradient (PDAPG) algorithm and a primal dual proximal gradient (PDPG-L) algorithm for solving nonsmooth nonconvex-(strongly) concave and nonconvex-linear minimax problems with coupled linear constraints, respectively. The iteration complexity of the two algorithms also have been proved to reach an ϵ -stationary point. To our knowledge, they are the first two algorithms with iteration complexity guarantee for solving the nonconvex minimax problems with coupled linear constraints.

1e - TC911 - 514A - **Variational Analysis, Variational Inequalities and Complementarity 1** Chair: JEIN SHAN CHEN

Zero duality gap for optimization problems in spaces without linear structure

– MONIKA SYGA, FACULTY OF MATHEMATICS AND INFORMATION SCIENCES, WARSAW UNIVERSITY OF TECHNOLOGY – Ewa Bednarczuk

We prove sufficient and necessary conditions ensuring zero duality gap for Lagrangian duality in some classes of nonconvex

optimization problems. To this aim, we use the ϕ -convexity theory and minimax theorems for ϕ -convex functions. The obtained zero duality results apply to infinite-dimensional linear optimization problems, including Kantorovich duality which plays an important role in determining Wasserstein distance.

Accelerating Communications in Decentralized Learning via Randomization – EDOUARD OYALLON, FLATIRON INSTITUTE – Adel Nabli, Eugene Belilovsky

This talk explores DADAO and A2CID2, two recent methods designed to address communication bottlenecks in distributed decentralized learning. DADAO stands out as the first decentralized, accelerated, primal, asynchronous optimization algorithm to minimize a sum of smooth and strongly convex functions. Meanwhile, A2CID2 utilizes continuous local momentum and randomized peer-to-peer operations to accelerate communications. These methods model operations within a decentralized framework using Stochastic Differential Equations, providing a robust theoretical basis while being straightforward to implement. By decoupling computation from communication, they effectively reduce communication demands and enhance scalability, which is amenable to large-scale deep learning applications. During the talk, I will examine the fundamental principles of these methods and critically assess their limitations compared to traditional stochastic synchronous settings.

On some new NCP functions – JEIN SHAN CHEN, NATIONAL TAIWAN NORMAL UNIVERSITY – A nonlinear complementarity problem (NCP) consists of nonlinear inequalities with nonnegativity and orthogonality conditions on multiple variables and given multivariate functions. Such problems arise in constrained optimization and equilibrium problems. Moreover, applications of NCPs in operations research, engineering, and economics motivated significant research efforts in the past decades, which resulted in various numerical techniques including the merit function approach, nonsmooth Newton method, and regularization approach.

In this talk, we present two classes of new NCP functions with their properties. In particular, we propose neural networks and inexact Levenberg-Marquardt method by using these NCP functions for solving nonlinear complementarity problems (NCPs). More specifically, for the first class of NCP functions, we build a new class of neural networks based on the smoothing method for NCP introduced by Hadou and Maheux. For the second class of NCP functions, which are smooth, coercive and strongly semismooth, we propose an inexact Levenberg-Marquardt method for solving nonlinear complementarity problems (NCPs). Different from existing exact/inexact Levenberg-Marquardt methods, the proposed method adopts a derivative-free line search to ensure its globalization.

In addition, we present numerical experiments to validate the theoretical results and to illustrate the difference in numerical performance of functions. Numerical performance and comparison are reported and demonstrated.

1e - WA338 - 514A - **Algorithms for optimization, monotone inclusions and variational inequalities** 3 Chair: WALAA MOURSI ←

Exotic error bounds, Karamata theory and convergence rate analysis – BRUNO LOURENÇO, INSTITUTE OF STATISTICAL MATHEMATICS – Tianxiang Liu

In this talk we overview recent results on convergence rates of algorithms for convex feasibility problems. Our main tools will be the framework of consistent error bound functions and Karamata theory. Karamata theory is a classical but relatively unknown tool inside the optimization community and it is quite helpful to analyze asymptotic behavior of functions. With that, we show new concrete convergence rate results for problems satisfying exotic non-Holderian error bounds such as logarithmic and entropic error bounds. These exotic error bounds may appear, for instance, in problems involving exponential cones and other non-semialgebraic objects. Time allowing, we will discuss our recent efforts in extending our techniques to common fixed point problems.

DEFBAL – a connection between the ADMM and forward-backward methods – JONATHAN ECKSTEIN, RUTGERS UNIVERSITY – Chang Yu, Felipe Atenas, Paulo J. S. Silva

This talk will demonstrate that the two minimization steps of the standard ADMM algorithm are equivalent to applying the standard forward-backward / proximal gradient method to a dual formulation of the augmented Lagrangian subproblem (this dual formulation always has the Lipschitz properties usually assumed in showing convergence of forward-backward methods). This insight allows the construction of a large family of ADMM-like methods by combining a range of proximal-gradient-like methods – for example, those involving Nesterov-class acceleration – with a variety of inexact augmented Lagrangian algorithms. We call this class of procedures “DEFBAL”, standing for “Dual Embedded Forward-Backward Augmented Lagrangian”. We give some possible applications and preliminary computational results, although any computational advantages are unclear at present.

The Douglas–Rachford algorithm for inconsistent problems – BETHANY CALDWELL, UNIVERSITY OF SOUTH AUSTRALIA – Regina Burachik, Yalçın Kaya, Walaa Moursi

The Douglas–Rachford algorithm has been successfully employed to solve convex optimization problems, or more generally find zeros of monotone inclusions. Recently, the behaviour of these methods in the inconsistent case, i.e., in the absence of solutions, has been a topic of interest as these problems appear regularly but can be difficult to analyze. We give a result for the strong convergence of the shadow sequence of the Douglas–Rachford algorithm in the possibly inconsistent case when one of the operators is uniformly monotone and 3* monotone but not necessarily a subdifferential. We also provide numerical experiments comparing the performance of the Douglas–Rachford algorithm for inconsistent finite- and infinite-dimensional linear-quadratic problems with the Peaceman–Rachford algorithm and Dykstra’s projection algorithm.

The Chambolle-Pock algorithm revisited: splitting operator and its range with applications – WALAA MOURSI, UNIVERSITY OF WATERLOO –

Primal-dual hybrid gradient (PDHG) is a first-order method

for saddle-point problems and convex programming introduced by Chambolle and Pock. Recently, Applegate et al. analyzed the behavior of PDHG when applied to an infeasible or unbounded instance of linear programming, and in particular, showed that PDHG is able to diagnose these conditions. Their analysis hinges on the notion of the infimal displacement vector in the closure of the range of the displacement mapping of the splitting operator that encodes the PDHG algorithm. In this talk, we develop a novel formula for this range using monotone operator theory. The analysis is then specialized to conic programming and further to quadratic programming (QP) and second-order cone programming (SOCP). A consequence of our analysis is that PDHG is able to diagnose infeasible or unbounded instances of QP and of the ellipsoid-separation problem, a subclass of SOCP.

1e - WB227 - 514A - **Algorithms for optimization, monotone inclusions and variational inequalities** 4 Chair: JUAN PABLO CONTRERAS [←](#)

Single-loop stochastic variance reduced method for variational inequalities – AFROOZ AFROOZ JALILZADEH, UNIVERSITY OF ARIZONA – In this talk, we consider the problem of variational inequalities (VIs). Within the class of monotone VIs with a finite-sum structure, we propose a novel single-loop stochastic variance-reduced algorithm equipped with the Bregman distance function. We demonstrate an optimal convergence guarantee for the proposed method. Furthermore, we investigate a structured class of non-monotone problems exhibiting weak Minty solutions and discuss the convergence rate of our proposed method. Numerical experiments will be presented to showcase the performance of the proposed algorithm compared with state-of-the-art methods.

Tikhonov regularized exterior penalty dynamics for constrained variational inequalities – SIQI QU, UNIVERSITY OF MANNHEIM – Mathias Staudigl

Solving equilibrium problems under constraints is an important problem in optimization and optimal control. In this context an important practical challenge is the efficient incorporation of constraints. We develop a continuous-time method for solving constrained variational inequalities based on a new penalty regulated dynamical system in a general potentially infinite-dimensional Hilbert space. In order to obtain strong convergence of the issued trajectory of our method, we incorporate an explicit Tikhonov regularization parameter in our method, leading to a class of time-varying monotone inclusion problems featuring multiscale aspects. Besides strong convergence, we illustrate the practical efficiency of our developed method in solving constrained min-max problems.

Synchronous and asynchronous gradient-response schemes for computing quasi-Nash equilibria under uncertainty – ZHUOYU XIAO, PENNSYLVANIA STATE UNIVERSITY – Uday Shanbhag In this work, we propose synchronous and asynchronous gradient-response schemes to solve stochastic nonconvex games. Our theoretical results show that under some reasonably mild conditions, both schemes generate sequences which converge to quasi-Nash equilibria almost surely. We further

explore some special cases for which we can obtain rate statements. Both the convergence and rate claims appear to be novel in the context of computing quasi-Nash equilibria. Numerical experiments are also provided. To our best knowledge, this represents the first the first work on gradient-response schemes for stochastic nonconvex games.

Stochastic Halpern iteration in non-Euclidean spaces and applications to synchronous Q-learning – JUAN PABLO CONTRERAS, PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE –

We analyze the overall oracle complexity of the stochastic Halpern iteration with minibatch, aiming to approximate fixed points of nonexpansive and contractive operators in a finite-dimensional space with a non-Euclidean norm. We outline conditions for almost sure convergence of the iterates and demonstrate that, under mild assumptions such as uniformly bounded variance, the expected fixed-point residual decays at a rate of $\tilde{O}(1/n)$. When approximating a fixed point within a tolerance $\varepsilon > 0$, our method exhibits an overall oracle complexity of $\tilde{O}(\varepsilon^{-5})$, surpassing the $\tilde{O}(\varepsilon^{-6})$ rate of the stochastic Krasnosel'skii-Mann iteration. Furthermore, we establish a lower bound of $\tilde{O}(\varepsilon^{-3})$, which applies to a wide range of algorithms, including all averaged iterations like Krasnosel'skii-Mann and Halpern. As an application, we propose new algorithms for synchronous Q-learning with average and discounted rewards. In particular, for average reward, our method improves upon the best-known sample complexity in the literature.

1e - WB329 - 515C - **Conic optimization** Chair: JAKUB MARECEK [←](#)

Extensions of constant rank qualification constraint condition to nonlinear conic programming – HECTOR RAMIREZ, UNIVERSIDAD DE CHILE –

We present new constraint qualification conditions for nonlinear conic programming that extend some of the constant rank-type conditions from nonlinear programming. Specifically, we propose a general and geometric approach, based on the study of the faces of the cone, for defining a new extension of this condition to the conic context. We then compare these new conditions with some of the existing ones, including the non-degeneracy condition, Robinson's constraint qualification, and the metric subregularity constraint qualification. The main advantage of the latter is that we are able to recast the strong second-order properties of the constant rank condition in a conic context. In particular, we obtain a second-order necessary optimality condition that is stronger than the classical one obtained under Robinson's constraint qualification, in the sense that it holds for every Lagrange multiplier, even though our condition is independent of Robinson's condition.

Learning quantum Hamiltonians at any temperature in polynomial time with Chebyshev and bit complexity – ALES WODECKI, CZECH TECHNICAL UNIVERSITY, PRAGUE – Jakub Marecek We consider the problem of learning local quantum Hamiltonians given copies of their Gibbs state at a known inverse temperature, following Haah et al. [2108.04842] and Bakshi et al. [arXiv:2310.02243]. Our main technical contribution is a new flat polynomial approximation of the exponential function

based on the Chebyshev expansion, which enables the formulation of learning quantum Hamiltonians as a polynomial optimization problem. This, in turn, can benefit from the use of moment/SOS relaxations, whose polynomial bit complexity requires careful analysis [O'Donnell, ITCS 2017]. Finally, we show that learning a k -local Hamiltonian, whose dual interaction graph is of bounded degree, runs in polynomial time under mild assumptions.

Time-varying semidefinite programming: path-following a Burer-Monteiro factorization – ANTONIO BELLON, UNIVERSITÄT AUGSBURG –

We present an online algorithm for time-varying semidefinite programs (TV-SDPs), based on the tracking of the solution trajectory of a low-rank matrix factorization, also known as the Burer-Monteiro factorization, in a path-following procedure. There, a path-following algorithm solves a sequence of linearized systems. This requires the introduction of a horizontal space constraint to ensure the local injectivity of the low-rank factorization. The method produces a sequence of approximate solutions for the original TV-SDP problem, for which we show that they stay close to the optimal solution path if properly initialized. Numerical experiments for a time-varying max-cut SDP relaxation demonstrate the computational advantages of the proposed method for tracking TV-SDPs in terms of runtime compared to off-the-shelf interior point methods.

Time-varying semidefinite programming: geometry of the trajectory of solutions – JAKUB MARECEK, CTU PRAGUE – Antonio Bellon, Didier Henrion, Vyacheslav Kungurtsev

In many applications, solutions of convex optimization problems need to be updated on-line, in response to time-varying inputs. We study behaviours of the solution trajectory, defined as the set of solutions depending on the time parameter. First, we consider parametric semidefinite programs, which are linear optimization problems in the semidefinite cone whose coefficients (input data) depend on a time parameter. As our main result, we show that only six distinct behaviors can be observed at a neighborhood of a given point along the solution trajectory. Each possible behavior is illustrated by an example. Second, we mention extensions to other time-varying conic optimization problems, including linear programming and polynomial optimization, and applications in transmission-constrained problems in power systems.

1e - ThA354 - 514C - **Recent advances in Variational Inequalities and Complementarity** Chair: JEAN PIERRE DUSSAULT 

The radius of statistical efficiency – MATEO DIAZ, JOHNS HOPKINS UNIVERSITY – Joshua Cutler, Dmitriy Drusvyatskiy

Classical results in asymptotic statistics show that the Fisher information matrix controls the difficulty of estimating a statistical model from observed data. In this work, we introduce a companion measure of robustness of an estimation problem: the radius of statistical efficiency (RSE) is the size of the smallest perturbation to the problem data that renders the Fisher information matrix singular. This quantity arises naturally as a statistical analogue of the "distance to ill-posedness", commonly used in numerical analysis and optimization. We com-

pute RSE up to numerical constants for a variety of test bed problems, including principal component analysis, generalized linear models, phase retrieval, blind deconvolution, and matrix completion. In all cases, the RSE quantifies the compatibility between the covariance of the population data and the underlying model parameter.

The Role of Subgradients in Second-Order Variational Analysis – EBRAHIM SARABI, MIAMI UNIVERSITY –

Any second-order variational construction is defined at a point in the graph of subgradient mappings. The second element of such points is a subgradient. Does it make a difference which subgradients we choose in these constructions? This talk is going to answer this question and demonstrate that certain subgradients, indeed, allow us to obtain more information about functions under consideration. Focusing mainly on the behavior of the second subderivative and subgradient proto-derivative of certain composite functions, we demonstrate that choosing the underlying subgradient, utilized in the definitions of these concepts, from the relative interior of the subdifferential mapping ensures stronger second-order variational properties such as strict twice epi-differentiability and strict subgradient proto-differentiability. Using this observation, we provide a simple characterization of continuous differentiability of the proximal mapping of our composite functions. As another application, we discuss the equivalence of metric regularity and strong metric regularity of a class of generalized equations at their nondegenerate solutions. This talk is based on joint works with Nguyen T. V. Hang.

Polyhedral Newton-min algorithms for complementarity problems – JEAN PIERRE DUSSAULT, UNIVERSITÉ DE SHERBROOKE – Jean Charles Gilbert, Mathieu Frappier

The semismooth Newton method is a very efficient approach for computing a zero of a large class of nonsmooth equations. When the initial iterate is sufficiently close to a regular zero and the function is strongly semismooth, the generated sequence converges quadratically to that zero, while the iteration only requires to solve a linear system.

If the first iterate is far away from a zero, however, it is difficult to force its convergence using linesearch or trust regions because a semismooth Newton direction may not be a descent direction of the associated least-square merit function, unlike when the function is differentiable. We explore this question in the particular case of a nonsmooth equation reformulation of the nonlinear complementarity problem, using the minimum function. We propose a globally convergent algorithm using a modification of a semismooth Newton direction that makes it a descent direction of the least-square function. Instead of requiring that the direction satisfies a linear system, it must be a feasible point of a convex polyhedron; hence, it can be computed in polynomial time. This polyhedron is defined by the often very few inequalities, obtained by linearizing pairs of functions that have close negative values at the current iterate; hence, somehow, the algorithm feels the proximity of a "negative kink" of the minimum function and acts accordingly. This is the plain PNM (Polyhedral Newton-min) algorithm.

In order to avoid as often as possible the extra cost of having to find a feasible point of a polyhedron, a hybrid algorithm (HNM) is also proposed, in which the Newton-min direction is

accepted if a sufficient-descent-like criterion is satisfied, which is often the case in practice. Global and fast convergence to regular points is proved.

PNM or HNM may be used for linear complementarity problems (LCPs) since only linear systems and polyhedral feasibility computations are involved, no LCP subproblem. Preliminary experiments on LCPs were encouraging, we observed that in most instances, the Newton-min direction was always accepted. Only in a few instances and in only one or two iterations, the algorithm needed expensive feasibility steps and these steps were helpful in avoiding failures.

Levenberg-Marquardt globalization of the polyhedral Newton-min algorithm for complementarity problems – BAPTISTE PLAQUEVENT JOURDAIN, INRIA PARIS, UNIVERSITÉ DE SHERBROOKE – Jean Pierre Dussault, Jean Charles Gilbert

Complementarity problems, in their rather general formulation, consist in finding some vector x such that $0 \preceq F(x) - G(x) \preceq 0$, where F and G are smooth vector functions. Optimality conditions for problems with nonlinear inequality constraints fall within this framework, but the problem has many other important applications. Remarkably, it is possible to reformulate the problem as a nonsmooth equation $H(x) = 0$. Then, a possible approach consists in applying nonsmooth Newton-like iterations to this equation. A method of choice is the semismooth Newton algorithm, which benefits from fast local convergence but lacks some global convergence property.

When $H = \min(F, G)$, one can overcome this drawback by using the polyhedral Newton-min (PNM) algorithm with line-search. The price to pay for this improvement is the need to find a point in a convex polyhedron (instead of solving a linear system) at each iteration. Furthermore, regularity conditions on the cluster points are required to guarantee that these are solutions. One goal of the presented work is to remedy to this latter restriction by using a Levenberg-Marquardt (LM) globalization technique. This extension is made difficult in particular because the LM method minimizes a sum of squared linearized functions, which mixes the components, whereas the NM algorithm is designed componentwise thus enjoys separability properties.

When at an iterate some components of F and G are equal, H is not differentiable. A technique used in the PNM algorithm deals with these components by combining F and G , while the NM algorithm only uses one of them. The proposed method follows a similar approach, using a componentwise convex combination of F and G . An appropriate choice of these weights allows one to obtain a descent curved path for the cost function, provided that the current iterate is not stationary. A tolerance is used to detect the proximity to the kinks of H , which intervenes in the analysis and is also relevant for numerical reasons.

1e - ThA912 - 514A - Variational Analysis, Variational Inequalities and Complementarity 2

Chair: TATJANA CHAVDAROVA

On projections onto Parametric Polyhedral Sets – KRZYSZTOF RUTKOWSKI, CARDINAL STEFAN WYSZYNSKI UNIVERSITY IN WARSAW –

In this presentation we are investigating Lipschitz continuity

of the projection of a given point onto a moving polyhedron (parametric polyhedral set), in the case when the parameter is included in the functions of the left and right-hand sides of a linear system of inequalities. The problem can be cast equivalently as a parametric variational inequality and related to the properties of the graphical subdifferential mapping. Basing on that result, the sufficient conditions for Lipschitzness of projection onto a moving polyhedron will be discussed. We will show that, whenever an equivalent representation of a moving polyhedron, called stable, exists, then the selected regularity conditions are enough to ensure local Lipschitzness of projection onto a moving polyhedron.

Douglas-Rachford splitting algorithm with line search – LINGLING XU, SCHOOL OF MATHEMATICAL SCIENCE, NANJING NORMAL UNIVERSITY

Douglas-Rachford (DR) splitting algorithm is a classical splitting algorithm for solving the zero point problem of two maximal monotone operators. The selection of step size is of great significance for the convergence speed of the algorithm and proper line search criteria can effectively improve the running speed of the algorithm. We proposed a DR splitting algorithm with line search, and proves the global convergence of the algorithm sequence and the sub-linear convergence rate of the function value corresponding to the sequence generated by the algorithm in the ergodic sense.

Primal-Dual Methods with Adjoint Mismatch - Existence of Stationary Points, Convergence and Error Bounds – FELIX SCHNEPPE, UNIVERSITY OF BREMEN –

The Chambolle-Pock method and the primal-dual Douglas-Rachford method are well-known standard methods to solve saddle-point problems of the form

$$\min_x \max_y G(x) + \langle Ax, y \rangle - F^*(y).$$

However, in practical applications like computed tomography, the adjoint operator is often replaced by a computationally more efficient approximation. This leads to an adjoint mismatch in the algorithms, which translates to a replacement of A^* with some linear operator V^* in the algorithms and to non-convergence in practical applications. In this talk, we analyze the convergence properties of the Chambolle-Pock method as well as the primal-dual Douglas-Rachford method under the presence of an adjoint mismatch and prove conditions under which the existence of a solution can still be guaranteed. We explain the difficulties in finding appropriate step sizes and present methods to calculate them. Furthermore, we discuss how the adjoint mismatch affects the implementation of the algorithm. Additionally, we show that we have to solve a variational inclusion instead of the former saddle-point problem, but we are able to present error bounds on the primal solution for both algorithms.

Last-Iterate Convergence of Saddle-Point Optimizers via High-Resolution Differential Equations – TATJANA CHAVDAROVA, CISPA AND SAARLAND UNIVERSITY –

Various first-order methods used for solving Variational Inequalities (VI) produce the same continuous-time ordinary differential equation (ODE) when derived simply. However, these

methods have different convergence behaviors, even in basic bilinear games. Consequently, while the ODE approach has been influential in analyzing methods minimizing real-valued functions, it has not had a comparable impact in VI optimization.

We utilize a framework from fluid dynamics, called High-Resolution Differential Equations (HRDEs), to formulate ODE models for VI optimization methods. Importantly, these HRDEs differ across methods. Furthermore, in bilinear games, the convergence properties of the HRDEs align with the characteristics of the corresponding discrete methods. We show that the HRDE for Optimistic Gradient Descent Ascent (OGDA) has last-iterate convergence for general monotone variational inequalities. Lastly, we establish convergence rates for the best-iterate convergence of OGDA, based solely on the first-order smoothness of the monotone operator.

1e - ThB226 - 512E - **Nonconvexity, stochasticity and hierarchy in optimization problems**
Chair: MATHIAS STAUDIGL

The inexact power augmented Lagrangian method for constrained nonconvex optimization – ALEXANDER BODARD, KU LEUVEN – Konstantinos Oikonomidis, Emanuel Laude, Patrinos Panagiotis

In this talk, we analyze an inexact augmented Lagrangian method with an unconventional augmenting term for solving a general class of nonconvex problems with nonlinear constraints. It is shown that this method achieves a better iteration complexity in terms of the constraint violation than its classical counterpart. Additionally, we provide a total complexity analysis when using an accelerated first-order method to solve the Hölder-smooth inner problems. This result suggests that the optimal power of the augmenting term depends on the desired tolerance for the suboptimality and constraint violation. We also discuss the convex setting, in which the method corresponds to a higher-order proximal point method, yielding a fast global sublinear rate and a local superlinear rate. Finally, numerical experiments validate the practical merits of our power augmented Lagrangian method on some problems of practical interest.

A quasi-Newton method for nonsmooth, non-convex, stochastic optimization – LUKE MARRINAN, PENN STATE UNIVERSITY –

We consider the minimization of a Lipschitz continuous and expectation-valued function over a closed and convex set. Our focus lies on obtaining both asymptotics as well as rate and complexity guarantees for computing an approximate stationary point (in a Clarke sense) via zeroth-order schemes. We adopt a randomized-smoothing-based approach reliant on minimizing a smooth approximation of our objective function. In such a setting, we develop a zeroth-order stochastic quasi-Newton scheme reliant on a combination of randomized and Moreau smoothing is proposed and analyzed, for which iteration and sample complexities are derived.

Zeroth-order federated methods for stochastic MPECs and nondifferentiable nonconvex hierarchical optimization – YUYANG QIU, RUTGERS UNIVERSITY – Uday Shanbhag, Farzad Yousefian
Motivated by the emergence of federated learning (FL), we

design and analyze communication-efficient decentralized computational methods for addressing four broadly applicable but hitherto unstudied problem classes in FL: (i) Nondifferentiable nonconvex optimization; (ii) Bilevel optimization; (iii) Minimax problems; and (iv) Two-stage stochastic mathematical programs with equilibrium constraints (2s-SMPEC). Notably, in an implicit sense, (ii), (iii), and (iv) are instances of (i). However, these hierarchical problems are often complicated by the absence of a closed form expression for the implicit objective function. Research on these problems has been limited and afflicted by reliance on strong assumptions, including the need for differentiability of the implicit function and the absence of constraints in the lower-level problem, among others. We address these shortcomings by making the following contributions. In (i), by leveraging convolution-based smoothing and Clarke’s subdifferential calculus, we devise a randomized smoothing-enabled zeroth-order FL method and derive communication and iteration complexity guarantees for computing an approximate Clarke stationary point. To contend with (ii) and (iii), we devise a unifying randomized implicit zeroth-order FL framework, equipped with explicit communication and iteration complexities. Importantly, our method utilizes delays during local steps to skip calls to the inexact lower-level FL oracle. This results in significant reduction in communication overhead. In (iv), we devise an inexact implicit variant of the method in (i). Remarkably, this method achieves a total communication complexity matching that of single-level nonsmooth nonconvex optimization in FL. This appears to be the first time that 2s-SMPECs are provably addressed in FL. We empirically validate the theoretical findings on instances of federated nonsmooth and hierarchical problems including training of ReLU neural networks, hyperparameter learning, fair classification, and Stackelberg-Nash-Cournot equilibrium seeking problem.

Hessian barrier algorithms for non convex conic optimization – MATHIAS STAUDIGL, UNIVERSITY OF MANNHEIM – Pavel Dvurechensky

A key problem in mathematical imaging, signal processing and computational statistics is the minimization of non-convex objective functions that may be non-differentiable at the relative boundary of the feasible set. This paper proposes a new family of first- and second-order interior-point methods for non-convex optimization problems with linear and conic constraints, combining logarithmically homogeneous barriers with quadratic and cubic regularization respectively. Our approach is based on a potential-reduction mechanism and, under the Lipschitz continuity of the corresponding derivative with respect to the local barrier-induced norm, attains a suitably defined class of approximate first- or second-order KKT points with worst-case iteration complexity $O(\epsilon^{-2})$ (first-order) and $O(\epsilon^{-3/2})$ (second-order), respectively. Based on these findings, we develop new path-following schemes attaining the same complexity, modulo adjusting constants. These complexity bounds are known to be optimal in the unconstrained case, and our work shows that they are upper bounds in the case with complicated constraints as well. To the best of our knowledge, this work is the first which achieves these worst-case complexity bounds under such weak conditions for general conic constrained non-convex optimization problems.

1e - ThB913 - 514A - **Variational Analysis, Variational Inequalities and Complementarity 3**
Chair: ABDERRAHIM JOURANI

Calm local optimality for nonconvex-nonconcave minimax problems – XIAOXIAO MA, UNIVERSITY OF VICTORIA – Wei Yao, Jane Ye, Zhang Jin

Nonconvex-nonconcave minimax problems have found numerous applications in various fields including machine learning. However, questions remain about what is a good surrogate for local minimax optimum and how to characterize the minimax optimality. Recently Jin, Netrapalli, and Jordan (ICML 2020) introduced a concept of local minimax point and derived optimality conditions for the smooth and unconstrained case. In this work, we introduce the concept of calm local minimax point, which is a local minimax point with a calm radius function. With the extra calmness property we obtain first and second order sufficient and necessary optimality conditions for a very general class of nonsmooth nonconvex-nonconcave minimax problem. Moreover we show that the calm local minimax optimality and the local minimax optimality coincide under a weak sufficient optimality condition for the maximization problem. This equivalence allows us to derive stronger optimality conditions under weaker assumptions for local minimax optimality.

On desingularizing functions in convex programming – ABDERRAHIM JOURANI, INSTITUT DE MATHÉMATIQUES DE BOURGOGNE, UNIVERSITÉ DE BOURGOGNE –

Over the past years, the idea of the KL property has become increasingly common in the optimisation area. It was used as tool guarantees that every sequence generated by many different descent algorithms enjoys finite length property. When verifying the KL property, one needs to find a desingularizing function, which is usually considered difficult. In light of this, one provides a desingularizing function that can be regarded as the smallest among all potential desingularizing functions. We investigate continuity, Lipschitz continuity, differentiability and subdifferentiability properties of this function as well as different characterizations of the so-called Kurdyka-Lojasiewicz-Hoffman constant by introducing a new class of functions with nonsmooth moderate behaviour. Based on these attributes, we provide a convergence analysis and an estimation of the convergence rate of an inexact descent methods for convex differentiable functions that covers a wide class of gradient descent methods. We also delves the complexity property of this method and its relationship with λ -one-dimensional worst-case proximal sequences.

1e - FA103 - 514A - **Recent advances in bilevel optimization** Chair: JAN HAROLD ALCANTARA

Moreau envelope based reformulation and algorithm for bilevel programs – SHANGZHI ZENG, UNIVERSITY OF VICTORIA –

Bilevel programming has emerged as a valuable tool for hyperparameter selection, a central concern in machine learning. In a recent study by Ye et al. (2023), a value function-based difference of convex algorithm was introduced to address bilevel programs. This approach proves powerful when dealing with scenarios where the lower-level problem exhibits convexity in

both the upper-level and lower-level variables. In this work, we expand the range of applications, now requiring convexity only in the lower-level variables of the lower-level program. We present a new single-level difference of weakly convex reformulation based on the Moreau envelope of the lower-level problem. We further develop a sequentially convergent Inexact Proximal Difference of Weakly Convex Algorithm (iP-DwCA). Additionally, for scenarios where all problem functions are smooth, we develop a single-loop gradient-based algorithm.

On first-order methods for stochastic bilevel optimization – JEONGYEOL KWON, UNIVERSITY OF WISCONSIN MADISON –

We study first-order algorithms for solving Bilevel Optimization with access to stochastic gradient oracles. The problem has been extensively studied in recent years; the main technical challenge is to keep track of lower-level solutions in response to the changes in the upper-level variables. Subsequently, all existing approaches tie their analyses to a genie algorithm that knows lower-level solutions and, therefore, need not query any points far from them. We consider a dual question to such approaches: suppose we have an oracle, which we call y^* -aware, that returns an estimate of the lower-level solution alongside the first-order gradient estimators locally unbiased within a reliable region. We study the complexity of finding stationary points with the proposed y^* -aware oracle; we derive the matching upper and lower bounds on the iteration complexities with access to first-order y^* -aware oracles without/with additional stochastic smoothness assumptions. Our upper bound result can be naturally extended to the standard unbiased oracles without the loss in the complexity, improving the best-known upper bounds by a polynomial factor. Furthermore, our upper bound suggests that the first-order methods are not necessarily slower than second-order methods under the fair stochastic smoothness comparison. On the other hand, our lower bound implies that any approach that simulates y^* -aware oracle must suffer the same lower bounds, and we propose the conjecture on the tightness of our results.

Relaxation methods for pessimistic bilevel optimization – ALAIN ZEMKOHO, THE UNIVERSITY OF SOUTHAMPTON –

When the lower-level optimal solution set-valued mapping of a bilevel optimization problem is not single-valued, we are faced with an ill-posed problem, which gives rise to the optimistic and pessimistic bilevel optimization problems, as tractable algorithmic frameworks. However, solving the pessimistic bilevel optimization problem is far more challenging than the optimistic one. Relaxation methods have appeared to be some of the most efficient techniques to solve the optimistic bilevel optimization problem in its Karush-Kuhn-Tucker (KKT) reformulation or the corresponding more general mathematical program with complementarity constraints. Inspired by this success, this work studies their potential in the context of the pessimistic bilevel optimization problems. To proceed, we consider the KKT reformulation of a pessimistic bilevel optimization problem, where all the functions involved are at least continuously differentiable, the lower-level problem is convex, and the Slater constraint qualification is satisfied. We then construct theoretical results ensuring that a sequence of global/local optimal solutions (resp. stationarity points) of the relaxations converges to the corresponding points for the KKT reformulation of the pessimistic bilevel optimization. The results are accom-

panied by some technical proofs ensuring that the relaxation algorithms are well-defined. In this talk, we will provide a brief overview of the results of this work, as well as some numerical results on the implementation of the algorithms.

Theoretical smoothing frameworks for general nonsmooth bilevel problems – JAN HAROLD ALCANTARA, RIKEN CENTER FOR ADVANCED INTELLIGENCE PROJECT – Akiko Takeda

The optimal value function approach provides a useful reformulation of the bilevel optimization problem, but its utility is often hindered by the nonsmoothness of the value function even in cases when the associated lower-level function is smooth. In this talk, we present two smoothing strategies for the value function associated with lower-level functions that are not necessarily smooth but are Lipschitz continuous. The first method employs a quadratic regularization approach for partially convex lower-level functions, while the second utilizes entropic regularization for general lower-level objective functions. Meanwhile, the property known as gradient consistency is crucial in ensuring that a designed smoothing algorithm is globally subsequently convergent to stationary points of the equivalent reformulation of the bilevel problem. With this in mind, we show that the proposed smooth approximations satisfy the gradient consistent property under certain conditions on the lower-level function.

1e - FB336 - 514A - **Nonsmooth and hierarchical optimization in machine learning** Chair: RAHUL PARHI 

First-order methods for bilevel optimization – ZHAOSONG LU, UNIVERSITY OF MINNESOTA – Sanyou Mei

Bilevel optimization, also known as two-level optimization, is an important branch within mathematical optimization. It has found applications across various domains, including economics, logistics, supply chain, transportation, engineering design, and machine learning. In this talk, we will present first-order methods for solving a class of bilevel optimization problems using either single or sequential minimax optimization schemes. We will also discuss the first-order operation complexity of these methods and present preliminary numerical results to illustrate their performance.

Classification under strategic adversary manipulation using pessimistic bilevel optimization – DAVID BENFIELD, UNIVERSITY OF SOUTHAMPTON – Alain Zemkoho

Adversarial machine learning concerns the situations where learners face attacks from active adversaries. In particular, the underlying distribution of the data is vulnerable to significant changes made by the adversary. We consider the case of test-time attacks, where such changes to the data occur in response to trained prediction models. Under these attacks, the accuracy of the prediction model at time of implementation is susceptible to significant decreases in accuracy. Such scenarios arise in applications such as spam email filtering, malware detection and fake-image generation. We model these interactions between the learner and the adversary as a game. While many game theoretic models exist, some of these assume that the players act simultaneously. However, in the case of adversarial learning, a perhaps more appropriate assumption is

that players can observe their opponent's actions before making their own. For example, spam email senders might probe an email filter by sending test emails before deploying their final products. As such, we model the interaction between the learner and adversary as a Stackelberg game, with the learner taking the role of the leader. The adversary, modelled as a data generator, takes the role of the follower, constructing their generator in response to the learner's prediction function. The game is formulated as a pessimistic bilevel model to realistically capture the antagonistic nature of adversarial learning scenarios. While existing pessimistic models rely on assumptions about the uniqueness of the lower-level solution, we will present our model and solution method which do not make such assumptions.

A Banach-space view of neural network training – RAHUL PARHI, UCSD –

In this talk, we consider the problem of training neural networks with common regularization schemes such as weight decay (which corresponds to the standard Tikhonov regularization). In the case of shallow neural networks, it turns out that this non-convex optimization problem can be lifted to a convex optimization problem posed over a particular Banach space. Our main result is a theorem that states that every solution to the non-convex neural network training problem is a solution to the optimization problem posed over this Banach space. In particular, the solution set to the Banach-space problem is nonempty, convex, and weak* compact and its extreme points are finite-width neural networks whose widths are bounded by the number of data in the original training objective. This Banach space does not coincide with classical function spaces studied in analysis as it is defined via a measure of smoothness in the domain of the Radon transform of the underlying functions. We will discuss several implications of this correspondence between neural network training problems and Banach-space optimization problems towards the understanding of the impressive performance seen by neural networks. These results could also motivate the design of new optimization algorithms for training neural networks. If time permits, we will discuss extensions of these results to other architectures.

1f - TA130 - 512E - **Fast randomized algorithms for distributed optimization** Chair: LAURENT CONDAT 

CMA Light: a minibatch algorithm for large-scale non-convex finite sum optimization – CORRADO COPPOLA, SAPIENZA UNIVERSITY OF ROME – Laura Palagi, Giampaolo Liuzzi, Lorenzo Papa, Irene Amerini

We consider minimizing the sum of a large number of smooth and possibly non-convex functions, which is the typical problem encountered in the training of deep neural networks on large-size datasets. By drawing inspiration from the Controlled Minibatch Algorithm (CMA) scheme proposed by Liuzzi et al. (2022), we define CMA Light, an incremental gradient-like method, which requires seldom computation of the objective function, needed only when light watchdog rules are not satisfied. Thanks to an extrapolation derivative-free linesearch, convergence still holds without any assumptions on the norm of the gradient. We firstly carry out numerical test on standard feed-forward neural networks to highlight the advantage

of CMA Light over CMA and other mini-batch algorithms on large-size datasets. Then, we present a thorough computational analysis of CMA Light performance on standard image classification tasks on CIFAR10 and CIFAR100 with deep convolutional neural networks and residual networks, showing that CMA Light can easily scale up to problem with 50-100 millions of variables and has an advantage over its state-of-the-art competitors. Finally, we present CMA Light randomized counterpart, which converges with standard assumptions on the norm of the gradient, and show its scaling capabilities on classification tasks with vision transformers, where the number of variables grows up to the order of billions.

On adaptive stochastic heavy ball momentum for solving linear systems – JIAXIN XIE, BEIHANG UNIVERSITY –

The stochastic heavy ball momentum (SHBM) method has gained considerable popularity as a scalable approach for solving large-scale optimization problems. However, one limitation of this method is its reliance on prior knowledge of certain problem parameters, such as singular values of a matrix. In this paper, we propose an adaptive variant of the SHBM method for solving stochastic problems that are reformulated from linear systems using user-defined distributions. Our adaptive SHBM (ASHBM) method utilizes iterative information to update the parameters, addressing an open problem in the literature regarding the adaptive learning of momentum parameters. We prove that our method converges linearly in expectation, with a better convergence bound compared to the basic method. Notably, we demonstrate that the deterministic version of our ASHBM algorithm can be reformulated as a variant of the conjugate gradient (CG) method, inheriting many of its appealing properties, such as finite-time convergence. Consequently, the ASHBM method can be further generalized to develop a brand-new framework of the stochastic CG (SCG) method for solving linear systems. Our theoretical results are supported by numerical experiments.

Optimal gradient tracking for decentralized optimization – MING YAN, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN –

In this talk, I will focus on solving the decentralized optimization problem of minimizing the sum of objective functions over a multi-agent network. The agents are embedded in an undirected graph where they can only send/receive information directly to/from their immediate neighbors. Assuming smooth and strongly convex objective functions, we propose an Optimal Gradient Tracking (OGT) method that simultaneously achieves the optimal gradient computation complexity and communication complexity. Its development involves two building blocks that are also of independent interest. The first one is another new decentralized gradient tracking method, SSGT, which achieves optimal gradient computation. The second one is a technique termed LCA, which can be implemented "looplessly" but achieves a similar effect by adding multiple inner loops of Chebyshev acceleration in the algorithm. This LCA technique can accelerate many other gradient tracking-based methods concerning the graph condition number.

Communication-efficient distributed optimization algorithms – LAURENT CONDAT, KAUST – Peter Richtárik

In the big data era, the explosion in size and complexity of

data arises in parallel to a shift towards distributed computations. In distributed optimization and machine learning, especially in federated learning, specific issues arise, such as decentralized data storage for privacy concerns. In this setting, a large number of machines perform computations in parallel and communicate back and forth with a distant server. Communication is typically costly and slow and forms the main bottleneck. To decrease it, two strategies are popular: 1) communicate less frequently; 2) compress the communicated vectors. I will present several randomized algorithms we developed recently in this area, with proved convergence guarantees and state-of-the-art communication complexity.

1f - TB131 - 512C - Beyond Standard Assumptions in Optimization: Structured Non-Convexity, Non-Monotonicity, and Non-Smoothness Chair: EDUARD GORBUNOV

Generalized-Smooth Nonconvex Optimization is As Efficient As Smooth Nonconvex Optimization – YI ZHOU, UNIVERSITY OF UTAH –

Various optimal gradient-based algorithms have been developed for smooth nonconvex optimization. However, many nonconvex machine learning problems do not belong to the class of smooth functions and therefore the existing algorithms are sub-optimal. Instead, these problems have been shown to satisfy certain generalized-smooth conditions, which have not been well understood in the existing literature. In this paper, we propose a notion of α -symmetric generalized-smoothness that extends the existing notions and covers many important functions such as high-order polynomials and exponential functions. We study the fundamental properties and establish descent lemmas for the functions in this class. Then, to solve such a large class of nonconvex problems, we design a special deterministic normalized gradient descent algorithm that achieves the optimal iteration complexity $\mathcal{O}(\epsilon^{-2})$, and also prove that the popular SPIDER variance reduction algorithm achieves the optimal sample complexity $\mathcal{O}(\epsilon^{-3})$ in the stochastic setting. Our results show that solving generalized-smooth nonconvex problems is as efficient as solving smooth nonconvex problems.

Stochastic Optimization under Hidden Convexity – ILYAS FATKHULLIN, ETH ZÜRICH – Niao He, Yifan Hu

In this work, we consider constrained stochastic optimization problems under hidden convexity, i.e., those that admit a convex reformulation via non-linear (but invertible) map $c(\cdot)$. A number of non-convex problems ranging from optimal control, revenue and inventory management, to convex reinforcement learning all admit such a hidden convex structure. Unfortunately, in the majority of applications considered, the map $c(\cdot)$ is unavailable or implicit; therefore, directly solving the convex reformulation is not possible. On the other hand, the stochastic gradients with respect to the original variable are often easy to obtain. Motivated by these observations, we examine the basic projected stochastic (sub-) gradient methods for solving such problems under hidden convexity. We provide the first sample complexity guarantees for global convergence in smooth and non-smooth settings. Additionally, in the smooth setting, we improve our results to the last iterate convergence in terms of

function value gap using the momentum variant of projected stochastic gradient descent.

Convergence of Douglas–Rachford Splitting and Primal-Dual Hybrid Gradient for Non-Monotone, Non-Smooth Problems – BRECHT EVENS, KU LEUVEN –

Despite the widespread use of splitting methods, their convergence analysis has predominantly been limited to convex/monotone settings. This presentation delves into the convergence properties of two such methods, Douglas-Rachford splitting (DRS) and the primal-dual hybrid gradient (PDHG) method, in the absence of monotonicity and smoothness. Specifically, we introduce the notion of semimonotonicity and establish conditions under which DRS and PDHG exhibit global convergence, even when dealing with the sum of two semimonotone operators. Through illustrative examples, we showcase the versatility of our theoretical framework in addressing problems emerging in optimization and reinforcement learning.

On the Lower Bound of Minimizing Polyak-Lojasiewicz functions – CONG FANG, PEKING UNIVERSITY –

Polyak-Lojasiewicz (PL) (Polyak, 1963) condition is a weaker condition than the strong convexity but suffices to ensure a global convergence for the Gradient Descent algorithm. In this paper, we study the lower bound of algorithms using first-order oracles to find an approximate optimal solution. We show that any first-order algorithm requires at least $\Omega\left(\frac{L}{\mu} \log \frac{1}{\epsilon}\right)$ gradient costs to find an ϵ -approximate optimal solution for a general L -smooth function that has an μ -PL constant. This result demonstrates the optimality of the Gradient Descent algorithm to minimize smooth PL functions in the sense that there exists a “hard” PL function such that no first-order algorithm can be faster than Gradient Descent when ignoring a numerical constant. In contrast, it is well-known that the momentum technique, e.g. Nesterov (2003, chap. 2), can provably accelerate Gradient Descent to $O\left(\sqrt{\frac{L}{\mu}} \log \frac{1}{\epsilon}\right)$ gradient costs for functions that are L -smooth and $\hat{\mu}$ -strongly convex. Therefore, our result distinguishes the hardness of minimizing a smooth PL function and a smooth strongly convex function as the complexity of the former cannot be improved by any polynomial order in general.

1f - WA334 - 512C - **Random subspace methods for continuous optimization** Chair: PIERRE LOUIS POIRION 

A Semismooth Newton Stochastic Proximal Point Algorithm with Variance Reduction – MICHAEL ULBRICH, TECHNICAL UNIVERSITY OF MUNICH – Andre Milzarek, Fabian Schaipp

We develop an implementable stochastic proximal point (SPP) method for a class of weakly convex, composite optimization problems. The proposed stochastic proximal point algorithm incorporates a variance reduction mechanism and the implicit SPP updates are computed by applying an inexact semismooth Newton method to a subproblem that is typically small-scale. We establish a convergence theory that takes the inexactness of the SPP steps into account and which is in accordance with existing convergence guarantees of (proximal) stochastic

variance-reduced gradient methods. Numerical experiments show that the proposed algorithm compares favorably with other state-of-the-art methods and achieves higher robustness with respect to the step size selection.

Applying Random projection techniques to large-scale nonconvex optimization problems – AKIKO TAKEDA, THE UNIVERSITY OF TOKYO / RIKEN –

Random projection techniques based on the Johnson-Lindenstrauss lemma are used for randomly aggregating the constraints or variables of optimization problems while approximately preserving their optimal values, which leads to smaller-scale optimization problems. In this talk, we show a few applications of random matrix techniques for constructing random subspace algorithms that iteratively solve smaller-scale subproblems and discuss their convergence speed. This talk is based on joint works with Ryota Nozawa, and Pierre-Louis Poirion.

Random projections for quadratically constrained quadratic programming – BENEDETTO MANCA, UNIVERSITÀ DEGLI STUDI DI CAGLIARI – Pierre Louis Poirion, Leo Liberti

We consider quadratically constrained quadratic programs (QCQP) in a general form with both linear and quadratic constraints. We will show that, if the number of quadratic constraints m and the number of decision variables n are large enough, there exist two random matrices T (with k rows and m columns) and R (with r rows and n columns) such that it is possible to apply them to the parameters and variables of the original problem in order to obtain another QCQP with fewer decision variables and quadratic constraints which is easier and faster to solve. We will provide theoretical results showing that the projected program is related to the original one in terms of both feasibility and optimality and a procedure that exploits a solution of the projected program to obtain a solution of the original one. Numerical experiments on some different classes of QCQPs will be presented.

Random subspace Newton method for unconstrained non-convex optimization – PIERRE LOUIS POIRION, RIKEN – Terunari Fuji, Akiko Takeda

Abstract: In this talk, we present a randomized subspace regularized Newton method for a non-convex function. We show that our method has global convergence under appropriate assumptions, and its convergence rate is the same as that of the full regularized Newton method. Furthermore, we can obtain a local linear convergence rate, under some additional assumptions, and prove that this rate is the best we can hope, under local strong convexity assumptions, when using random subspace. Finally, we will prove that under rank deficiency conditions, one can obtain super-linear convergence in the subspace setting.

1f - WB49 - 512C - **Splitting and Randomized Methods 1** Chair: DEREN HAN 

Improved L-BFGS Methods for Large-Scale Unconstrained Optimization – MEHIDDIN AL BAALI, SULTAN QABOOS UNIVERSITY –

New techniques for modifying the limited memory L-BFGS method for large-scale unconstrained optimization will be considered. They modify the L-BFGS update based on testing

appropriate conditions in a sense to be defined. One technique reduces the number of the BFGS updates without destroying the quality of the L-BFGS Hessian. In another technique, the difference in gradients of the objective function, which appears in the BFGS updating formula, is modified by certain formulae whenever necessary so that a sufficient positive curvature condition is satisfied. A possible combination of the best features of the above two techniques will be stated as one for the L-BFGS method. Numerical results will be described to show that the proposed techniques improve the performance of the L-BFGS method substantially in several cases.

A Decentralized Primal-Dual Method with Quasi-Newton Tracking – LIPING WANG, NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS –

This talk considers the decentralized optimization problem of minimizing a finite sum of strongly convex and twice continuously differentiable functions over a fixed-connected undirected network. A fully decentralized primal-dual method (DPDM) and its generalization (GDPDM), which allows for multiple primal steps per iteration, are proposed. In our method, both primal and dual updates use second-order information while Hessian approximations are constructed by simple algebraic calculations and at most matrix-vector products. Specifically, the primal update combines the single loop Jacobi with the BFGS method for efficient computation and distributed communication. The dual update newly embodies a second-order correction built with a new dual gradient variation with Barzilai-Borwein (BB) stepsize. Additionally, on each node, the local direction asymptotically approaches the centralized quasi-Newton direction. Under some mild assumptions, decentralized primal-dual method (DPDM) and its generation (GDPDM) are shown to have a global linear convergence rate for solving strongly convex decentralized optimization problems. Numerical results on linear regression problem illustrate that GDPDM is less impacted by deterioration of condition number of the objective function than existed methods. The numerical behavior on logistic regression demonstrates progressing performance of decentralized primal-dual method in iterative convergence as well as distributed communication.

Group SLOPE Penalized Low-Rank Tensor Regression – ZIYAN LUO, BEIJING JIAOTONG UNIVERSITY –

We aim to seek a selection and estimation procedure for a class of tensor regression problems with multivariate covariates and matrix responses, which can provide theoretical guarantees for model selection in finite samples. Considering the frontal slice sparsity and low-rankness inherited in the coefficient tensor, we formulate the regression procedure as a group SLOPE penalized low-rank tensor optimization problem based on an orthogonal decomposition, namely TgSLOPE. This procedure provably controls the newly introduced tensor group false discovery rate (TgFDR), provided that the predictor matrix is column orthogonal. Moreover, we propose conditions to guarantee the global identifiability of the coefficient tensor to be estimated based on the so-called Kruskal rank, and establish the asymptotically minimax convergence with respect to the TgSLOPE estimate risk. For efficient problem resolution, we equivalently transform the TgSLOPE problem into a difference-of-convex (DC) program with the level-coercive objective function. This allows us to solve the reformulation problem of TgSLOPE by

an efficient proximal DC algorithm (DCA) with global convergence, provided that the predictor matrix is full column rank. Numerical studies conducted both on synthetic data and a real human brain connection data illustrate the efficacy of the proposed TgSLOPE estimation procedure.

Revisit Iterative Complexity of Some Projection Methods for Monotone Linear Variational Inequalities – DEREN HAN, BEIHANG UNIVERSITY –

For monotone linear variational inequality problems, projection methods play an important role, due to its low computational cost when the projection onto the set is simple. Though its global convergence is well understood, its iteration complexity is established until the recent work of Chen, Fu, He and Yuan (Journal of Optimization Theory and Applications, 2017). While the non-ergodic complexity is measured by the natural residual, the ergodic complexity is measured by the restricted dual gap function. Moreover, the non-ergodic complexity is measured with the best case (the minimal of the iterate), and we are more interest in the last iterate. In this paper, we give a simple proof of the ergodic complexity measured by the natural residual, and the non-ergodic complexity of the last iterate (for Algorithm-I). Under some additional conditions, we also presented their linear rate of convergence.

1f - ThA128 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization I Chair: NICOLAS LOIZOU

Global-Local Smoothness: Line Search can really help! Really! – MARK SCHMIDT, UBC –

Iteration complexities for first-order optimization algorithms are typically stated in terms of a global Lipschitz constant of the gradient, and optimal results can typically be achieved using fixed step sizes. But many objective functions that arise in practice have regions with small Lipschitz constants where larger step sizes can be used. Many local Lipschitz assumptions have thus been proposed, which lead to results showing that adaptive step sizes and/or line searches yield improved convergence rates over fixed step sizes. However, these faster rates tend to depend on the path taken by the algorithm, which makes it difficult to compare the iteration complexities of different methods.

We consider a simple characterization of global and local smoothness that only depends on properties of the function. This allows explicit lower and upper bounds on iteration complexities in terms of problem-dependent constants, which allows us to compare iteration complexities between algorithms. Under this assumption it is straightforward to show the advantages of line searches over fixed step sizes, of line searches with unbounded step sizes over those that bound the maximum step size, and even that in some settings gradient descent with line search has a better iteration complexity than accelerated gradient methods with fixed step sizes.

What is the right extension of Polyak's step-size to mirror descent? – RYAN D'ORAZIO, MILA, UNIVERSITÉ DE MONTRÉAL – Nicolas Loizou, Issam Laradji, Mitliagkas Ioannis

We investigate the convergence of stochastic mirror descent (SMD) under interpolation in relatively smooth and smooth

convex optimization. In relatively smooth convex optimization we provide new convergence guarantees for SMD with a constant stepsize. For smooth convex optimization we propose a new adaptive stepsize scheme — the mirror stochastic Polyak stepsize (mSPS). We show convergence to a neighborhood that vanishes under interpolation. Consequently, these results correspond to the first convergence guarantees under interpolation for the exponentiated gradient algorithm for fixed or adaptive stepsizes. Although mSPS provides one generalization of the Polyak stepsize, it may not be best suited for relatively smooth problems, we discuss potential alternatives that be more compatible with relative smoothness assumptions.

Scaling Law: Finding Compute Optimal Curves on a Simple Model with SGD – COURTNEY PAQUETTE, MCGILL UNIVERSITY –

We describe a program of analysis of stochastic gradient descent on a high-dimensional least squares problem with power law random features. First, we give description of these loss curves that can be analyzed precisely. This is a simple two-hyperparameter family of optimization problems, which displays 5 distinct phases of loss curves; these phases are determined by the relative complexities of the target, data distribution, and whether these are ‘high-dimensional’ or not (which in context can be precisely defined). In each phase, we also give, for a given fixed compute budget, the optimal random-feature dimensionality. In this sense, we predict the best allocation of parameters for a given compute budget.

Joint work with Elliot Paquette (McGill), Jeffrey Pennington (Google Deepmind), and Lechao Xiao (Google Deepmind).

Generalized Polyak Step Size for First Order Optimization with Momentum – XIAOYU WANG, THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY – MIKAEL JOHANSSON, TONG ZHANG

In machine learning applications, it is well known that carefully designed learning rate (step size) schedules can significantly improve the convergence of commonly used first-order optimization algorithms. Therefore how to set step size adaptively becomes an important research question. A popular and effective method is the Polyak step size, which sets step size adaptively for gradient descent or stochastic gradient descent without the need to estimate the smoothness parameter of the objective function. However, there has not been a principled way to generalize the Polyak step size for algorithms with momentum accelerations. This paper presents a general framework to set the learning rate adaptively for first-order optimization methods with momentum, motivated by the derivation of Polyak step size. It is shown that the resulting techniques are much less sensitive to the choice of momentum parameter and may avoid the oscillation of the heavy-ball method on ill-conditioned problems. These adaptive step sizes are further extended to the stochastic settings, which are attractive choices for stochastic gradient descent with momentum. Our methods are demonstrated to be more effective for stochastic gradient methods than prior adaptive step size algorithms in large-scale machine learning tasks.

1f - ThB50 - 512C - Splitting and Randomized Methods 2 Chair: YANNAN CHEN

Uniform Function Estimators in Reproducing

Kernel Hilbert Spaces, with applications in stochastic optimization – ALOIS PICHLER, UNIVERSITY OF TECHNOLOGY CHEMNITZ –

We reconstruct functions, which are observed with superimposed errors at random locations. The problem is considered in reproducing kernel Hilbert spaces (RKHS). It is demonstrated that the estimator, which is often derived by employing Gaussian random fields, converges in the mean norm of the RKHS to the conditional expectation and this implies local and uniform convergence of this function estimator. By preselecting the kernel, the problem does not suffer from the curse of dimensionality.

We analyze the statistical properties of the estimator. We derive convergence properties and provide a conservative rate of convergence for increasing sample sizes.

Data-Driven Minimax Optimization with Expectation Constraints – XUDONG LI, FUDAN UNIVERSITY – SHUOGUANG YANG, GUANGHUI LAN

Attention to data-driven optimization approaches, including the well-known stochastic gradient descent method, has grown significantly over recent decades, but data-driven constraints have rarely been studied, because of the computational challenges of projections onto the feasible set defined by these hard constraints. In this talk, we focus on the non-smooth convex-concave stochastic minimax regime and formulate the data-driven constraints as expectation constraints. The minimax expectation constrained problem subsumes a broad class of real-world applications, including data-driven robust optimization, optimization with misspecification, and Area Under the ROC Curve (AUC) maximization with fairness constraints. The new model distinguishes itself from classic minimax optimization problems in at least two aspects: (1) it deals explicitly with challenging expectation constraints and is a more realistic model; (2) it is flexible and well-suited to data-driven modeling. Then, we propose a class of efficient primal-dual algorithms to tackle the minimax expectation constrained problem. Without assuming bounded second moments of the generated dual iterates, we show that our algorithms achieve the optimal $cO(1/\sqrt{N})$ rate of convergence for objective optimality gap, duality gap, and feasibility residuals. We also demonstrate the practical efficiency of our algorithms by conducting numerical experiments on various large-scale real-world applications.

A Scalable Optimization Approach for the Multilinear System Arising from Scattered Data Interpolation – YANNAN CHEN, SOUTH CHINA NORMAL UNIVERSITY –

Scattered data interpolation aims to reconstruct a continuous (smooth) function that approximates the underlying function by fitting (meshless) data points. Here, “scattered” means that the data sites have no structure or order between their relative locations. There are extensive applications of scattered data interpolation in computer graphics, fluid dynamics, inverse kinematics, machine learning, etc. When data sites enjoy a well mesh geometry, plenty of methods such as wavelets, multivariate splines, and finite elements have been used for the interpolation problem. However, in the context of scattered data interpolation, meshless methods including radial basis functions and kernel-based approximations are promising. We focus on kernel-based methods. In this talk, we consider the novel generalized Mercer kernel in the reproducing kernel Banach space for

scattered data interpolation. The system of interpolation equations is formulated as a multilinear system with a structural tensor, which is an absolutely and uniformly convergent infinite series of symmetric rank-one tensors. Then we design a fast numerical method for computing products between the structural tensor and any vectors in arbitrary precision. Whereafter, a scalable optimization approach equipped with limited-memory BFGS and Wolfe line-search techniques is customized for solving these multilinear systems. Using the Lojasiewicz inequality, we prove that the proposed scalable optimization approach is a globally convergent algorithm and processes a linear or sub-linear convergence rate. Numerical experiments illustrate that the proposed scalable optimization approach can improve the accuracy of interpolation fitting and computational efficiency.

1f - ThC129 - 512C - Adaptive, Line-search, and Parameter-Free Stochastic Optimization
 II Chair: FRANCESCO ORABONA 

The Price of Adaptivity in Stochastic Convex Optimization – OLIVER HINDER, UNIVERSITY OF PITTSBURGH – Yair Carmon

We prove impossibility results for adaptivity in non-smooth stochastic convex optimization. Given a set of problem parameters we wish to adapt to, we define a “price of adaptivity” (PoA) that, roughly speaking, measures the multiplicative increase in suboptimality due to uncertainty in these parameters. When the initial distance to the optimum is unknown but a gradient norm bound is known, we show that the PoA is at least logarithmic for expected suboptimality, and double-logarithmic for median suboptimality. When there is uncertainty in both distance and gradient norm, we show that the PoA must be polynomial in the level of uncertainty. Our lower bounds nearly match existing upper bounds, and establish that there is no parameter-free lunch.

Coin Sampling: Learning-Rate-Free Optimization on the Space of Probability Measures – LOUIS SHARROCK, LANCASTER UNIVERSITY –

The task of sampling from a target probability distribution whose density is only known up to a normalisation constant is of fundamental importance to computational statistics and machine learning. There are various popular approaches to this task, including Markov chain Monte Carlo (MCMC) and variational inference (VI). More recently, there has been growing interest in developing hybrid sampling methods which combine the non-parametric nature of MCMC with the parametric approach used in VI. In particular, particle based variational inference (ParVI) methods approximate the target distribution using an ensemble of interacting particles, which are deterministically updated by iteratively minimising a metric such as the Kullback-Leibler divergence. Unfortunately, such methods invariably depend on hyperparameters such as the learning rate, which must be carefully tuned by practitioners in order to ensure convergence to the target distribution at a suitable rate.

In this talk, we introduce a suite of new sampling algorithms which are entirely learning rate free. Our approach leverages the perspective of sampling as an optimisation problem over the space of probability measures, and existing ideas from parameter free optimisation. We discuss how to establish the convergence of our algorithms under suitable assumptions

on the target distribution. We then illustrate the performance of our approach on a range of numerical examples, including several high dimensional models and datasets, demonstrating comparable performance to existing state-of-the-art sampling algorithms, but with no need to tune a learning rate.

Parameter-free projected gradient descent with application to contextual bandits with knapsacks – EVGENII CHZHEN, CNRS –

I will describe a fully parameter-free version of AdaGrad (norm), which is adaptive to the distance between the initialization and the optimum, and to the sum of the square norm of the subgradients. Describing an extension of our approach to stochastic optimization, I will highlight the utility of the approach for the problem of contextual bandits with knapsacks.

1f - FA51 - 512C - Splitting and Randomized Methods 3 Chair: CHUNFENG CUI 

Optimal and approximate optimal parameters of the SOR-like iteration method for solving absolute value equations – CAIRONG CHEN, FUJIAN NORMAL UNIVERSITY – Bo Huang, Dongmei Yu, Deren Han

The SOR-like iteration method is an efficient method for solving the NP-hard absolute value equations. However, the SOR-like iteration method is one-parameter-dependent and it is an important problem to determine the optimal iteration parameter. In this talk, we will revisit the convergence results of the SOR-like iteration method and give a new proof. The new proof exhibits some insights in determining the convergent region and the optimal iteration parameter. Along this line, the iteration-independent optimal parameter and the approximate optimal parameter are explored. Numerical results are presented to demonstrate our theoretical results.

A regularized splitting method for sums of multi-block maximal monotone operators and its relationship with symmetric ADMM – XINGJU CAI, NANJING NORMAL UNIVERSITY –

The monotone inclusion problem has wide applications in mathematical optimization and control. In this paper, we consider designing a splitting method for finding a zero point of $\sum_{i=1}^m A_i x + Bx$ ($m \geq 2$), where B is β - cocoercive. We introduce a new space reformulation technique to transform this problem into a three-operator problem in $\mathcal{H}^{m-1}$. Based on the proposed space reformulation technique, we present a regularized splitting method. This method involves the computation of B in a forward step and the parallel computation of the A_i 's in a backward step. We also prove the convergence of the proposed method and the sublinear convergence rate for the fixed point residuals under mild conditions in infinite-dimension Hilbert space. Our method extends the Douglas-Rachford splitting method and the Davis-Yin three operator splitting method. It is worth noting that symmetric ADMM is a special case of our method in the context of convex optimization. Finally, we apply our method to the mean-variance optimization problem, inverse problems in imaging, and the soft-margin support vector machine problem with nonsmooth hinge functions. We then compare its performance with the existing methods in the literature.

Extended alternating structure-adapted proximal gradient algorithm for multiblock non-convex optimization – CHUNFENG CUI, BEIHANG UNIVERSITY – Ying Gao, Wenxing Zhang, Deren Han

Alternating structure-adapted proximal (ASAP) gradient algorithm has garnered significant interest on solving nonconvex nonsmooth block-regularized problems recently. However, multiblock nonseparable structure hinders the applications of ASAP on far-reaching practical problems. In this paper, following the seminal work (M. Nikolova and P. Tan, SIAM J Optim, 29:2053-2078) of ASAP, we develop an extended ASAP for solving nonconvex nonsmooth problems with multiblock nonseparable structure. By exploiting the properties of partial subdifferential for nonsmooth alternating scheme, our developed convergence analysis covers a variety of nonsmooth nonseparable coupling functions. Thus expanding the applicability scope of the model, particularly in multimodal data fusion. Under the Aubin property on partial subdifferential mapping, the global convergence of extended ASAP is established based on Kurdyka-Lojasiewicz property. Furthermore, the sublinear convergence rate of extended ASAP is built upon the proximal point mapping algorithmic framework and some mild conditions. Numerical simulations on multimodal data fusion demonstrate the compelling performance of the proposed method.

1f - FB52 - 512C - **Splitting and Randomized Methods 4** Chair: JINGYA CHANG

A Recursive Multilevel Algorithm for Deep Learning – ISABEL JACOB, TECHNICAL UNIVERSITY DARMSTADT – Stefan Ulbrich

As the use cases for neural networks become increasingly complex, modern neural networks must also become deeper and more intricate to keep up, indicating the need for more efficient learning algorithms. Multilevel methods, traditionally used to solve differential equations using a hierarchy of discretizations, offer the potential to reduce computational effort.

In this talk, we combine both concepts and introduce a multilevel stochastic gradient descent algorithm that accelerates learning through a multilevel strategy. A gradient correction term is needed to establish first-order consistency. We discuss convergence of the method in the case of a deterministic gradient correction as well as a stochastic gradient correction under additional conditions including step size regularization and an angle condition.

To demonstrate the usefulness of our approach, we apply it to residual neural networks in image classification. The resolution of the images is utilized to generate data sets of varying complexity, which are then used to build a hierarchy of neural networks with a decreasing number of variables. Additionally, we construct corresponding prolongation and restriction operators. Numerical results are presented.

Improved variance reduction extragradient method with line search for stochastic variational inequalities – TING LI, BEIHANG UNIVERSITY – Xingju Cai, Yongzhong Song, Yumin Ma

In this paper, we investigate the numerical methods for solving stochastic variational inequalities. Using line search scheme, we propose an improved variance based stochastic extragradient method with different step sizes in the prediction and

correction steps. The range of correction step size which can guarantee the convergence is also given. For the initial line search step size of each iteration, an adaptive method is adopted. Rather than the same scale for each reduction, a proportional reduction related to the problem is used to meet the line search criteria. Under the assumptions of Lipschitz continuous, pseudo-monotone operator and independent identically distributed sampling, the iterative complexity and the oracle complexity are obtained. When estimating the upper bound of the second order moment of the martingale difference sequence, we give a more convenient and comprehensible proof instead of using the Burkholder-Davis-Gundy inequality. The proposed algorithm is applied to fractional programming problems and the l_2 regularized logistic regression problem. The numerical results demonstrate its superiority.

Differentially Private Federated Learning via Inexact ADMM with Multiple Local Updates – JINGYA CHANG, GUANGDONG UNIVERSITY OF TECHNOLOGY –

Differential privacy is a data privacy protection technique that can be applied to a federated learning model and is designed to maximize the protection of individual privacy when data is analyzed and processed. However, while ensuring strong data privacy, differential privacy technology hinders achieving higher learning performance. We propose an imprecise alternating direction multiplier algorithm with multiple local updates for federated learning. The random noise generated by the Gaussian distribution with variance varying with time and the random noise generated by the Laplacian distribution are added to the output of the subproblems during iteration. We show that our algorithm provides (ϵ, δ) -differential privacy and ϵ -differential privacy each iteration, respectively, where ϵ is the privacy budget controlled by the user, and the privacy parameter δ represents the probability of failure that does not meet the approximate differential privacy definition. The convergence property and convergence rate are analyzed under the condition that the objective function is smooth convex, non-smooth convex and strongly convex. Using MNIST data set for image classification, it is shown that our algorithm has higher accuracy and faster convergence rate than the existing DP algorithm when achieving the same data privacy level. The trade-off between classification accuracy and privacy protection are demonstrated numerically.

1g - MB302 - 515A - **Challenging DFO and SBO Applications 1** Chair: YVES LUCET

Optimization under hidden constraints and its application to the robust design of wind turbines – MORGANE MENZ, IFP ENERGIES NOUVELLES – Delphine Sinoquet, Miguel Munoz Zuniga

The robust design optimization of offshore wind turbines generally requires numerous computationally expensive code simulations associated to the input values of both design variables and environmental wind load variables. However, the simulators encounter simulation crashes due to convergence issues for some values of the environmental and design inputs. These failures correspond to a hidden constraint (not known in advance by the user) and might be as costly to evaluate as a feasible simulation. The latter failures must be used in a wise way to target feasible input areas and thus to avoid unnecessary costly simulation failure during the robust optimization itera-

tions. Therefore, we propose a methodology for derivative-free optimization in the presence of hidden constraints learned with a Gaussian Process Classifier. We also propose an enhanced strategy based on an enrichment criterion focused on improving the classifier when it is useful for optimization convergence. We will present our methodology and different algorithms benchmarks in the presence of simulation crashes on toy problems and a wind turbine design application.

Black-box optimization for the design of a jet plate for impingement cooling – FILIPPO MARINI, UNIVERSITY OF BOLOGNA – Lorenzo Cocchi, Margherita Porcelli, Elisa Riccietti

In this talk, we present an industrial application that exploits black-box optimization in the design of a cooling system for a nozzle in a gas turbine. We describe a black-box function that simulates an impingement cooling system starting from a well-known model that correlates the design features of the cooling system with efficiency parameters. Then we provide a definition of the constrained optimization problem and a solver based on the Brute Force Optimizer. Finally, we show the effectiveness of the proposed strategy on test cases of practical interest.

Blackbox optimization for origami-inspired multistable structures – LUCA BOISNEAULT, POLYTECHNIQUE MONTRÉAL – Charles Audet, David Melancon

Multistable mechanical systems exhibit more than one stable configuration where the elastic energy is locally minimized. To realize such systems, the art of origami has been proposed as a versatile platform to design deployable structures with both compact and functional stable states. Conceptually, a multistable origami motif is constructed of two-dimensional surfaces connected by one-dimensional fold lines. This leads to stable configurations exhibiting zero-energy local minima. Physically, origami-inspired structures are three-dimensional, comprising facets and hinges fabricated in a distinct stable state where residual stresses are minimized. This leads to the dominance of one stable state over others. To improve mechanical performance, one can solve the constrained optimization problem of maximizing the multistability of origami structures, defined as the amount of elastic energy required to switch between stable states, while ensuring materials used for the facets and hinges remain within their elastic regime. In this study, the Mesh Adaptive Direct Search (MADS) algorithm, a blackbox optimization technique, is used to solve the optimization problem. Initially, the bistable waterbomb-base origami motif is selected as a case-study to present the methodology. The elastic energy of this origami pattern under deployment is calculated via Finite Element (FE) simulations which serve as the blackbox in the MADS optimization loop. To validate the results, optimized waterbomb-base geometries are built via Fused Filament Fabrication and their response under loading is characterized experimentally on a Uniaxial Test Machine. The methodology is then extended to more complex origami patterns, such as the Kresling cylinder, as well as functional origami-inspired structures, including deployable shelters and multi-modal robotic arms. Ultimately, our method offers a general framework for optimizing multistability in mechanical systems, presenting opportunities for advancement across various engineering applications.

Convex optimization of vertical alignment of roads – YVES LUCET, UBC –

The road design problem is split between corridor selection, horizontal alignment (computing a satellite view of the road), vertical alignment (where to cut and where to fill to obtain a vertical profile of the road), and earthwork (what material to move from which section to what section). The vertical alignment problem consists in minimizing the road construction costs under regulatory and safety constraints. The variables include the amount of material to cut or fill at each section, and where to move materials from sections to sections. The problem may be solved as part of a 3D alignment optimization problem as a nonlinear discontinuous optimization problem usually tackled with heuristics (genetic algorithms, particle swarm...). When the horizontal alignment is fixed, for example when an existing road needs to be renovated, only the vertical alignment is optimized. In that case, the problem becomes a mixed-integer linear program.

We will present two approaches that result in a convex optimization problem. The first one computes side-slope angles to estimate volumes using trapezoids. The volumes are the limit of volumes computed using a mixed-integer linear programming model when the number of slabs tends to infinity. Unfortunately, the convexity of this model breaks down when multiple materials are considered. The second convex optimization model can handle multiple materials by directly approximating the side slope areas using a quadratic regression model. Both models are convex quadratically constrained quadratic programs solved by CPLEX or Gurobi.

1g - TA306 - 515C - Large-scale and Parallel DFO Chair: YOUSSEF DIOUANE 

A derivative-free approach to partitioned optimization – PIERRE YVES BOUCHET, POLYTECHNIQUE MONTRÉAL, GERAD – Charles Audet, Loïc Bourdin

We introduce the "partitioned optimization" framework to solve some large-scale derivative-free optimization problems. A typical usecase is when the objective function has an explicit structure consisting in a full-dimensional component simple to minimize and a challenging nonsmooth component depending on a few variables only (let us call them the nonsmooth variables). We partition the variables space into a continuum of hyperplanes, each consisting on the subspace leaving all the nonsmooth variables fixed to a given set of values. As a result, for each hyperplane, the restriction of the objective function to that hyperplane is smooth so a dedicated algorithm can compute a minimizer. Then, we use a derivative-free algorithm to identify the hyperplane which minimizes the minimum of the restriction of the objective function to that hyperplane. The "partitioned optimization" framework formalizes this strategy and generalizes its applicability to a broad class of large-scale derivative-free problems. Our framework is applicable to any problem such that, first, the difficult components have an explicit small-dimensional dependency and, second, the remaining components are accessible and may be minimized at an acceptable cost. Some numerical examples are provided to highlight a gain in performance. In particular, we apply the framework to solve a class of discontinuous optimal control problems that makes usual strategies from the literature to fail.

Blockwise direct search method – HAITIAN LI, POLYU – Zaikun Zhang

Direct search methods are one class of derivative-free optimization algorithms that evaluate the objective function only based on the comparison of function values. We introduce a new framework of direct search method called blockwise direct search (BDS), which divides the searching set into blocks. For every iteration, each block has its step size. We develop this framework into a solver, which is open-source and easy to use. In numerical experiments, we observe that BDS can also be compared with model-based methods in some cases. In addition, our solver shows its efficiency and stability under noise without introducing specific techniques. BDS can also be used to tune the hyperparameters itself to improve the performance.

Using concurrent objective evaluations in derivative-free optimization – JEFFREY LARSON, ARGONNE NATIONAL LABORATORY –

We consider the case where concurrent evaluations of an objective function can be performed. We are particularly interested in scenarios where the objective function is expensive to evaluate, with each evaluation possibly using considerable parallel computing resources or relying on evaluating a real-world system or experiment. As such, the number of parameters being optimized is relatively few, as is the number of concurrent evaluations that can be performed. We conduct extensive numerical tests to study the utilization of such concurrency for both direct-search and model-based trust-region methods. We demonstrate how increasing levels of concurrency affect the efficiency of these algorithms. We investigate the impact of varying levels of concurrency and how they affect the observed performance of these algorithms. Numerical tests are run on a collection of synthetic benchmark problems and real-world simulations. Additionally, we propose a set of guidelines for choosing between direct-search and model-based methods based on the problem characteristics and available computational resources. The utilization of parallel computing frameworks in our approach showcases a practical and scalable solution for complex simulation optimization problems. Application problems include beamline design and operation in high-energy physics as well as the design of stellarator and tokamak fusion power plants.

A direct-search method for decentralized derivative-free optimization – YOUSSEF DIOUANE, POLYTECHNIQUE MONTRÉAL – Elhoucine Bergou, Vyacheslav Kungurtsev, Clément Royer

Derivative-free algorithms are particularly useful to tackle optimization problems in which the function to optimize arises from complex, expensive procedures that prevent from applying classical, derivative-based algorithms. In data-driven systems, this can manifest by having to perform calculations stored on different computer nodes, which significantly hardens the optimization process.

In this talk, we will present a direct-search framework for minimizing a sum of functions that are dispatched across a network of communicating agents. The proposed method combines the use of positive spanning sets, a standard feature of direct-search techniques, together with stepsize sequences that are prone to decentralized optimization. We establish global convergence of the proposed algorithm, and provide a numer-

ical comparison illustrating the benefit of our approach over strategies that explicitly attempt to approximate derivatives.

1g - TB322 - 515C - Recent Progresses in DFO

1 Chair: GIAMPAOLO LIUZZI

The cosine measure relative to a subspace – GABRIEL JARRY BOLDOC, SAINT FRANCIS XAVIER UNIVERSITY – Charles Audet, Warren Hare

The cosine measure is a tool that tells you how well a set of directions is covering the space R^n . In this talk, we extend the concept of cosine measure by defining the cosine measure relative to a subspace. This novel definition might be useful for subspace decomposition optimization methods. We propose a deterministic algorithm to compute it and discuss the situation in which the set of directions is infinite.

A (forgotten) subspace framework for large-scale derivative-free optimization – ZAIKUN ZHANG, THE HONG KONG POLYTECHNIC UNIVERSITY –

We re-introduce a subspace framework for solving problems of 10^4 variables without using derivatives. Studied in Sec. 5.3 of [Zhang, On Derivative-Free Optimization Methods (in Chinese), PhD thesis, CAS, 2012] and presented in ICCOPT 2013 (Lisbon), it remains nearly unknown to the community. It was implemented by Zhang in MATLAB in 2011, ported to Module-3 by Nystroem (Intel) in 2017, and included in cm3 in 2019 (<https://github.com/modula3/cm3/blob/master/caltech-other/newuoa/src/NewUOAs.m3>).

Parallel versions of the mesh adaptive direct search algorithm – SAMUEL MENDOZA, POLYTECHNIQUE MONTRÉAL – Sébastien Le Digabel, Antoine Lesage Landry

This presentation surveys the different parallel variants of the mesh adaptive direct search (mads) algorithm for constrained blackbox optimization. These problems can inherently imply high computational costs due to the possible large number of variables and multi-modality of the search space. In addition, the potential time-intensive nature and time heterogeneity of the blackboxes defining the problem prompts the need for efficient implementations. With the increasing use of high-performance computing, parallelism emerges as an actionable solution to mitigate computation time. The reviewed methods employ diverse levels of parallelism and distinct parallel strategies to effectively tackle each aspect outlined above. The presentation details the practical implementations, provides computational results, and offers insights into the advantages and limitations of each mads parallel method.

Worst case complexity bounds for linesearch-type derivative-free algorithms – GIAMPAOLO LIUZZI, DIAG SAPIENZA UNIVERSITY OF ROME – Andrea Brilli, Morteza Kimiaei, Stefano Lucidi

In this talk we are concerned with derivative-free methods for the solution of unconstrained black-box optimization problems. We show that derivative-free algorithms which are based on a linesearch-type extrapolation technique with sufficient decrease have the same worst case complexity proved for direct search methods. Furthermore, thanks to the linesearch approach with sufficient decrease, they also have the property that the number of iterations (in the worst case) for which the norm of the

gradient is above a given threshold is of the order of ϵ^{-2} . This last property considerably enriches the worst case analysis of derivative-free algorithm and, to the best of our knowledge, is new in this context. It is worth noticing that the property characterizes the behavior of the derivative-free algorithm better than the usual complexity result. Indeed, typical complexity results give the number of iterations required to drive the norm of the gradient below a prefixed tolerance for the first time. If we let the method run, the norm of the gradient might well rise above the tolerance again. The property we prove in this paper indicates that the total number of iterations with a gradient norm above a specified tolerance is bounded by a constant that depends on ϵ^{-2} .

1g - TC305 - 515C - **Multiobjective DFO Chair:**
LUDOVIC SALOMON 

An inexact restoration direct multisearch filter approach to constrained optimization – EVERTON SILVA, NOVA SCHOOL OF SCIENCE AND TECHNOLOGY – Ana Luisa Custodio

Direct Multisearch (DMS) is a class of methods for multiobjective derivative-free optimization that has a well-established convergence analysis and competitive computational implementations, being often used as benchmark for new algorithms or in practical applications. From a theoretical point of view, DMS was developed for continuous optimization with general constraints, using an extreme barrier approach where only feasible points are evaluated. In this work, we propose the integration of an inexact restoration filter approach in DMS, to address optimization problems with general constraints. Like in any filter approach, violations of the relaxable constraints are addressed as an additional objective that needs to be minimized. The inexact restoration approach attempts to recover feasibility when the poll center is infeasible. Under mild assumptions, we prove that the so-called DMS-FILTER-IR algorithm generates feasible and/or infeasible subsequences of iterates that converge to either a Pareto stationary point, in the feasible case, or to a Pareto stationary point for the problem that only considers the unrelaxable constraints, potentially serving as a Pareto stationary point of the original problem, in the infeasible case. We will detail the proposed algorithm, provide theoretical results on convergence, and report numerical experiments that state the good performance of this approach to address multiobjective problems with general constraints.

A comparison of search strategies for multiobjective direct search algorithm – LUDOVIC SALOMON, POLYTECHNIQUE MONTRÉAL – Christophe Tribes, Sébastien Le Digabel

The last decade has seen the development of new convergence-based derivative-free multiobjective algorithms. Most of them are extensions of reliable single-objective methods. Direct search algorithms are iterative procedures that alternate between search and poll steps. The poll, a local step around the current incumbent, guarantees their convergence. The search step is optional. However, well-designed search strategies can significantly improve the performance of these methods. Search strategies have been extensively studied for single-objective optimization. The proposal of such strategies for multiobjective optimization is in its infancy. In this talk, we propose several extensions of search strategies for multiobjective optimization: once is based on the quadratic models. Numerical experiments

illustrate the performance of these new approaches.

Gradient-based and derivative-free multi-objective optimization via iterative single-objective optimization: MO-BFGS and MO-SLSQP – DIMO BROCKHOFF, INRIA AND IP PARIS –

Many multiobjective solvers approach the simultaneous optimization of two or more objective functions by solving varying single-objective formulations (such as a weighted sum of the objective values) iteratively. Their scalar formulations of the multiobjective problem are thereby typically independent of any solutions found and not easy to set in practice (in particular when solving non-linear objective functions). Evolutionary multiobjective optimization (EMO) algorithms, on the contrary, typically "evolve" a fixed set of μ solutions (aka their "population") in parallel towards the Pareto set—exploiting information about already explored search space regions in terms of set-dominance and diversity criteria and without the need of a priori information from a human.

In this talk, I will discuss another possibility: optimizing iteratively a set-based scalarizing function which dynamically changes with the increasing set of found solutions and which is based on the hypervolume indicator. This allows to use efficient single-objective solvers without the need to specify (human) preferences in classical scalarizing approaches. In numerical experiments with SLSQP and L-BFGS as single-objective solvers, we observe how the new iterative approach performs in both gradient-based and derivative-free mode on a variety of test functions and compare the two resulting algorithms MO-SLSQP and MO-BFGS to a multitude of previously ran algorithms.

This is joint work with Anne Auger, Nikolaus Hansen, and Baptiste Plaquevent-Jourdain.

1g - WA309 - 515C - **Multifidelity and Bilevel DFO Chair:** FRANCESCO RINALDI 

Multifidelity Monte Carlo estimates for accelerating stochastic optimization with trust regions – YUNSOO HA, NATIONAL RENEWABLE ENERGY LAB – Julianne Mueller

In recent years, multifidelity stochastic optimization problems have gained significant attention as they arise in a wide variety of application areas including virtual biofuel engineering, the design of autonomous laboratory experiments, controlling the energy demand of buildings, multi-scale modeling for scale-up challenges, and many more. Multifidelity stochastic optimization has the capacity to reduce the computational workload by exploiting information from lower-fidelity models during the optimization of the high-fidelity target. However, there are still persistent challenges to contend with, including the crucial issue of determining the sample sizes to obtain accurate function estimates for the optimization task while operating within the constraints of a limited computational budget in the context of multifidelity models. In this talk, we present our ongoing work in designing a sample-efficient trust region algorithm utilizing multifidelity Monte Carlo estimates with the aim of achieving almost sure convergence to the first-order critical point. We will show how the proposed sampling method can enable us to optimally adjust the fidelity of the function value estimates, thus minimizing the computational resource requirements. We

will also present preliminary results of our research on benchmark problems and compare with state of the art stochastic optimization approaches.

Hierarchically constrained multi-fidelity blackbox optimization – XAVIER LEBEUF, POLYTECHNIQUE MONTRÉAL – Charles Audet, Sébastien Le Digabel, Miguel Diago, Stéphane Alarie

This work introduces a novel multi-fidelity blackbox optimization algorithm designed to alleviate the resource-intensive task of evaluating infeasible points. This algorithm is an intermediary component bridging a direct search solver and a blackbox, resulting in reduced computation time per evaluation, all while preserving the efficiency and convergence properties of the chosen solver. This is made possible by assessing feasibility through a broad range of fidelities, leveraging information from cost-effective evaluations before committing to a full computation. These feasibility estimations are generated through a hierarchical evaluation of constraints, tailored to the multi-fidelity nature of the blackbox problem, and defined by a biadjacency matrix, for which we propose a construction. A series of computational tests using the NOMAD solver on the solar family of blackbox problems are conducted to validate the approach. The results show a significant improvement in solution quality when an initial feasible starting point is known in advance of the optimization process. When this condition is not met, the outcomes are contingent upon certain properties of the blackbox.

Zeroth-order bilevel optimization – SAEED GHADIMI, UNIVERSITY OF WATERLOO – Alireza Aghasi

Studying bilevel optimization problems has attracted much interest over the past few years as they arising in several application domains such as automated machine learning and hyperparameter tuning. The main challenge in designing algorithms for bilevel problems is to estimate the gradient of upper level function which is itself a function of the gradient of the optimal solution to the lower problem. This estimation involves calculating blocks of the Hessian matrix of the lower level problem and their inverse which is computationally expensive. To overcome this computational challenge, fully first-order and partial zeroth-order methods have been developed which still require some unbiased gradient estimators. In this talk, we present a fully zeroth-order stochastic approximation algorithm for solving stochastic bilevel optimization problems assuming that neither the upper/lower loss functions, nor their unbiased gradient estimates are available. To do so, we first generalize the Gaussian convolution technique to the functions with two block variables and establish all corresponding relationships between such functions and their smoothed Gaussian approximations. Using two independent smoothing parameters for each block, will give us the flexibility of exploiting zeroth-order derivative estimates over just one block. Using these results, we estimate the first- and second-order derivatives, provide our zeroth-order double-loop algorithm, and establish its non-asymptotic convergence analysis.

Inexact direct-search methods for bilevel optimization problems – FRANCESCO RINALDI, UNIVERSITY OF PADOVA – Youssef Diouane, Vyacheslav Kungurtsev, Damiano Zeffiro

In this work, we introduce new direct-search schemes for the

solution of bilevel optimization (BO) problems. Our methods rely on a fixed accuracy blackbox oracle for the lower-level problem, and deal both with smooth and potentially nonsmooth true objectives. We thus analyze for the first time in the literature direct-search schemes in these settings, giving convergence guarantees to approximate stationary points, as well as complexity bounds in the smooth case. We also propose the first adaptation of mesh adaptive direct-search schemes for BO. Some preliminary numerical results on a standard set of bilevel optimization problems show the effectiveness of our new approaches.

1g - WC323 - 515C - **Recent Progresses in DFO**
2 Chair: ANA LUISA CUSTODIO 

Improving MADS search step with a quadratic programming solver – CHRISTOPHE TRIBES, POLYTECHNIQUE MONTRÉAL – Ludovic Salomon, Tangi Migot

Our presentation will discuss improvements in the resolution of constrained blackbox optimization (BBO) problems using the mesh adaptive direct search (MADS) algorithm implemented in the Nomad software package. MADS alternates between Search and Poll steps to generate trial points. The presentation focuses on the quadratic model search. This search uses a subset of all the previously obtained blackbox evaluations to build function models and solves an optimization sub-problem. This can be done with the MADS algorithm but depending on the problem at hand the results are not satisfactory. In particular, the COCO platform BBOB-constrained Sphere problem in 10 dimensions with one linear inequality constraint has a single known optimum. However, MADS converges very slowly to the sub-problem optimum. This work focuses on combining a Quadratic Programming (QP) solver and MADS for the sub-problem. We will present the candidate QP solvers and their implementation in Nomad. Results showing the benefits of combining a QP solver and MADS to will be given.

Advancing derivative-free optimization: exploring logarithmic barrier and exterior penalty in direct search – ANDREA BRILLI, SAPIENZA UNIVERSITY OF ROME – Giampaolo Liuzzi, Ana Luisa Custodio, Everton Silva

This research explores an approach to addressing mathematical optimization problems characterized by nonlinear objective functions and constraints, particularly in situations where derivatives are unavailable. A Sequential Penalty Approach is introduced, incorporating a Logarithmic Barrier for managing inequality constraints and an Exterior Penalty for equality constraints. Traditionally, optimization problems without access to derivative information pose significant challenges. Our proposed method aims to overcome these hurdles by introducing a sequential penalty framework that effectively addresses both inequality and equality constraints. The Logarithmic Barrier ensures a smooth handling of inequality constraints, while the Exterior Penalty seamlessly accommodates equality constraints, providing a comprehensive solution for a wide range of nonlinear optimization challenges. This strategy, proposed for the first time in a derivative-free setting in the framework of linesearch methods, is adapted and incorporated into SID-PSM algorithm, a generalized pattern search method, allowing to effectively handle general nonlinear constraints. Under reasonable assumptions regarding the smoothness of the functions,

convergence is established, without any convexity assumptions. Empirical validation is conducted through experimentation on a diverse set of optimization problems. The results corroborate not only the efficacy of the proposed approach but also its practical utility. The direct search algorithm, when complemented by the Logarithmic Barrier and Exterior Penalty, consistently exhibits promising numerical results, demonstrating the method's potential.

A hybrid direct search and projected simplex gradient method for convex constrained optimization – ANA LUISA CUSTODIO, NOVA SCHOOL OF SCIENCE AND TECHNOLOGY – Evelin Krulikowski, Marcos Raydan

In this work, we propose a new Derivative-free Optimization (DFO) approach for solving convex constrained minimization problems. The feasible set is assumed to be the nonempty intersection of a finite collection of closed convex sets, such that the projection onto each of these individual convex sets is simple and inexpensive to compute. The proposed iterative approach alternates between steps that use Directional Direct Search (DDS), considering adequate poll directions, and a Spectral Projected Gradient (SPG) method, replacing the real gradient by a simplex gradient, under a DFO approach. The simplex gradient is computed at no extra cost in terms of functions evaluations, by reusing previously evaluated points. In the SPG steps, if the convex feasible set is simple, then a direct projection is computed. If the feasible set is the intersection of finitely many convex simple sets, then Dykstra's alternating projection method is applied. Convergence properties are established under standard assumptions usually associated to DFO methods. Some preliminary numerical experiments are reported to illustrate the performance of the proposed algorithm, in particular by comparing it with a classical DDS method. Our results indicate that the hybrid algorithm is a robust and effective approach for derivative-free convex constrained optimization.

1g - ThA325 - 515C - **DFO based on Approximations to Derivatives 1** Chair: STEFAN M. WILD



Is building first-order simulation oracles really worth it? – SARA SHASHAANI, NORTH CAROLINA STATE UNIVERSITY – Raghu Pasupathy, Yunsoo Ha
Stochastic simulations or black-box noisy (non-convex) problems often do not provide direct gradient observations. However, conventional wisdom asserts that having gradients should lead to better convergence and complexity results. This knowledge has motivated work towards building first-order simulation oracles that compute function sample path derivatives by employing infinitesimal perturbation analysis or likelihood ratio methods. These first-order simulation oracles can be costly and inflexible, adding an eligibility requirement to deem a function suitable for derivative-based models. The natural question is, is it worth the trouble building the first-order capability in simulation oracles and will that guarantee better complexity? We find that the answer is not necessarily. Concretely, we show that unless the function sample paths are themselves smooth with uniformly bounded Lipschitz constants, there is no gain in having a first-order simulation oracle. That is, stochastic derivative-free algorithms that do not need or use direct gra-

dients will lead to the same complexity results that one would attain with stochastic derivative-based algorithms. This finding is important in many cases where the smoothness of the stochastic process is clearly lost. An example is when the objective functions are quantiles or conditional expectations rendering the sample path functions piecewise linear or step functions. In such settings, our results imply that a derivative-free solver will be just as efficient as a derivative-based solver. The main vehicle for our analysis and result is common random numbers and the consequent reduced variance in gradient and function estimates.

Adaptive stochastic subspace descent – KILLIAN WOOD, UNIVERSITY OF COLORADO BOULDER – Stephen Becker

Many optimization problems, such as hyper-parameter learning and PDE constrained optimization, have costly objective functions for which the gradient is unavailable. We propose a novel descent method, Adaptive Stochastic Subspace Descent (A-SSD), which uses randomized directional derivative approximates combined with an adaptive Polyak step-size policy. We show that the step-size policy is the optimal solution to a truncated linear model over the random subspace. Finally, we demonstrate empirically that A-SSD is competitive against commonly used derivative-free methods such as BOBYQA and conjugate directions for functions in dimensions exceeding 100.

Self-Correcting DFO Algorithm: An Attempt at a Convergence Rate – JIAQI WANG, CORNELL UNIVERSITY –

We will discuss steps complexity analysis (in terms of function evaluations) of model-based trust region methods. The one method for which complex has been established previously is from the paper "Trust-Region Methods Without Using Derivatives: Worst Case Complexity and the NonSmooth Case" by Garmanjani, Judice and Vicente. It employs a criticality step where instead of trying to make progress the algorithm's focus is shifted to generating a fully linear model. In contrast we study alternative methods where the algorithm always tries to make progress and accepts steps if such progress happens to occur. One of the methods, makes on model improvement iteration, when necessary, following methodology by Powell. The other method maintains geometry via self-correcting property following the paper "Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization" by Scheinberg and Toint. We will also provide numerical comparison of these methods.

Derivative-free stochastic optimization via adaptive sampling strategies – STEFAN M. WILD, LAWRENCE BERKELEY NATIONAL LABORATORY – Raghu Bollapragada, Cem Karamanli

We present a novel derivative-free optimization framework for solving unconstrained stochastic optimization problems. Many problems in fields ranging from simulation optimization to reinforcement learning involve settings where only stochastic function values are obtained via an oracle with no available gradient information, necessitating the usage of derivative-free optimization methodologies. Our approach involves estimating gradients using stochastic function evaluations and integrating adaptive sampling techniques to control the accuracy in these stochastic approximations. We consider various gradient estimation techniques including finite difference, Gaussian smooth-

ing, randomized coordinate finite difference, and randomized subspace finite difference methods. We provide theoretical convergence guarantees for our framework and analyze the worst-case iteration and sample complexities associated with each gradient estimation method. Finally, we demonstrate the empirical performance of the methods on logistic regression and nonlinear least squares problems.

1g - ThB300 - 515C - **Stochastic DFO 1 Chair:**
MATT MENICKELLY 

Convergence analysis of CMA-ES – ARMAND GISSLER, ECOLE POLYTECHNIQUE – Anne Auger, Nikolaus Hansen

The Covariance Matrix Adaptation Evolution Strategies (CMA-ES) is a black-box derivative-free stochastic optimization method that has been widely applied across various domains. Despite empirical evidence of its linear convergence to the solution of many optimization problems, including the minimization of nonsmooth nonconvex multimodal ill-conditioned functions, a mathematical proof was lacking. I will present recent advancements that aim to fill this gap. By normalizing the states of CMA-ES while minimizing a spherical function (monotonously transformed Euclidean norm), we define a time-homogeneous Markov chain that—when setting correctly the hyperparameters of CMA-ES—converges at a geometric rate towards a stationary probability distribution. A law of large numbers concludes our convergence proof and provides a convergence rate for spherical objective functions. Due to invariance properties of the algorithm, this result applies as well to ellipsoid (monotonously transformed convex-quadratic) objective functions, for which the covariance matrix learns second-order information through the approximation of the inverse Hessian.

Q-fully quadratic modeling and its application in a random subspace derivative-free method – YIWEN CHEN, UNIVERSITY OF BRITISH COLUMBIA – Warren Hare, Amy Wiebe

Model-based derivative-free optimization (DFO) methods are an important class of DFO methods that are known to struggle with solving high-dimensional optimization problems. Recent research has shown that incorporating random subspaces into model-based DFO methods has the potential to improve their performance on high-dimensional problems. However, most of the current theoretical and practical results are based on linear approximation models due to the complexity of quadratic approximation models. In this talk, we propose a random subspace derivative-free trust-region algorithm based on quadratic approximations. Unlike most of its precursors, this algorithm does not require any special form of objective function. We study the geometry of sample sets, the error bounds for approximations, and the quality of subspaces. In particular, we provide a technique to construct Q-fully quadratic models, which is easy to analyze and implement. We present an almost-sure global convergence result of our algorithm and give an upper bound on the expected number of iterations to find a sufficiently small gradient. We also develop numerical experiments to compare the performance of our algorithm using both linear and quadratic approximation models. The numerical results demonstrate the strengths and weaknesses of using quadratic approximations.

A new two-dimensional model-based subspace method for large-scale unconstrained derivative-free optimization: 2D-MoSub – PENGCHENG XIE, CHINESE ACADEMY OF SCIENCES – Ya Xiang Yuan

Derivative-free optimization methods have gained attention due to their applicability in scenarios where derivative information is either unavailable or expensive to compute. For existing techniques for model-based derivative-free optimization, the large-scale problems are still bottlenecks. One important approach is the use of subspace methods, which involves minimizing the objective function in a low-dimensional subspace at each iteration to obtain the next iteration point. This talk will introduce our method 2D-MoSub (2-dimensional model-based subspace method), which is a novel derivative-free optimization method based on the subspace method for general unconstrained optimization and especially aims to solve large-scale derivative-free optimization problems. 2D-MoSub combines 2-dimensional quadratic interpolation models and trust-region techniques to iteratively update the points and explore the 2-dimensional subspace. 2D-MoSub's framework includes initialization, constructing the interpolation set, building the quadratic interpolation model, performing trust-region trial steps, and updating the trust-region radius and subspace. This talk will introduce the framework and computational details of 2D-MoSub, and discuss the poisedness and quality of the interpolation set on the corresponding 2-dimensional subspace. This talk will also analyze some properties of 2D-MoSub, including the model's approximation error with projection property and the algorithm's convergence. Numerical results demonstrate the effectiveness and efficiency of 2D-MoSub for solving a variety of unconstrained optimization problems.

Noise-aware derivative-free optimization – MATT MENICKELLY, ARGONNE NATIONAL LABORATORY – Jeffrey Larson, Jiahao Shi

In this work, we develop and analyze a noise-aware model-based trust-region (MBTR) method for derivative-free optimization. By a noise-aware algorithm, we refer to any algorithm that takes (a dynamic) estimate of noise present in blackbox function values. While we intend for the definition of noise to be fairly encompassing, the theory on which we build especially considers cases where the noise can be bounded in absolute value, or where the noise is generated from some distribution exhibiting a subexponential tail. Our algorithm builds on standard MBTR methods. However, guided by theoretical convergence results, our method explicitly decouples the trust-region radius parameter from the model-building process, and we additionally pay far more attention to the geometry of interpolation sets used in model-building than do standard implementations of MBTR methods. Our convergence results provide high-probability rates to neighborhoods of local optima, and an implementation of our method does quite well compared to derivative-free methods that are not explicitly noise-aware. This work was particularly motivated by variational quantum algorithms (VQAs), a paradigm of hybrid quantum-classical computing, in which measurements acquired on a quantum device are effectively treated as a blackbox function. In the VQA setting, the noise present in these blackbox function values is an aggregation of both stochastic noise arising from probabilistic realizations of state vectors, as well as non-negligible hardware noise from near-term quantum devices that is gener-

ally independent of the optimization process. We demonstrate the utility of our novel method on optimization problems that arise in VQAs.

1g - ThC303 - 515C - **Challenging DFO and SBO Applications 2** Chair: EDWARD HALLÉ HANNAN

Management of categorical variables in a mixed-variable (surrogate-based) evolutionary algorithm (Part I) – SYLVERIO POOL MARQUEZ, UNIVERSITÉ DE NAMUR – Charlotte Beauthier, Rajan Filomeno Coelho, Caroline Sainvitu, Annick Sartenauer

The development of methods combining numerical simulation tools and advanced optimization algorithms plays a crucial role in exploring new concepts. In practice, various industrial design problems have continuous, discrete, and categorical variables. From an engineering point of view, the specific case of categorical variables is of great practical interest by its ability to represent the choice of a material or an engine architecture, the shape of a cross-section, etc. Our objective is to manage efficiently and concurrently these different types of variables within optimization processes using either an evolutionary algorithm or a combination of surrogate models with this algorithm. The developments presented are carried out in Minamo, the in-house design space exploration and multi-disciplinary optimization platform of Cenaero.

In this first part, a mixed-variable encoding, first introduced in [Wang et al., 2021] for a particle swarm optimization (PSO) algorithm, is adapted for an evolutionary algorithm (EA). In short, each possibility, named "attribute", of a categorical variable is associated to a real value defined in $[0,1]$. This value is updated during the optimization process and is exploited when the genetic operators are applied on the categorical variables involved. The higher the value is for an attribute, the more this attribute will be favored by such operators. The main difference with the method proposed in [Wang et al., 2021] is the use of these values inside the crossover operator of a EA. Additionally, new ways to update the values associated with each attribute are introduced by taking the fitness of the points sharing this attribute also into account.

In order to analyze the influence of the different techniques proposed, a benchmark is presented where the considered test problems come from structural design problems. This set of test problems is used to compare the performance of the different variants proposed to handle categorical variables with the current EA used within Minamo. Moreover, a PSO algorithm with the original method proposed in [Wang et al., 2021] is added to this benchmark to analyze the difference with the optimization performance coming from the proposed methods for the EA. Finally, the open-source algorithm NOMAD [Le Digabel, 2011] is also included in this benchmark.

Management of categorical variables in a mixed-variable (surrogate-based) evolutionary algorithm (Part II) – CHARLOTTE BEAUTHIER, CENAERO – Rajan Filomeno Coelho, Sylverio Pool Marquez, Caroline Sainvitu, Annick Sartenauer

The development of methods combining numerical simulation

tools and advanced optimization algorithms plays a crucial role in exploring new concepts. In practice, various industrial design problems have continuous, discrete, and categorical variables. From an engineering point of view, the specific case of categorical variables is of great practical interest by its ability to represent the choice of a material or an engine architecture, the shape of a cross-section, etc. Our objective is to manage efficiently and concurrently these different types of variables within optimization processes using either a genetic algorithm or a combination of surrogate models with this algorithm. The developments presented are carried out in Minamo, the in-house design space exploration and multi-disciplinary optimization platform of the research center Cenaero. In this second part, a refinement of the notion of distance between variables is proposed. Indeed, since the construction of the surrogate models and the optimization search require the comparison of points with variables of different types, an adequate approach is to define a specific distance, namely here the Heterogeneous Euclidean-Overlap Metric, as proposed in [McCane and Albert, 2008] and [Wilson and Martinez, 1997]. In our approach, we propose to redefine the notion of distance between the possible string values of a categorical variable (named "attributes"), through the original concept of "affinity". The notion of affinities between attributes can be interpreted as a weighted relationship between attributes. These affinities are usually defined based on a physical intuition of the designer, but we also consider here how to set these affinities mathematically, using clustering or projection methods. Indeed, affinities are generally implicitly associated to the behavior of one or several outputs (objectives or constraints) that behave(s) similarly for various attributes. Some attributes can therefore be declared as "close to each other" because they have a similar impact on one or more quantities (even if these attributes can also be far away for other quantities). We therefore slightly modify the overlap distance accordingly, by allowing the notion of affinities between attributes of categorical variables. In order to study the impact of the use of affinities (on the quality of the trained surrogate models or on the convergence performance in a surrogate-based optimization process), specific test problems coming from structural design framework (on which these affinities can be defined) are studied, and numerical results are presented.

A graph-structured distance for heterogeneous datasets with meta variables – EDWARD HALLÉ HANNAN, GERAD, POLYTECHNIQUE MONTRÉAL – Charles Audet, Sébastien Le Digabel, Youssef Diouane

Heterogeneous datasets emerge in various machine learning or optimization applications that feature different data sources, various data types and complex interrelationships between variables. In practice, heterogeneous datasets are often partitioned into smaller well-behaved ones that are easier to process. However, some applications involve expensive-to-generate or limited size datasets, which motivates methods that utilize heterogeneous datasets in their entirety. This last remark is particularly important for blackbox (or simulation-based) optimization that tackles objective functions and constraints that may require hours, or even days, to evaluate. The first main contribution of this work is a modelling graph-structured framework that generalizes state-of-the-art hierarchical, tree-structured, or variable-size frameworks. This framework models domains

that involve heterogeneous datasets in which variables may be continuous, integer, or categorical, with some identified as meta if their values determine the inclusion/exclusion or affect constraints of other so-called decreed variables. Excluded variables are introduced to manage variables that are included in some points, but excluded in others. The second main contribution is the graph-structured distance that compares extended points with any combination of included and excluded variables: any pair of points can be compared, allowing to work directly in heterogeneous datasets with meta variables. The contributions are illustrated with some regression experiments, in which the performance of a multilayer perceptron w.r.t. to its hyperparameters is modeled with inverse distance weighting and K-nearest neighbors models.

1g - FA326 - 515C - **DFO based on Approximations to Derivatives 2** Chair: ANNE AUGER ←

Black-Box Optimization Algorithms for Problems with Convex Regularizers – YANJUN LIU, PRINCETON UNIVERSITY – Lindon Roberts

In this talk, we present a derivative-free trust region framework for problems which are the sum of a smooth nonconvex black-box function and a non-smooth convex regularizer. We assume the regularizer is known and has an easy-to-evaluate proximal operator. Our method, with global convergence and worst-case complexity guarantees, is more practical than existing methods for these problems as it allows inexact calculations of stationary measures (which do not have a closed-form expression). When specialized to the case of regularized nonlinear least-squares problems, we provide implementable procedures for computing inexact stationary measures and achieving the required trust-region subproblem accuracy. We show numerical results for nonlinear least-squares with L1-norm regularization.

Zeroth-order low-rank Hessian estimation via matrix recovery – TIANYU WANG, FUDAN UNIVERSITY – Zicheng Wang, Jiajia Yu

A zeroth-order Hessian estimator aims to recover the Hessian matrix of an objective function at any given point, using as few finite-difference computations as possible. We study zeroth-order Hessian estimation for low-rank Hessians, from a matrix recovery perspective. Our challenge lies in the fact that traditional matrix recovery techniques are not directly suitable for our scenario. They either demand incoherence assumptions (or its variants), or require an impractical number of finite-difference computations in our setting. To overcome these hurdles, we employ zeroth-order Hessian estimations aligned with proper matrix measurements, and prove new recovery guarantees for these estimators. More specifically, we prove that for a Hessian matrix $H \in \mathbb{R}^{n \times n}$ of rank r , $\mathcal{O}(nr^2 \log^2 n)$ proper zeroth-order finite-difference computations ensure a highly probable exact recovery of H . Compared to existing methods, our method can greatly reduce the number of finite-difference computations, and does not require any incoherence assumptions.

Derivative-free implementation of the regularized Newton method with lazy approximated Hessians – GEOVANI GRAPIGLIA, UCLOUVAIN –

In this work, we develop a zeroth-order (derivative-free) implementation of the cubically regularized Newton method for solv-

ing general non-convex optimization problems. Specifically, we consider the class of unconstrained problems whose objective functions are twice continuously differentiable with Lipschitz continuous Hessians. The proposed method employs finite difference approximations of gradient vectors and Hessian matrices. We use a special adaptive search procedure in our algorithms, which simultaneously fits both the regularization parameters of the models and the stepsizes of the finite difference approximations. Thanks to this procedure our scheme does not require the knowledge of the actual Lipschitz constants. Additionally, we equip our algorithm with the lazy Hessian update, meaning that the method reuses a previously computed Hessian approximation matrix for several iterations. In terms of worst-case complexity, we prove that the proposed method is able to find an ϵ -approximate stationary point using at most $\mathcal{O}(n^{3/2} \epsilon^{-3/2})$ function evaluations, where n is the dimension of the problem. This complexity bound improves existing bounds in terms of the joint dependence on n and ϵ for derivative-free methods applied to the minimization of functions with Lipschitz continuous Hessians. Illustrative numerical results confirm our theoretical findings.

On the connection between affine-invariance, convergence and learning second order information – ANNE AUGER, INRIA AND ECOLE POLYTECHNIQUE –

Invariance is a general concept that is fundamental in many domains. In statistics, machine learning, decisions taken as a result of an algorithm based on data should not be affected by simple transformations on the input data like reordering or translation (as long as the changes on the data are reflected in the outcome of the decision). In optimization, the performance of an algorithm on a function should not be greatly affected if this function is translated, scaled or even transformed by an affine transformation of the search space. Invariance is particularly important as it guarantees generalization of results observed from a dataset, from an experiment, from an optimization to a wider class of problems induced by the invariance property. In numerical optimization, affine-invariance in the search space is the ultimate invariance that implies all other classical invariances.

In this presentation, after motivating invariance, I will discuss the link between affine-invariance, convergence and learning of second order information using two classes of algorithms: on the one-hand quasi-Newton algorithms and more recent derivative-free stochastic algorithms like variants of the CMA-ES. For this latter example, the algorithm learns the inverse Hessian on strictly convex-quadratic functions in a comparison-based setting, i.e. without use of derivative nor of function value.

1g - FB327 - 515C - **Stochastic DFO 2** Chair: WARREN HARE ←

Stochastic derivative-free optimization algorithms using random subspace strategies – KWASSI JOSEPH DZAHINI, ARGONNE NATIONAL LABORATORY – Stefan M. Wild

We propose trust-region and direct-search frameworks for large-scale stochastic derivative-free optimization by introducing the STARS and StoDARS algorithms. While STARS achieves scalability by minimizing random models that approx-

imate the objective in low-dimensional affine subspaces thus significantly reducing per-iteration costs in terms of function evaluations, these goals are achieved by StoDARS through the exploration of the decision space by means of poll directions generated in random subspaces. These subspaces and their dimension are chosen via Johnson-Lindenstrauss transforms such as those obtained from Haar-distributed orthogonal random matrices. The quality of the subspaces and sets of poll directions, as well as the accuracies of estimates and models used by the algorithms are required to hold with sufficiently high, but fixed, probabilities. Convergence and complexity results are obtained for both methods using martingale theory. In particular, by leveraging the ability of StoDARS to generate a dense set of poll directions, its almost sure convergence to Clarke stationary points is established. Moreover, the analysis of second-order behavior of the well-known mesh adaptive direct-search algorithms using a second-order-like extension of the Rademacher's theorem-based definition of the Clarke sub-differential (so-called generalized Hessian) is extended to the StoDARS framework, making it the first in a stochastic direct-search setting to the best of our knowledge.

Non-convergence analysis of probabilistic direct search – CUNXIN HUANG, THE HONG KONG POLYTECHNIC UNIVERSITY – Zaikun Zhang

Direct search is a popular method in derivative-free optimization. Probabilistic direct search has attracted increasing attention in recent years due to both its practical success and theoretical appeal. It is proved to converge under certain conditions at the same global rate as its deterministic counterpart, but the cost per iteration is much lower, leading to significant advantages in practice. However, a fundamental question has been lacking a systematic theoretical investigation: when will probabilistic direct search fail to converge? We answer this question by establishing the non-convergence theory of probabilistic direct search. We prove that probabilistic direct search fails to converge if the searching set is probabilistic ascent. Our theory not only deepens our understanding of the behavior of the algorithm, but also clarifies the limit of reducing the cost per iteration by randomization, and hence provides guidance for practical implementations of probabilistic direct search.

Full-low evaluation methods for bound and linearly constrained derivative-free optimization – CLÉMENT ROYER, UNIVERSITÉ PARIS DAUPHINE PSL – Oumaima Sohab, Luis Nunes Vicente

Derivative-free optimization consists in finding the best value of an objective function without relying on derivatives. To tackle such problems, one may build approximate derivatives, using for instance finite-difference estimates. One may also design algorithmic strategies that perform space exploration and seek improvement over the current point. The first type of strategy often provides good performance on smooth problems but at the expense of more function evaluations. The second type is cheaper and typically handles non-smoothness or noise in the objective better. Recently, full-low evaluation methods have been proposed as a hybrid class of DFO algorithms that combine both strategies, respectively denoted as Full-Eval and Low-Eval. In the unconstrained case, these methods showed promising numerical performance.

In this talk, we present an extension of the full-low evaluation framework to bound and linearly constrained derivative-

free optimization problems. We derive convergence results for an instance of this framework, that combines finite-difference quasi-Newton steps with probabilistic direct-search steps. The former are projected onto the feasible set, while the latter are defined within tangent cones identified by nearby active constraints. We also showcase the practical advantages of our instance over algorithms that rely solely on Full-eval or Low-eval iterations on standard linearly constrained problems, that we adapt to introduce noisy evaluations as well as non-smoothness.

Expected decrease for derivative-free algorithms using random subspaces – WARREN HARE, UNIVERSITY OF BRITISH COLUMBIA – Lin-don Roberts, Clement Royer

Derivative-free algorithms seek the minimum of a given function based only on function values queried at appropriate points. Their performance is known to worsen as the problem dimension increases. Recent advances in developing high dimension derivative-free techniques have tackled this issue by working in low-dimensional subspaces that are drawn at random in an iterative fashion. In this talk, we present analysis for derivative-free algorithms that employing random subspaces to obtain understanding of the expected decrease achieved per function evaluation. We develop an analysis for derivative-free algorithms (both direct search and model based approaches) employing random subspaces. Our results leverage linear local approximations of smooth functions to obtain understanding of the expected decrease achieved per function evaluation. Although the quantities of interest involve multidimensional integrals with no closed-form expression, a relative comparison for different subspace dimensions suggest that low dimension is preferable. Numerical computation of the quantities of interest confirm the benefit of operating in low-dimensional subspaces.

1h - MB88 - 515C - Nonsmooth PDE Constrained Optimization 1 Chair: HARBIR ANTIL



PDE Constrained Optimization and Digital Twins – HARBIR ANTIL, GEORGE MASON UNIVERSITY –

With recent advancements in computing resources and interdisciplinary collaborations, a new research field called Digital Twins (DTs) is starting to emerge. Data from sensors located on a physical system is fed into its DT, the DT in-turn help make decisions about the physical system. This cycle then continues for the life-time of the physical system. A typical example is a bridge or a human heart.

In many cases, these problems can be cast as PDE-constrained optimization (PDECO) problems. This talk begins by discussing the role of PDECO in DTs. All the aforementioned decisions must be made while accounting for the underlying uncertainties. In this vein, a risk-averse optimization based framework is considered. To overcome the high-dimensional costs (due to uncertainty), a tensor train decomposition framework is introduced.

A fundamental challenge for these problems is the ambiguity in the underlying distribution. The last part of the talk will introduce a “Rockafellian” framework for PDECO that is robust to inaccuracies in the precise form of the problem un-

certainty. The framework is based on problem relaxation and involves optimizing a bivariate objective functional that features both a standard control variable and an additional perturbation variable that handles the distributional ambiguity.

Semi-smoothness in Infinite Dimensional Non-Smooth Optimization – ANTON SCHIELA, UNIVERSITY OF BAYREUTH –

We consider local convergence of Proximal-Newton Methods in Hilbert spaces. Fast local convergence of the unregularized algorithm can be shown straightforwardly under a semi-smoothness assumption on the differentiable part of the objective. However, transition to fast local convergence of a globalized algorithm is surprisingly subtle and requires additional insight into the problem. In this talk we will elaborate on this topic, where non-smooth optimization and semi-smooth Newton meet.

A numerical solution approach for non-smooth optimal control problems based on the Pontryagin maximum principle – DANIEL WACHSMUTH, UNIVERSITY OF WUERZBURG –

Optimal control problems subject to partial differential equations appear in many applications. In recent years, interest has shifted to non-smooth optimal control problems. In this talk, we present an algorithm to numerically solve non-smooth and non-convex optimal control problems. The cost functional is assumed to be of integral type with possibly non-smooth and non-convex integrands. As state equation we choose semilinear elliptic or parabolic partial differential equations. Then the method is applicable to a wide range of problems, including problems with integer-valued controls. For problems with this structure the celebrated Pontryagin maximum principle is a necessary optimality condition, which has no analogue in finite-dimensional optimization. The method we present is a gradient-like scheme that is based on the Pontryagin maximum principle, and includes a Armijo-type linesearch procedure. The descent direction is motivated by topological derivatives. We discuss convergence properties of the method: weak limit points are stationary in the following sense. If the state part of the functional is convex then the iterates converge to a global solution of the problem. Otherwise the maximum principle is fulfilled for weak limit points up to an epsilon. We emphasize that the method is applicable to problems, where existence of optimal controls cannot be proven. In addition, we present results of numerical experiments.

Interweaved first-order methods for PDE-constrained and bilevel optimisation – TUOMO VALKONEN, ESCUELA POLITÉCNICA NACIONAL – Ensio Suonperä, Bjørn Jensen

PDE-constrained optimisation problems, as well as bilevel optimisation problems—which cover nonsmooth PDE constraints, such as Bingham flow—have traditionally been solved with one of two approaches: (a) Newton-type methods applied to sufficiently smooth optimality conditions for a constrained problem formulation, or (b) through treating the inner problem or partial differential equation through its solution mapping. Unless derivative-free methods are applied, the latter generally involves calculating both the solution mapping and its derivative to a high precision on each step of the outer optimisation method.

Recently in bilevel optimisation research, especially as applied to machine learning, so-called single-loop approaches have been introduced, to reduce the computational cost of frequently solving the inner problem. On each step of the outer method, such overall methods only take a single step of a conventional optimisation method towards the solution inner problem, bridging the gap between the two aforementioned approaches. The same principle can be applied to PDE-constrained optimisation. As we have recently shown, significant performance improvements can be obtained by interweaving the steps of conventional iterative solvers (Jacobi, Gauss–Seidel, conjugate gradients) for both the PDE and its adjoint, with the steps of the optimisation method. Moreover, in this talk, we demonstrate how the adjoint equation in bilevel problems can also benefit from such interweaving with conventional linear system solvers.

1h - TA90 - 512B - Nonsmooth PDE Constrained Optimization 2 Chair: ANNIKA SCHIEMANN

A hybrid physics-informed neural network based multiscale solver as a PDE constrained optimization problem – MICHAEL HINTERMÜLLER, WEIERSTRASS INSTITUTE BERLIN –

We study physics-informed neural networks (PINNs) constrained by partial differential equations (PDEs) and their application in approximating multi scale PDEs. The latter may arise in multi scale material models where, e.g. in a two-scale setting, the coarse scale model is accessible via the FEM whereas the fine scale problem is approached by a NN technique. In fact, from a continuous perspective, our formulation corresponds to a non-standard PDE-constrained optimization problem with a PINN-type objective. From a discrete standpoint, the formulation represents a hybrid numerical solver that utilizes both neural networks and finite elements. We propose a function space framework for the problem and develop an algorithm for its numerical solution, combining an adjoint-based technique from optimal control with automatic differentiation. As a proof of concept, our multi scale solver is applied to a heat transfer problem with oscillating coefficients, where the neural network approximates a fine-scale problem, and a coarse-scale problem constrains the learning process. We show that incorporating coarse-scale information into the neural network training process through our modelling framework acts as a preconditioner for the low-frequency component of the fine-scale PDE, resulting in improved convergence properties and accuracy of the PINN method. The relevance of the hybrid solver to numerical homogenization is discussed.

Decomposition in Integer Optimal Control with Total Variation Regularization – PAUL MANNS, TU DORTMUND UNIVERSITY – Robert J Baraldi

We are interested in solving mathematical programs that feature distributed and discrete-valued decision variables that allow to control a dynamical system that is modeled by means of a PDE constraint. The distributed decision variables are measurable functions and regularized with a total variation penalty, that is the value of the regularizer is the sum of the interface areas of the level sets of decision variables weighted by the jump

heights over the respective interfaces. In particular, the level sets have finite perimeters. In the multi-dimensional case, current discrete optimization-based techniques enable us to solve medium-sized problem instances to optimality in an acceptable amount of time in the context of descent algorithms. Realistic problem sizes may currently be deemed intractable, however. In order to overcome this issue, we propose and analyze a domain decomposition technique that allows to replace the solution of large-scale problem instances to many medium-sized problems that are coupled through overlaps of the subdomains on which the smaller problems are defined. To this end, we employ covering arguments inside of an algorithmic framework to localize the optimality and sufficient decrease conditions for a superordinate trust-region strategy. We employ an algorithm leans on greedy approaches in coordinate descent algorithms in order to obtain convergence to stationary points.

Integer linear programming for two-dimensional trust-region subproblems arising in integer optimal control – MARVIN SEVERITT, TU DORTMUND UNIVERSITY – Paul Manns

We consider subproblems arising from a trust-region algorithm for mixed integer optimal control problems with total variation regularization and integer-valued control functions. After discretization we obtain a class of integer linear programs. We discuss the NP-hardness for the one-dimensional and the two-dimensional case. In the one-dimensional case, the subproblems can be transformed into shortest path problems from a source to a sink on a layered graph. This graph-based approach allows us to solve the one-dimensional problem in pseudo-polynomial time. We shortly discuss why this approach does not work for the two-dimensional case. Instead, we observe that the underlying polyhedrons exhibit structural restrictions in the vertices regarding which variables can attain fractional values at the same time. We use a graph-based view of the problem and this special property of the underlying polyhedron to obtain valid cutting planes to tighten the linear programming relaxation. Furthermore, we briefly discuss how good primal solutions and a branching rule can be derived from the one-dimensional case.

Integer Optimal Control Problems with TV regularization on Multi-dimensional Domains – ANNIKA SCHIEMANN, TU DORTMUND UNIVERSITY – Paul Manns

We consider integer optimal control problems with total variation regularization on multi-dimensional domains and derive first-order optimality conditions by means of local variations of the level sets of the feasible control functions. In order to solve this class of problems, we present a function space trust-region algorithm as well as an outer approximation algorithm to solve such problems after finite-element discretization. The control functions are discretized with integer-valued functions that are constant on the cells of a cubic mesh. We introduce a discretized total variation whose discretization is coupled to but varying from the discretization of the control functions. The coupling of the discretizations allows to recover the total variation of discontinuous integer-valued control functions with the integer-valued discretized control functions in the sense of Gamma-convergence. We add a constraint to the discretized problems that vanishes in the limit and that enforces compactness which guarantees the existence of minimizers. We prove that those minimizers converge to minimizers of the original

problem when the mesh sizes of the coupled discretizations are driven to zero. We will conclude by presenting some numerical examples that confirm our theoretical results.

1h - TB91 - 512B - **Nonsmooth PDE Constrained Optimization 3** Chair: HENRIK WYSCHKA 

An Inexact Trust-Region Algorithm for Nonsmooth Risk-Averse Optimization – DREW KOURI, SANDIA NATIONAL LABORATORIES –

Many practical problems require the optimization of systems (e.g., differential equations) with uncertain inputs such as noisy problem data, unknown operating conditions, and unverifiable modeling assumptions. In this talk, we formulate these problems as infinite-dimensional, risk-averse stochastic programs for which we minimize a quantification of risk associated with the system outputs. For many popular risk models, the resulting risk-averse objective function is not differentiable, significantly complicating the numerical solution of the optimization problem. Unfortunately, traditional methods for nonsmooth optimization converge slowly (e.g., sublinearly) and consequently are often intractable for problems in which the objective function and any derivative information is expensive to evaluate. To address this challenge, we introduce a novel trust-region algorithm for solving large-scale nonsmooth risk-averse optimization problems. This algorithm is motivated by the primal-dual risk minimization algorithm and employs smooth approximate risk measures at each iteration. In addition, this algorithm permits and rigorously controls inexact objective function value and derivative (when available) computations, enabling the use of inexpensive approximations (e.g., adaptive discretizations). We discuss convergence of the algorithm under mild assumptions and demonstrate its efficiency on various examples from PDE-constrained optimization.

Control in the Coefficients of an Obstacle Problem – NICOLAI SIMON, UNI HAMBURG – Winfried Wollner

In this talk we discuss the introduction of a control variable into the coefficients of an optimal control problem governed by an obstacle problem. As is known, the associated solution operator of the obstacle problem is not Gateaux differentiable, therefore we utilize a regularization approach to compute first order limiting optimality conditions for this problem.

The novelty and focus of this talk is the introduction of a matrix valued control into the coefficients of the problem. To handle the resulting multiplicative coupling of the control and the gradient of the state, we will utilize H-convergence arguments in a bootstrapping approach and prove strong L^p convergence for this coefficient control variable.

A Discontinuous Galerkin Method for Optimal Control of the Obstacle Problem – ROHIT KHANDELWAL, GEORGE MASON UNIVERSITY –

In this talk, we discuss the quasi-optimal a priori error estimates for an optimal control problem constrained by an elliptic obstacle problem where the finite element discretization is carried out using the symmetric interior penalty discontinuous Galerkin method. The main proofs are based on the improved L^2 -error estimates for the obstacle problem, the discrete maximum principle, and a well-known quadratic growth property.

The standard (restrictive) assumptions on mesh are not assumed here.

Limiting Descent Directions in p-Harmonic Shape Optimization – HENRIK WYSCHKA, UNIVERSITÄT HAMBURG – Peter Marvin Müller, Thomas Rung, Winnifried Wollner

Shape optimization constrained to partial differential equations is a vibrant field of research with high relevance for industrial-grade applications. Recent developments suggest that using a p-harmonic approach to determine descent directions is superior to classical Hilbert space methods. This applies in particular to the representation of kinks and corners in occurring shapes. However, the approach requires the solution of a vector-valued p-Laplace problem with a boundary force for each descent direction. We present an algorithm to solve these problems for finite p efficiently and discuss extensions to the limit setting. A key challenge in this regard is that the Lipschitz deformations obtained as solutions in limit setting are in general non-unique. Thus, we focus on solutions which are in a sense limits to solutions for finite p and aim to preserve mesh quality throughout the optimization.

1h - TC320 - 512B - Analysis of Algorithms for PDE Constrained Optimization 1 Chair: STEFAN ULBRICH

Snapshot generation for POD models of semilinear parabolic optimal control problems based on a simplified Newton method – STEFAN ULBRICH, TECHNICAL UNIVERSITY OF DARMSTADT – Paul Manns

In PDE-constrained optimization, proper orthogonal decomposition (POD) provides a surrogate model of a (potentially expensive) PDE discretization on which optimization iterations are executed. Because POD models usually provide good approximation quality only locally, they have to be updated during optimization. Updating the POD model is usually expensive, however, and therefore often impossible in a model-predictive control (MPC) context. Thus, reduced models of mediocre quality might be accepted. We take the view of a simplified Newton method for solving semilinear evolution equations to derive an algorithm that can serve as an offline phase to produce a POD model. Approaches that build the POD model with impulse response snapshots can be regarded as the first Newton step in this context. In particular, POD models that are based on impulse response snapshots are extended by adding a second simplified Newton step. This procedure improves the approximation quality of the POD model significantly by introducing a moderate amount of extra computational costs during optimization or the MPC loop. We illustrate our findings with an example satisfying our assumptions.

On the well-posedness of the Mortensen observer for a defocusing cubic wave equation – JESPER SCHRÖDER, TECHNISCHE UNIVERSITÄT BERLIN –

In this presentation the analytical background of nonlinear observers based on minimal energy estimation is discussed. Originally, such strategies were proposed for the reconstruction of the state of finite dimensional dynamical systems by means of a

measured output where both the dynamics and the output are subject to white noise. Our work aims at lifting this concept to a class of partial differential equations featuring deterministic perturbations using the example of a wave equation with a cubic defocusing term in three space dimensions. In particular, we investigate local regularity of the corresponding value function and consider operator Riccati equations to characterize its second spatial derivative.

Adaptive Randomized Sketching for Nonsmooth Dynamic Optimization – ROBERT J BARALDI, SANDIA NATIONAL LABORATORIES – Drew Kouri, Harbir Antil

Dynamic optimization problems arise in many applications, such as optimal flow control, full waveform inversion, and medical imaging. Despite their ubiquity, such problems are plagued by significant computational challenges. For example, memory is often a limiting factor when determining if a problem is tractable, since the evaluation of derivatives requires the entire state trajectory. Many applications additionally employ nonsmooth regularizers such as the L1-norm or the total variation, as well as auxiliary constraints on the optimization variables. We introduce a novel trust-region algorithm for minimizing the sum of a smooth, nonconvex function and a nonsmooth, convex function that addresses these two challenges. Our algorithm employs randomized sketching to store a compressed version of the state trajectory for use in derivative computations. By allowing the trust-region algorithm to adaptively learn the rank of the state sketch, we arrive at a provably convergent method with near optimal memory requirements. We demonstrate the efficacy of our method on a few control problems in dynamic PDE-constrained optimization.

1h - WA914 - 512B - Optimal Control, PDE Constrained Optimization, and Multi-level Methods 1 Chair: SEPEHR MOALEMI

Image Registration Using Optimal Control of a Linear Hyperbolic Transport Equation – JOHANNES HAUBNER, UNIVERSITY OF GRAZ – Bastian Zapf, Lukas Baumgärtner, Stephan Schmidt

Image registration is crucial in many imaging applications such as medical imaging or computer vision. The goal of finding a suitable transformation between two images poses similar restrictions and regularity requirements on the set of admissible transformations as shape optimization problems. In the scope of this talk, we build on an approach that models image registration as an optimization problem that is constrained by a linear hyperbolic transport equation. We use a higher-order discontinuous Galerkin finite element method for discretization and motivate the numerical upwind scheme and its limitations from the continuous weak space-time formulation of the transport equation. Moreover, we build on recent theoretical results to model the optimization problem. To discuss the potential of the proposed algorithm, we apply it to patient specific brain mesh generation from magnetic resonance images (MRI). This can be a time consuming task and require manual corrections, e.g., for meshing the ventricular system or defining subdomains. The idea is to use the registration of an input MRI to a respective target in order to obtain a new mesh from a high-quality template mesh.

Robust Optimal Control Through Scenario-

Based Solutions to Semi-Infinite Programs With Existence Constraints – JAD WEHBEH, IMPERIAL COLLEGE LONDON – Eric C. Kerrigan

Robust optimal control problems, when expressed in discrete time, generally take a min-max form over continuous variable spaces. In this talk, we focus on the problem of choosing decision variables that optimize the performance of a system in the worst-case realization of uncertainties drawn from bounded continuous spaces. We also consider systems with non-unique state trajectories that arise from the introduction of modelling variables through smoothing approaches, as well as existence constraints, which are useful in describing various logical conditions. Under certain assumptions on the continuity of the functions involved, we demonstrate how this problem can be cast as a semi-infinite program (SIP) and solved using a local reduction technique, which replaces the infinite-dimensional constraints by an optimally chosen finite scenario set. The solution procedure is therefore split into a number of lower level problems that find the uncertainty scenarios that maximize the constraint violations given the current best guess for the decision variables, and a high level problem that minimizes the original objective under the obtained scenario set. The existence constraints, meanwhile, can be reformulated as separate SIPs, and the worst-case violations for these constraints can be obtained by solving these nested SIPs using the same local reduction approach. Our approach solves a wide range of problems and exhibits better scaling than scenario trees or traditional global optimization techniques. We illustrate this through a number of examples drawn from discrete-time robust optimal control.

Towards robust optimization of chromatographic separation processes with flow reversal – DOMINIK H. CEBULLA, INSTITUTE FOR MATHEMATICAL OPTIMIZATION, TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG – Christian Kirches, Andreas Potschka

Column liquid chromatography plays an important role in the downstream processing of biopharmaceuticals, where the goal is to capture and purify a target protein from a mixture. Our goal is to employ a model-based approach for process optimization in order to improve the quality of the product, while also achieving further economical and ecological benefits.

Rate models in combination with suitable reaction schemes that model the specific adsorption process are often employed to describe chromatographic processes. Regarding the optimization of such processes, due to the presence of discrete control decisions, we are eventually facing mixed-integer optimal control problems governed by advection-diffusion-reaction-type partial differential equations with high nonlinearities. Furthermore, since at least one flow reversal is typically performed in practical applications to obtain sharper elution profiles, we additionally have to deal with switching dynamics. Lastly, it is important to determine robust solutions in order to safeguard against, e.g., uncertain model parameters, such as reaction rates and feed composition.

In this talk we present developments towards robustly optimal switching control applied to chromatographic separation processes and discuss the obtained results.

Input-Output Stability of First-Order Optimization Algorithms: A Passivity Approach

– SEPEHR MOALEMI, MCGILL UNIVERSITY – James Richard Forbes

In this presentation, the stability of popular first-order optimization algorithms is examined through the lens of passivity. The passivity theorem ensures input-output stability of a passive plant connected in a negative feedback loop with a very strictly passive (VSP) system. Combining existing work on control interpretation of first-order optimization algorithms with loop transformation techniques, it is shown that gradient descent (GD) can be rendered passive, while the more recent triple momentum (TM) method is input strictly passive (ISP). It is shown that the sector boundedness of the gradient of an L-smooth, m-strongly convex function renders it being VSP. Therefore, the passivity theorem can ensure input-output stability of GD when the gradient is L-smooth, m-strongly convex.

1h - WB321 - 512B - Challenging PDE Constrained Applications Chair: ANTON BONDAREV



Low-rank tensor approximation for risk averse optimization – AKWUM ONWUNTA, LEHIGH UNIVERSITY –

In this talk, we discuss the solution of high-dimensional risk-averse optimization problems governed by differential equations (ODEs and/or PDEs) under uncertainty. As an example, we focus on the so-called Conditional Value-at-Risk (CVaR), but the approach is equally applicable to other coherent risk measures. Both the full and reduced space formulations are considered. Our approach is based on tensor train approximations of random fields discretized using stochastic collocation. To avoid the non-smoothness of the objective function underpinning the CVaR, we propose an adaptive strategy to select the width parameter of the smoothed CVaR to balance the smoothing and tensor approximation errors. Moreover, an unbiased Monte Carlo CVaR estimate can be computed by using the smoothed CVaR as a control variate. To accelerate the computations, we introduce an efficient preconditioner for the Karush-Kuhn-Tucker (KKT) system in the full space formulation. We also present numerical experiments to demonstrate that the proposed method enables accurate CVaR optimization constrained by PDEs with random coefficients.

Optimal Control of the Generalized Riemann Problem for Hyperbolic Systems of Conservation Laws – JANNIK BREITKOPF, TECHNICAL UNIVERSITY DARMSTADT – Stefan Ulbrich

In this talk, we analyze optimal control problems for quasilinear strictly hyperbolic systems of conservation laws where the control is the initial state of the system. The problem is of interest, for example, in the context of fluid mechanics or traffic flow modelling. Similar problems for scalar conservation laws have already been studied. However, the case of hyperbolic systems is more involved due to the coupling of the characteristic fields.

We begin our analysis by considering the Generalized Riemann Problem, which has a piecewise smooth initial state with exactly one discontinuity. This is a natural choice since it is well known that solutions to hyperbolic conservation laws generally develop discontinuities even for smooth data. For piecewise C^1 initial data we obtain the existence, uniqueness and

stability of an entropy solution by a careful fixed point argument built on the associated Riemann Problem with piecewise constant initial states. The construction yields insights into the structure and regularity of the solution and provides a foundation to derive differentiability results of the control-to-state mapping.

The entropy solution is piecewise C^1 . Its smooth parts are separated by C^2 curves which are either shock curves or boundaries of rarefaction waves. In a subsequent step, we show that these curves depend differentiably on the initial state. This allows the transformation to a fixed space-time domain as a reference space. In this reference space, we can show that the transformed solution depends differentiably on the initial state in the topology of continuous functions. For this, a detailed knowledge of the structure of the solution and the behaviour of the shock curves is crucial. As an immediate consequence, the differentiability of tracking type functionals for the optimal control problem follows.

A Variational Calculus for Optimal Control of Networks of Scalar Conservation or Balance Laws – MARCEL MAXIMILIAN STEINHARDT, TECHNICAL UNIVERSITY DARMSTADT – Stefan Ulbrich

Networks of scalar conservation or balance laws provide models for vehicular traffic flow, supply chains or transmission of data. Such networks usually consist of initial boundary value problems (IBVPs) of scalar conservation or balance laws on every edge coupled by node conditions. For their optimal control a variational calculus is desirable that implies differentiability of objective functionals w.r.t. controls. In the last decade research on IBVPs successfully introduced a variational calculus which implies differentiability of objective functionals of tracking type and also yields an adjoint based gradient representation for the functional. This talk presents recent progress in an extension of these results to networks of scalar conservation or balance laws. Regarding node conditions we introduce a framework for their representation compatible with the known approach on single edges which allows us to extend the desired results from IBVPs. Therefore, in the optimal control on networks of scalar conservation or balance laws we obtain continuous Fréchet differentiability for functionals of tracking-type w.r.t. controls and an adjoint based gradient representation on the network.

Parameter Identification in Boundary Value Problems for Fiber Spinning – LUKAS KANNENGISSER, TRIER UNIVERSITY – Nicole Marheineke, Raimund Wegener

Industrial manufacturing processes are often described by systems of differential equations with respective material laws. Since these material laws depend on material-specific parameters, which are frequently unknown or can only be obtained through a range of expensive experiments, there is a need for numerical parameter estimation strategies. In this context, we consider a fiber spinning process modeled by a boundary value problem (BVP) of ordinary differential equations where some material parameters need to be estimated on the basis of few experimental data. We have developed a numerical method involving collocation and continuation methods to solve BVPs arising in a generalized Newtonian setting. Furthermore we are able to calculate gradients of the BVP solutions with respect to the material parameters and apply nonlinear optimization

techniques to optimize the material parameters of interest with respect to given measurement data. In this talk we aim to replace the computationally demanding BVPs through surrogate models, leading to fast online optimization. Furthermore we investigate the training process of these surrogates and transfer learning approaches for different fiber spinning configurations.

1h - ThA361 - 512B - Analysis of Algorithms for PDE Constrained Optimization 2 Chair:
ARSEN HNATIUK

Sensor path planning for optimal experimental design – THOMAS LYNN, ARGONNE NATIONAL LABORATORY – Nicole Aretz, Sven Leyffer

In scenarios where static sensors are impractical, moving sensors are used to gather data to solve inverse problems. We optimize the path that such a mobile sensor takes for various optimal experimental design objectives applied to Bayesian inverse problems. We demonstrate our approach using a reaction-advection-diffusion equation that models the spread of a pollutant using a finite element approach. We impose constraints on the observation path such as a system of differential algebraic equations governing the movement of the sensor and obstacle avoidance constraints.

Low-rank Gradient Flow – a First Order Algorithm for Non-convex Optimization – ANDREI DRAGANESCU, UNIVERSITY OF MARYLAND, BALTIMORE COUNTY –

We introduce a novel class of first-order methods for unconstrained optimization, called low-rank gradient flows (LRGFs). The idea behind these methods is to construct at every optimization step a low-rank quadratic surrogate for the cost function, followed by an analytic solve for the gradient flow on the surrogate model; the optimization step concludes with a line search on the curve representing the gradient flow. It is shown that the above steps are condensed in a very simple formula for the gradient flow, at a cost per step that is comparable to that of a nonlinear conjugate gradient algorithm. The fact that the line search is conducted along a curve distinguishes LRGF from other first order optimization methods, where the line search is conducted along a search direction, that is, a straight line. This may also help LRGF better navigate the geometry of the cost function, allowing the method to avoid local minima more often than other first-order methods, as shown by numerical experiments. For higher dimensional problems the convergence can be accelerated using a multilevel strategy based on reduced order models.

Optimal Experimental Design for Universal Differential Equations – CHRISTOPH PLATE, OTTO VON GUERICKE UNIVERSITY MAGDEBURG – Carl Julius Martensen, Sebastian Sager

We are interested in using universal differential equations to model complex dynamical processes, e.g. in chemical engineering or medicine. Universal differential equations, or hybrid models, combine the advantages of first-principles models and data-driven approaches. Often the dynamical process is only partially known, i.e. it may consist of subprocesses that are not yet fully understood or difficult to describe from first principles. For these unknown or uncertain subprocesses, universal approximators, such as neural networks, can be embedded in

the model and trained on data. Since this training procedure requires experimental data, which are usually costly to collect, it is necessary to carefully design the experimental setup in advance. This amounts to deciding when to measure and how to stimulate the dynamic process in order to collect data with a high information content. This is addressed by the concept of optimal experimental design (OED), leading to a challenging, specifically structured optimal control problem. We formulate the OED problem for universal differential equations, where the number of parameters to be estimated can be very large, and discuss the properties of the resulting optimization problem. We also propose different methods to reduce the complexity resulting from the large number of parameters.

Lazified Generalized Conditional Gradient – ARSEN HNATIUK, HUMBOLDT UNIVERSITÄT ZU BERLIN –

The primal-dual active point strategy (PDAP) is known to achieve asymptotically linear convergence rates on the task of minimizing the sum of a smooth, convex loss function and a non-smooth convex regularizer over a Banach space, possibly with PDE constraints. This approach outperforms other non-accelerated conditional gradient methods at the cost of requiring solutions to two finite-dimensional optimization problems in each iteration. One of these problems is smooth and possibly non-convex, while the other is convex and non-smoothly regularized. In the present work, we investigate the properties of a lazified approach, where the subproblems are not required to be solved exactly. We demonstrate that in this setting there exist practical error tolerances on the subproblems such that asymptotical linear convergence is preserved.

1h - ThB318 - 512B - **Shape Optimization Chair:**
THEODOR KOMANN 

Level Set-Based Topology Optimization Of 3D Magnetic Circuits With Mechanical Constraints – ZAKARIA HOUTA, ENSEEIHT LAPLACE, UNIVERSITY OF TOULOUSE – Nicolas Lebbe, Thomas Huguet, Frédéric Messine

A topology optimization problem is a problem where the variables allow to define optimally the structure of certain parts of an object. This type of optimization is essential in many applications, such as the design of the structure of aircraft wings or electric motors. In our case, we are interested in optimizing the design of a 3D magnetic circuit which resembles a Hall effect thruster. This circuit contains a variable area, where, at any point, the presence of matter or vacuum is described by a level set function, the value of which is decided by the optimization process. The objective function to be optimized is defined from the magnetic field calculated in a target area. This magnetic field must correspond as well as possible to a fixed objective field in order to ensure proper thruster operation. This optimization problem is a PDE constrained optimization problem, where the PDEs in this case are Maxwell equations (of the magnetostatic).

One of the difficulties encountered in manufacturing a so-optimized magnetostatic structures, is that they are not necessarily mechanically stable. In order to take this mechanical constraint into account, we have developed a level set-based topology optimization algorithm which relies on gradient information and on numerical simulations of both the mechanical

deformation and the magnetostatic behavior of the structure. By comparing the designs obtained with those from magnetostatic optimization alone, our approach proves some effectiveness in obtaining robust designs.

PDE constrained shape optimization in the space of piecewise-smooth shapes – KATHRIN WELKER, TU BERGAKADEMIE FREIBERG –

Shape optimization is concerned with identifying shapes, or subsets of $\mathbb{R}$, behaving in an optimal way with respect to a given physical system. It has been an active field of research for the past decades and is used for example in engineering. Many relevant problems in the area of shape optimization involve a constraint in the form of a partial differential equation (PDE). Theory and algorithms in shape optimization can be based on techniques from differential geometry, e.g., a Riemannian manifold structure can be used to define the distances of two shapes. Thus, shape spaces are of particular interest in shape optimization.

In this talk, we apply the differential-geometric structure of Riemannian shape spaces to the theory of classical PDE constrained shape optimization problems. We also present a space containing shapes in $\mathbb{R}^2$ that can be identified with a Riemannian product manifold but at the same time admits piecewise smooth curves as elements. We present algorithms to solve PDE constrained (multi-)shape optimization problems and give numerical results of these algorithms.

Mesh Denoising and Inpainting using the Total Variation of the Normal and a Shape Newton Approach – ROLAND HERZOG, HEIDELBERG UNIVERSITY –

We present a novel approach to denoising and inpainting problems for surface meshes. The purpose of these problems is to remove noise or fill in missing parts while preserving important features such as sharp edges. A discrete variant of the total variation of the unit normal vector field serves as a regularizing functional to achieve these goals. In order to solve the resulting problem, we use a version of the split Bregman (ADMM) iteration adapted to the problem. A new formulation of the total variation regularizer, as well as the use of an inexact Newton method for the shape optimization step, bring significant speed-ups compared to earlier methods. Numerical examples are included, demonstrating the performance of our algorithm with some complex 3D geometries.

Robust Optimization of Electric Machines: A Gradient-Based Approach for Parameter and Shape with Isogeometric Analysis – THEODOR KOMANN, TU DARMSTADT – Stefan Ulbrich

In the field of electric machine design, optimizing geometry and parameters is crucial and the consideration of uncertainties like manufacturing tolerances is important. This work presents a robust optimization framework, which is applied to design a 3-phase, 6-pole Permanent Magnet Synchronous Motor (PMSM) and addresses these uncertainties. Our approach uses Isogeometric Analysis (IGA) for the two-dimensional construction of the geometry and the discretization of the governing state equation. To optimize the performance of a PMSM, we combine shape and parameter optimization, which allows for a larger design space. To robustify the optimal design against uncertainty we apply a robust optimization technique that uses a min-max formulation. To solve this bilevel problem, we work

with the value functions of the lower level maximization problems and apply Danskin's Theorem for the computation of generalized derivatives. Additionally, the adjoint method is employed to efficiently solve the lower level problems with gradient based optimization. In the end we present numerical results, that show the efficiency of our approach.

1h - ThC319 - 512B - Multi-Level and Parallel Methods for PDE Constrained Optimization

Chair: WING HONG FELIX KWOK



Multigrid-in-time methods for nonlinear optimization of dynamical systems – RADOSLAV VUCHKOV, SANDIA NATIONAL LABORATORIES –

Fast methods for nonlinear optimization with ordinary differential equations (ODEs) and partial differential equations (PDEs) are critical to the solution of optimal control, optimal design and inverse problems in fluid, flight, structural and electrodynamics. To compute a step, based on the derivatives of the cost function and the governing differential equations, conventional methods rely on the forward and backward (adjoint) time evolution of the dynamical system. The serial nature of time stepping introduces an inherent computational bottleneck, resulting in long solution times even for relatively simple problems governed by scalar ODEs. In contrast, matrix-free composite-step sequential quadratic programming (SQP) methods [1,2] offer an interesting alternative. Specifically, they solve a sequence of convex quadratic optimization problems that lead to Karush Kuhn Tucker (KKT) saddle-point systems in which the linearized forward and adjoint time evolutions are fully coupled. This time-domain coupling, combined with the convexity of the quadratic subproblems, enables a multigrid-in-time approach. To develop our multigrid-in-time method, we first express the KKT systems in their block-tridiagonal form, where the diagonal blocks maintain the saddle-point structure on each time interval, and the off-diagonal block expose the time coupling. From the tridiagonal form, we construct block-Jacobi and block-Gauss-Seidel fixed-point schemes. The block-Jacobi scheme is used as the pre- and post-smoother for the proposed multigrid method. This smoother is trivially parallelizable in the time domain. A Symmetric block-Gauss-Seidel (SGS) scheme is used as the preconditioner for a coarse-grid solver based on an inner GMRES iteration. The SGS scheme, which is also investigated in [3], serializes in the time domain, however its cost is acceptable as it is applied to only a handful of time steps. Our prolongation and restriction operators are based on simple one-dimensional grid projection and injection operations. We use the multigrid method as the preconditioner for an outer flexible GMRES (FGMRES) iteration. We demonstrate excellent algorithmic scalability of the proposed multigrid-in-time approach, and the overall SQP method, on nonlinear optimization problems with ODE and PDE constraints. More concretely, the scalability extends to (i) the inner coarse-grid SGS-preconditioned GMRES method, (ii) the outer multigrid-preconditioned FGMRES method, and (iii) the SQP method. Our results also point to a very mild influence of dynamical system coefficients on the number of FGMRES iterations.

A full approximation scheme multilevel method for solving nonlinear variational inequalities – ED BUELER, UNIVERSITY OF ALASKA

FAIRBANKS – Patrick Farrell

The full approximation scheme constraint decomposition (FASCD) multilevel method for solving variational inequalities (VIs) is a joint extension of both the full approximation scheme (FAS) multigrid technique for nonlinear partial differential equations, due to A. Brandt, and the constraint decomposition (CD) method introduced by X.-C. Tai for VIs arising in optimization. It extends the CD idea by exploiting the telescoping nature of certain subset decompositions arising from multilevel mesh hierarchies. When a reduced-space (active set) Newton method is applied as a smoother, with work proportional to the number of unknowns on a given mesh level, FASCD V-cycles exhibit nearly mesh-independent convergence rates. The full multigrid cycle version is observed to be an optimal solver. Example problems include differential operators which are symmetric linear, nonsymmetric linear, and nonlinear, in unilateral and bilateral VI problems.

Pipeline waveform relaxation for optimal control – WING HONG FELIX KWOK, UNIVERSITÉ LAVAL –

(Note: this abstract may be updated in the future). When solving optimal control problems where the system is governed by wave propagation phenomena, each gradient calculation requires integrating the forward and adjoint PDEs. To add spatial and time parallelism to these forward and adjoint solves, we propose a waveform relaxation method with adaptive pipelining, as described in Kwok and Ong [SISC 2019]. We will study the convergence behaviour of this method applied to the wave equation and the theoretical speedup achievable when solving the overall optimal control problem.

2a - MB157 - 511F - Vehicle Routing Problems under Uncertainty Chair: PEDRO MUNARI



Approximating the recourse cost function of the vehicle routing problem with stochastic demands. – MATHEUS JUN OTA, UNIVERSITY OF WATERLOO – Ricardo Fukasawa, Aleksandr M. Kazachkov

The vehicle routing problem with stochastic demands (VRPSD) generalizes the classic capacitated vehicle routing problem (CVRP) by considering customer demands as random variables that are revealed on vehicle arrival. Our work focuses on modeling VRPSD as a two-stage stochastic program, where a recourse policy describes the recourse actions that should be taken whenever a violation of the vehicle capacity constraint is observed. The objective function then combines the initial planned routes cost (first-stage cost) and the expected recourse cost (second-stage cost). Some state-of-the-art exact algorithms for the VRPSD are branch-and-cut algorithms that use integer-L shaped-based cuts for approximating the (expected) recourse cost function. In this presentation, I will discuss our current progress towards deriving stronger inequalities for better approximating this recourse cost function.

Specifically, we propose to approximate the recourse cost function with the value function of a network-flow type linear program and we then project this value function into the space of the original variables via Benders decomposition. However, due to the size of the network, a standard implementation of the Benders decomposition method can be too time-consuming. Instead, we generate the optimality/feasibility cuts by looking

at a cone-relaxation of the original network-flow formulation. This approach has the potential to generate stronger cuts by solving smaller linear programs. Furthermore, by leveraging well known results on the combinatorial structure of network-flow polyhedra, we prove that, under some conditions, it is possible to improve the big-M coefficients of previously proposed integer-L shaped cuts. Computational experiments are still in progress and show the potential of this approach.

Distributionally Robust Cyclic Inventory Routing – ALBERT SCHROTENBOER, EINDHOVEN UNIVERSITY OF TECHNOLOGY – Menglei Jia, Feng Chen

Stochastic inventory routing concerns the joint distribution (via vehicle routes) and replenishment of inventory in one warehouse, multiple retailer systems. In such systems, practical considerations impose that vehicle routes conform to a cyclic schedule, for example, according to a weekly repeating pattern, so that upstream and downstream planning processes can be managed efficiently. In this work, we explore a new variant of the stochastic cyclic inventory routing problem by imposing that retailer demand comes from a distributionally robust ambiguity set. This raises several fundamental questions, including what the structure is of optimal distributionally robust inventory replenishment policies for the joint replenishment of retailers in a vehicle route, and how this information can be conveyed into a branch-and-price scheme for optimal inventory routing. At the conference, we will provide insight into these fundamental questions, and provide several approaches based on a price-and-branch scheme.

Polynomial-time separation algorithms based on closed-form equations for the RVRP under travel time uncertainty – RAFAEL AJUDARTE DE CAMPOS, UNIVERSITÉ LAVAL – Pedro Munari, Leandro Coelho

We study the vehicle routing problem with time windows (VRPTW) under travel time uncertainty considering different uncertainty sets. We propose a new polynomial-time separation algorithm that can be used with any convex uncertainty set and easily incorporated into branch-and-cut, branch-and-price, or heuristic algorithms. This flexibility distinguishes the proposed algorithm from previous works in the robust VRPTW and the robust traveling salesman problem (TSPTW) literature, which mostly rely on the traditional cardinality-constrained uncertainty set or, more recently, on the knapsack uncertainty set. In addition to these uncertainty sets, we consider the ellipsoidal, factor models, and discrete uncertainty sets using instances from the VRPTW literature. We develop tailored branch-and-cut and branch-and-price-and-cut methods to validate the proposed separation algorithm for these sets. Finally, we evaluate the impact of the studied uncertainty sets on the trade-off between the robustness of the solution and its cost.

Pickup and Delivery Vehicle Routing Problems under Uncertainty via Robust Optimization – PEDRO MUNARI, FEDERAL UNIVERSITY OF SÃO CARLOS –

We address variants of the vehicle routing problem (VRP) under uncertain demand and travel times, considering pickup and delivery operations. Incorporating uncertainty into the formulations and solution methods for these problems poses addi-

tional challenges compared to other VRP variants due to the interdependence of pickup and delivery nodes, non-monotonic loads in vehicles, and the additional decision on the vehicle load when leaving the depot. To overcome these challenges, we introduce new approaches based on dynamic programming to effectively incorporate uncertainty in different variants of pickup and delivery operations via robust optimization. We consider the pickup and delivery problem with time windows (PDPTW) and the one-commodity pickup and delivery vehicle routing problem (1-PDVRP), and ways to incorporate uncertainty into compact formulations as well as tailored branch-and-cut methods. The results of computational experiments with benchmark instances show the efficacy of our approaches and their superior performance in comparison to existing approaches.

2a - TA188 - 513B - Network-flow based MIP formulations Chair: FRANÇOIS CLAUTIAUX

A Branch-and-Price Approach for a Consolidation-based Formulation of the Scheduled Service Network Design Problem – MIKE HEWITT, LOYOLA UNIVERSITY CHICAGO – François Clautiaux, Simon Belieres

The Scheduled Service Network Design Problem (SSNDP) is an optimization problem that prescribes the transportation routes through a network of shipments that are small relative to vehicle capacity. The classical integer programming formulation of the SSNDP captures consolidation opportunities by modeling the routing of shipments on a directed time-expanded network consisting of nodes and arcs. In such a network, activities at a terminal at different points in time are modeled with different nodes. It has been observed that instances of such a formulation present multiple computational difficulties, including huge size due to an a priori enumeration of time points, symmetry, and weak linear programming relaxations.

Recently, a new formulation for the SSNDP has been proposed that differs from the classical time-expanded network formulation in two fundamental ways. The first is that it is not defined on a time-expanded network at all. By avoiding the use of a time-expanded network, both of the first two computational challenges are avoided. The second is that it is based on an a priori enumeration of consolidations, wherein a consolidation is defined as a set of shipments that dispatch on the same physical move at the same time. By modeling explicitly with consolidations, vehicle capacity needs can be encoded in the data of an instance, leading to a much stronger linear programming relaxation

In this presentation we present a Branch-and-Price based approach to solving instances of this formulation that involves generating consolidations dynamically through the solution of a pricing problem. As the pricing problem takes the form of a quadratic knapsack problem, we present a speed-up technique that leverages the observation that a subset of a consolidation is itself a consolidation. With an extensive computational study we compare the effectiveness of the solution approach with a state of the art solution method for the SSNDP.

Exact Methods For Electric Routing Operations: A Subpath-Based Pricing Problem Decomposition – SEAN LO, MIT – Alexandre Jacquillet

We propose the electric vehicle routing problem (EVRP) as a general model for electrified routing operations, with applications in trucking for long-haul logistics, robotics, and machine scheduling. The EVRP admits a path-based set-partitioning formulation with infinitely many variables, due to the presence of both discrete routing decisions (as in VRP) and continuous charging decisions (new to EVRP). We develop a two-level label-setting algorithm for the pricing problem arising in column generation, which generates non-dominated subpaths and stitches them together. In so doing, we propose a new paradigm of domination for paths and subpaths, which ensures finite convergence and correctness of column generation even with infinitely many variables. We demonstrate how two-level label-setting can be applied to ng-route relaxations, via additional ng-set resources, and tighten these relaxations iteratively until one obtains elementary paths. We also demonstrate how our two-level label-setting can support cutting planes which strengthen the linear relaxation of EVRP, via additional cut resources. Numerical results show that our approach outperforms state-of-the-art algorithms and scales to realistic problem sizes.

A network-flow formulation for a scheduling problem in the car industry – FRANÇOIS CLAUTIAUX, UNIVERSITÉ DE BORDEAUX – Siham Essoudaigui, Alain Nguyen, Ruslan Sadykov, Younes Nawel

In this presentation, we study an optimization problem that arises in the context of the automobile industry. When a new model of vehicle has just been conceived, it needs to undergo a series of tests before being released on the market. This phase is costly since specific vehicles called prototypes are built exclusively for test purposes. The validation tests carried out in the various facilities are numerous and are repeated on the different mechanical parts of the prototypes (thermal, acoustic, driving, sound insulation, safety or endurance tests, for example). The goal of the company is to minimize the number of tests realized after their due dates, and the number of prototypes to manufacture.

We address the problem by means of a path-flow formulation. In our path-flow model, the sequence of jobs assigned to a machine corresponds with a path in a network, where nodes are related to states representing a partial solution for a machine, and arcs correspond to decisions of assigning a job to a machine in this state. Our contribution is to propose an efficient algorithm to produce columns in the column generation process. In our approach, configurations are thus only considered implicitly, i.e., they are a byproduct of the set of jobs assigned to each machine. To satisfy the compatibility constraints, we have to embed additional information into the states to avoid assigning a set of jobs to a machine for which no possible configuration is possible. These constraints are embedded into a labelling algorithm, which is used to produce feasible assignments of jobs to a machine. From an algorithmic point of view, this calls for a new specific labelling process that we propose in this paper. In our method, new specific resource consumption and dominance checks have to be designed to account for this specificity.

Dynamic programming based arc-flow formulations for cutting and packing problems – LUIS MARQUES, INRIA CENTER OF UNIVERSITY OF BORDEAUX –

We study arc-flow formulations obtained from transition graphs or transition hypergraphs of dynamic programs. These formulations are known for their generally strong linear relaxations and can be solved directly by a general purpose mixed integer linear programming (MILP) solver. An advantage of the MILP formulation is that one can include additional linear constraints in the form of resource constraints. In the absence of these, arc-flow formulations in graphs – resp. hypergraphs – form a totally unimodular – resp. totally dual integral – system, meaning the solution of the linear relaxation of the formulation is integer. A drawback of the MILP formulation is its size which can be pseudo-polynomial or exponential. We focus on cutting and packing problems with temporal constraints, i.e., temporal knapsack problem, temporal bin packing problem and their variants. We introduce preprocessing techniques and dominance rules to reduce the size of the transition graphs and hypergraphs: symmetries reduction, equivalent states aggregation using methods from the literature and detection of states that can be replaced by other states where some constraints are relaxed. Numerically, we confirm the strength of the linear relaxation of the arc-flow formulations, we observe that the sizes of the arc-flow formulations are significantly reduced by our techniques and show that these formulations are competitive and can outperform popular compact formulations from the literature.

2a - TB915 - 513B - **Mixed Integer Linear Programming 1** Chair: QUOC TUNG LE

A column generation approach for a challenging real-world routing problem – NILS HASSAN QUTTINEH, LINKÖPING UNIVERSITY – Marduch Tadaros, Athanasios Migdalas, Torbjörn Larsson

We present the Hierarchical Multi-Switch Multi-Echelon Vehicle Routing Problem, a new variant of the well-known Vehicle Routing Problem. It is a real-world problem originating from the policies of a Nordic distribution company. The problem includes a single depot, a non-predetermined hierarchy of intermediate facilities, and two different fleets, consisting of homogeneous depot and homogeneous local vehicles, which are pulling swap-bodies. Depot vehicles with attached swap bodies depart from the central depot. They can either visit customers directly if only one swap-body is attached or visit one or two consecutive switch points in order to transfer one or two loaded swap-bodies to a corresponding number of local vehicles, which are subsequently routed to customers while the depot vehicle itself proceeds to serve customers with the remaining loaded swap-body. Two column generation formulations of the problem are proposed and evaluated, both with respect to their respective bounding quality and best found feasible solutions. The main focus is on the excellent lower bounding quality from the column generation models. As a bonus, we find good integer solutions.

Complexity of Determining Unboundedness in Linear Bilevel Optimization – BÁRBARA RODRIGUES, UNIVERSITY OF EDINBURGH – Margarida Carvalho, Miguel Anjos

Despite their widespread applications, bilevel problems are known to be challenging to solve in practice, as even in the linear case bilevel problems are strongly NP-hard. However,

unboundedness in these problems is often overlooked, with research typically assuming boundedness of their feasible sets. This presentation addresses this gap by exploring unboundedness in linear bilevel optimization and its computational complexity. We introduce the notion of direction of unboundedness to linear bilevel programs, and show that determining unboundedness in the case without upper-level constraints is in NP. Additionally, we prove that the decision problem of knowing whether a linear bilevel program with linking upper-level constraints is unbounded is NP-hard.

On the geometric and computational complexity of polynomial bilevel optimization – QUOC TUNG LE, TOULOUSE SCHOOL OF ECONOMICS – Jérôme Bolte, Edouard Pauwels, Samuel Vaiter

Bilevel optimization is an important mathematical tool to model phenomena in many domains, such as economic game theory, decision science, and machine learning, to name but a few. Despite its importance, efficient and scalable algorithms for bilevel optimization are mostly developed for the (strong) convexity of the lower-level problem case, which is unrealistic for many practical tasks. In the quest to understand more general bilevel problems, we relax the lower level strong convexity and consider polynomial bilevel optimization, i.e., polynomial objective functions and constraints. We focus on the worst-case analysis of this class of problems, from both geometric and computational viewpoints. Our analysis suggests that even the algebraic rigidity of polynomials does not exclude extreme pathologies induced by the bilevel optimization. More specifically, we demonstrate that any semi-algebraic function can be represented as the objective of a polynomial bilevel problem. This discovery implies that solving polynomial bilevel optimization is equivalent to optimizing general semi-algebraic functions. We obtain other sharp variations of this result by considering relevant properties of the lower problem, such as convexity or feasible set compactity. In addition, we show the Σ_2^P -hardness of polynomial bilevel optimization, characterizing polynomial bilevel problems as vastly more challenging than NP-complete problems (under reasonable hardness assumptions).

2a - WA231 - 511E - Bilevel Optimization with Discrete Decisions (I) Chair: MARTINA CERULLI



Playing Stackelberg Security Games in perfect and imperfect formulations – MARTIN LABBE, FREE UNIVERSITY OF BRUSSELS (ULB) – Pamela Bustamante Faundez, Víctor Bucarey, Vladimir Marianov Kluge, Fernando Ordóñez

Anticipating the strategies of potential attackers is crucial for protecting critical infrastructure. We can represent the challenge of the defenders of such infrastructure as a Stackelberg security game. The defender must decide how to allocate limited resources to protect specific targets, aiming to maximize their expected utility (such as minimizing the extent of damage) and considering that attackers will respond in a way that is most advantageous to them. We consider two different cases according to whether the defender's set of mixed strategies can be described with a perfect formulation or not. For both cases, we present exact algorithms to determine the optimal mixed strategies of the defender and validate them through computa-

tional experiments. In particular, we solve a Stackelberg security game that aims to protect targets with a defined budget, i.e. the defender pure strategies must satisfy a knapsack constraint. This game is solved with a branch and price approach.

A Toll-Setting Problem with Gamma-Robust Wardrop Equilibrium Conditions – YASMINE BECK, TRIER UNIVERSITY – Martin Labbe, Martin Schmidt

We consider a multi-commodity traffic network in which a toll-setting authority aims to maximize revenues by imposing tolls on certain arcs of the network. Users of the traffic network act in the sense of Wardrop's user equilibrium so that their individual travel costs are minimized. The problem can be seen as a single-leader multi-follower problem with the toll-setting authority acting as the leader and the network users acting as the followers. We model this setting as a mathematical problem with equilibrium constraints (MPEC), for which we present a single-level mixed-integer nonlinear reformulation. Moreover, we consider the setting in which the network users face uncertainties regarding their travel costs, which we address using a Gamma-robust approach. Our preliminary numerical results illustrate the impact of addressing uncertainties in comparison to nominal solutions.

BOBILib - A Bilevel Optimization Benchmark Instance Library – MARTIN SCHMIDT, TRIER UNIVERSITY – Thomas Kleinert, Ivana Ljubic, Ted Ralphs, Johannes Thürauf

The field of computational mixed-integer linear bilevel optimization made significant advances over the last 15 years. Many new problem classes have been tackled and a significant number of algorithmic techniques have been introduced. Consequently, we can solve new types of or simply larger problems compared to what was possible in the 2000s. However, what was missing up to now, was a well-curated and publicly available test-set of instances that the community can use to develop, test, and compare novel methods. In this talk, we present the Bilevel Optimization Benchmark Instance Library (BOBILib), which aims for filling the above mentioned gap. BOBILib contains over 2000 instances ranging from general mixed-integer linear bilevel problems to specific interdiction problems. In this talk, we will discuss the structure of the library, the type of instances involved, and the data format in which the problems are given.

A Branch-and-Cut Algorithm for Mixed-Integer Bilevel Linear Optimization Based on Improving Directions – FEDERICO BATTISTA, LEHIGH UNIVERSITY –

Mixed integer bilevel linear optimization problems aim to determine the optimal strategy of the leader in a two-player game, while considering that a follower will also react optimally in response. To adapt the well-known branch-and-cut framework widely used for solving mixed integer linear optimization problem (MILP) to the bilevel setting, it requires an oracle for checking whether the response obtained by solving a relaxation is optimal to follower's problem, given a strategy for the leader. Typically, this is accomplished by solving follower's problem (which is an MILP) to optimality. In this work, we propose a closely related alternative, relying instead on an oracle for determining the existence of an improving feasible direction.

The advantage of such an oracle is that it unifies the feasibility check with generation of strong valid inequalities in a way that we conjecture will lead to improvements in empirical performance. Preliminary numerical results are discussed using a modified version of the open source solver MiBS.

2a - WA916 - 513B - Mixed Integer Linear Programming 2 Chair: VANESSA ANN BEDDOE SAWKMIE

Lifting of the cover inequalities for the robust knapsack problem – YOUNGJOO ROH, SEOUL NATIONAL UNIVERSITY – Junyoung Kim, Kyungsik Lee
In this study, we consider the robust knapsack problem introduced by Bertsimas and Sim [Operations Research 52(1) pp.35-53, 2004]. Robust cover inequalities are well-known valid inequalities for the robust knapsack problem. These inequalities can be strengthened using lifting methods that improve each coefficient of inequality by utilizing the optimal objective values of lifting problems. Since obtaining an optimal solution for every lifting problem can be time-consuming, we propose an efficient alternative lifting method using the upper bounds of the lifting problems. Through computational tests, we demonstrate that our lifting method outperforms existing methods in terms of computational efficiency, while the effectiveness of the resulting lifted robust cover inequalities remains competitive compared to those generated by existing lifting methods.

A novel Pareto-optimal cut selection strategy for Benders Decomposition – FLORIAN ROESEL, FRIEDRICH ALEXANDER UNIVERSITÄT ERLANGEN NÜRNBERG –

Benders Decomposition (BD) has been used to solve a variety of optimization problems. The performance of BD is strongly influenced by the choice of the cuts to be added during the course of the algorithm.

Several algorithmic approaches to make this choice are known. The so-called Magnanti-Wong method aims for Pareto-optimal Benders cuts. Cuts based on minimal infeasible subsystems of a modified version of the Benders subproblem have proven to provide strong cuts. Recently, methods using facets of the subproblem's value function's epigraph as Benders cuts have been developed.

We contribute to the field of cut selection strategies for BD by developing a new notion of Pareto-optimality, that incorporates, in contrast to the Magnanti-Wong notion of Pareto-optimality, the behavior of the cuts on the whole feasible domain of the Benders masterproblem. We prove that cuts that are non-dominated according to our notion are non-dominated in the sense of Magnanti-Wong as well. We provide an approach to calculate non-dominated cuts according to our notion, that solves a single linear program that has one variable and one constraint more than the standard dual Benders subproblem.

We benchmark our cut selection strategy against the other known strategies on various instances. Some instances are taken from the MIPLib, others are multi-commodity flow network design problems, and others are randomly generated MIPs. The computational results show, that a hybrid selection strategy consisting of minimal infeasible system cuts and our

cut selection strategy is the most efficient in terms of running time. Further, the results show, that our cut selection strategy is the most efficient measured in the number of Benders cuts that are needed to solve a problem.

Bulk Optimization approaches to Graph Problems – ENRICO MALAGUTI, DEI UNIVERSITY OF BOLOGNA – Michele Monaci, Paolo Paronuzzi

We consider classical graph problems in a bulk optimization setting. In a nutshell, the idea is as follows: we want solve a graph problem (e.g., perfect matching or spanning tree), but there is uncertainty on the actual edges that will be available in different scenarios. Therefore, we identify a minimum cost subset of edges such that the resulting graph contains the structure we are interested in (i.e, a perfect matching or a spanning tree), under each scenario. Although the paradigm was studied for what concerns its theoretical properties, and some approximation algorithms were proposed, we are not aware of any computational approach. Therefore, we propose a unified approach based on Benders decomposition. In this approach, the master problem encapsulates variables determining which edges to activate, while the subproblems verify the admissibility of the master solution in each scenario. The transformation of the subproblems into efficiently solvable flow problems enhances computational tractability.

Route 'em and Count 'em : A Two-Stage Stochastic Programming Model in Undersea Warfare – VANESSA ANN BEDDOE SAWKMIE, UNIVERSITY OF WISCONSIN MADISON –

The detection of targets in undersea warfare requires successful detection by an active search asset. Maximizing detection likelihood requires strategic placement and routing of the search assets, each of which has unique search capabilities, in the decision theatre over the planning horizon. We develop a two-stage integer stochastic programming model designed to maximize the expected number of targets detected, considering both the stochastic nature of target movements and the probabilistic detection capabilities of the assets. Leveraging the problem structure, we develop a closed-form solution for the second stage of the stochastic program, which allows for large-scale instances of the model to be solved efficiently via Benders decomposition. We conduct computational experiments to demonstrate the efficacy of the approach.

2a - WB236 - 513D - Bilevel Optimization with Discrete Decisions (II) Chair: JOHANNES THÜRAUF

Network Relaxations for Discrete Bilevel Optimization under Linear Interactions – LEONARDO LOZANO, UNIVERSITY OF CINCINNATI – Andre Augusto Cire, David Bergman

Integer bilevel programming problems are known to be very challenging due to the lack of strong relaxations that can be efficiently computed. We propose single-level representations of integer bilevel programming problems that rely on network flow-based approximations of the follower's value function using decision diagrams. We then show how we can derive scalable relaxations from this representation by using relaxed decision diagrams, yielding dual bounds. We experimentally compare our approach with state-of-the-art bilevel programming solvers and show that we can obtain competitive results for

certain problem classes.

Using Disjunctive Cuts in a Branch-and-Cut Method to Solve Convex Integer Nonlinear Bilevel Problems – ANDREAS HORLÄNDER, TRIER UNIVERSITY – Ivana Ljubic, Martin Schmidt

We present a branch-and-cut method for solving convex integer nonlinear bilevel problems, i.e., bilevel models with nonlinear but convex objective functions and constraints in both the upper and the lower level. To this end, we generalize the idea of using disjunctive cuts to cut off integer-feasible but bilevel-infeasible points. These cuts can be obtained by solving a cut-generation problem, which itself is a single-level but nonconvex integer nonlinear problem. We show that this cut-generation problem can be decomposed into a series of smaller subproblems, that can all be solved in parallel and for which we can state assumptions under which all of them are convex integer nonlinear problems. Moreover, we develop additional algorithmic techniques such as tailored pruning rules to further speed up our method. We finally prove the correctness of the method and test it in a numerical study that shows the applicability of the method.

Interdiction of minimum spanning trees and other weighted matroids – NOAH WENINGER, UNIVERSITY OF WATERLOO – Ricardo Fukasawa

In the minimum spanning tree (MST) interdiction problem, the objective is to maximize the MST weight of a graph by deleting some edges. Each deleted edge incurs a cost, and there is a limit on the total expenditure. MST interdiction is known to be NP-complete even when all costs are 1 and weights are 0 or 1 (Fredrickson and Solis-Oba 1999). Since MSTs are minimum weight bases in the graphic matroid, this problem can be seen as a special case of the matroid interdiction problem, in which the objective is to maximize the minimum basis weight of a matroid by deleting some elements.

We develop a new branch-and-bound framework for solving matroid interdiction to optimality. For the case of MST interdiction, we show how to compute strong upper bounds by using dynamic programming over the solutions to minimum cut problems in related graphs. Furthermore, we show that a similar dynamic programming algorithm is able to exactly solve partition matroid interdiction, showing that partition matroid interdiction is only weakly NP-hard. For other matroids, we provide a set of tools that can be utilized to derive similar results.

Our algorithm achieves state-of-the-art computational performance for MST interdiction. In our preliminary computational tests, our algorithm was able to solve all instances from the literature in only a few minutes; many of these instances could not be solved within 3 hours by the previous best algorithms (Wei et al 2021). Our algorithm quickly solves typical instances with thousands of edges. We hope to apply our framework to more matroids and related problems in the future.

Combinatorial pricing problems: methods using embedded dynamic programming models – MARGARIDA CARVALHO, UNIVERSITÉ DE MONTRÉAL – Quang Minh Bui, José Neto

A combinatorial pricing problems (CPP) is a bilevel problem where we have a finite set of items, for which a leader sets

the price of some of them (i.e. tolled items). Then, a follower observes the prices set by the leader and determines a subset of items to select by solving a combinatorial optimization problem. The latter must be expressed as a binary linear program. We take the perspective of the leader, whose objective is to set the prices of tolled items in order to maximize the revenue related to the items selected by the follower. In this talk, we present a single-level reformulation of the CPP by rewriting the follower's problem as a longest path problem using a dynamic programming model. In particular, we show the reformulation using the well-known decision diagrams and a novel dynamic programming model called the selection diagrams. Due to the single-level reformulations non-compact size, we provide a cutting-plan methodology for solving them. Finally, we validate our general-purpose framework by testing it on different CPPs.

2a - WB917 - 513B - Mixed Integer Linear Programming 3 Chair: DAVID WARME

A block decomposition algorithm for an extension of the vertex cover problem – JUAN FRANCISCO CORRECHER, UNIVERSITAT DE VALÈNCIA – Mercedes Landete, Juanjo Peiró, Hande Yaman

The relationship between locations and their connections can be sometimes modelled using a generalization of the vertex cover problem. In this talk, we will present a block decomposition algorithm for a variant of the problem and discuss some preliminary numerical results. The efficiency of this method will be compared to non-decomposition procedures on several computational experiments. Finally, we will also propose further improvements and research directions.

Preliminary results on an extension of the vertex cover problem – JUANJO PEIRÓ, UNIVERSITAT DE VALÈNCIA – Juan Francisco Correcher, Mercedes Landete, Hande Yaman

In some contexts, we can model the relationship between locations and their connections using some extensions of the vertex cover problem. In this talk, we revisit some of these extensions, provide preliminary results that can be used for solutions methods, and try to connect them with well-known polyhedra such as the boolean quadric polytope. We present computational results from a branch-and-cut algorithm in which we show the applicability of our proposals.

Analyzing the Sensitivity of Integer Linear Programs via Optimization Oracles – ERIK JANSEN, TU DARMSTADT – Marc Pfetsch

We are interested in determining the set of objective vectors under which the given integer solution stays optimal. The set of all such objective vectors forms the normal cone at the solution with respect to the integer hull. Since the integer hull is not known in general and expensive to calculate, we are looking for an efficient way to calculate such cones or rather its polar, the radial cone.

We present an approach which calculates all facet-defining inequalities of the radial cone. The algorithm combines existing algorithms modified to fit the purpose. Since we only require the existence of an optimization oracle, we can calculate the radial cone for all linear optimization problems over polyhedra, including mixed integer problems.

The general idea of our algorithm is to utilize an optimization oracle to calculate – preferably neighboring – solutions and incrementally extend a current sub cone of the radial cone with these solutions. We will address the handling of potential problematic cases such as unbounded directions and non-full-dimensional polyhedra. Finally, we will use this algorithm to analyze the degeneracy of ILPs and present first computational results.

1,000,000 Point Steiner Tree Solved to Optimality – DAVID WARME, GROUP W –

Given a finite set N of points in the plane called "terminals," the Euclidean Steiner tree problem seeks a set of line segments connecting all the terminals having minimum total Euclidean length. The shortest such tree usually requires additional Steiner points serving as junctions that reduce the total interconnect length. In an optimal tree, Steiner points always have degree 3, with incident line segments meeting at 120-degree angles. This classical NP-hard problem is one of the oldest optimization problems in all of mathematics.

We describe a recently solved instance containing 1,000,000 terminals uniformly distributed in the unit square. It has been called "the largest classical NP-hard problem instance ever solved to optimality."

This computational breakthrough resulted from mathematical analyses called "strength indicators," which provide a quantitative assessment of the computational "strength" of a given inequality with respect to the underlying polyhedron over which optimization is performed. These indicators revealed previously unknown properties of the Spanning Tree in Hypergraph Polytope which were exploited to remarkable computational advantage.

2a - WC160 - 513B - **Network Design Chair:**
NAVNEET VIDYARTHI

ML-based Algorithm for Solving the Transmission Network Expansion Planning Problem – PABLO FELIPE OTEÍZA, UNIVERSITY OF SANTA MARIA CHILE – Mehrdad Pirnia, Ricardo Alvarez

This presentation proposes a novel algorithm for solving the multi-year Transmission Network Expansion Planning (TNEP) problem, that combines the use of ML, together with Column Generation (CG) algorithm. Using the Column Generation technique, the TNEP problem is decomposed into a Master Problem, which is computationally inexpensive to solve, and several subproblems, one for each year, which are of MILP nature. To increase convergence of the CG algorithm, we solve a linearly relaxed version (LP) of the subproblems and trained a binary classifier using random forest for estimating the optimal values of the investment (binary) variables. Input data of the classifier are the results of the linearized version of the subproblems, which are characterized into key features with physical meaning. Then, the subproblems are optimized again, but fixing the values of the investment variables according to the output of the classifier, i.e., by resolving an LP formulation of the subproblems. This strategy allows solving the subproblems much faster than its MILP counterpart while providing, at the same time, a feasible solution, which is fed into the master problem. One key advantage of this strategy is that, upon convergence, the optimality of the solution is guaranteed. This

proposal was tested in a case study based on the HRP 38-bus test system for different scenarios. The results showed that the binary classifier reached an accuracy above 96% for estimating the value of the binary variables, thus providing good-quality solutions of the subproblems. This allowed, in turn, reduction in the computational time required for solving the TNEP problem in around 53%, which shows the great benefits of the proposed approach.

Optimizing Electric Vehicle Fleet Operations for Sustainable Last-mile Delivery Under Uncertainty – JAEHEE JEONG, IVEY BUSINESS SCHOOL – Bissan Ghaddar, Nicolas Zufferey, Jatin Nathwani

In recent years, electric vehicles (EVs) have emerged as a preferred mode of transportation due to their advantages over traditional fuel-based vehicles, particularly concerning environmental impact. However, despite their benefits, EV adoption faces limitations. Among those limitations, this work focuses on driving range uncertainty stemming from various exogenous and endogenous factors impacting energy consumption. To address this challenge, this talk introduces a two-stage robust optimization framework tailored for the electric vehicle routing problem. The proposed approach proposes a column-and-constraint generation framework incorporating variable neighborhood search and alternating direction techniques, to tackle the proposed optimization model. Computational experiments demonstrate the economic efficiency and robustness of the proposed framework.

Discrete Location Problem with Probabilistic Service Level Constraints – NAVNEET VIDYARTHI, CONCORDIA UNIVERSITY AND CIRRELT –

We study a general class of stochastic capacitated facility location problems arising in the design and reconfiguration of e-commerce supply chains. It considers probabilistic service level constraints to ensure the stochastic delivery times of customers' orders, defined as the sum of waiting time (including service time) at the processing facility plus shipping time to reach the customer location, is within a prescribed time limit with a probability greater or equal to a threshold value. We provide alternative polyhedral representations of these highly nonlinear and non-convex probabilistic constraints, and develop an exact branch-and-cut algorithm for solving the resulting reformulations. The proposed algorithm is enhanced with several algorithmic refinements to accelerate its convergence. We perform extensive computational experiments based on real location data from the United States to evaluate the performance of our algorithm and to provide insights on the optimal structure of solution networks. Results obtained on large-scale instances with up to 2,500 customers and 225 potential facilities under different service level scenarios confirm the effectiveness of the proposed algorithm.

2a - ThA237 - 513B - **Recent Advances in Combinatorial Optimization Chair:** AHMAD ABDI

Lagrangian relaxation and duality for multiobjective integer programming – SAUMYA SINHA, UNIVERSITY OF MINNESOTA –

Multiobjective integer (linear) programs (MOIPs), as the name suggests, simultaneously optimize multiple objective functions

over a set of linear constraints and integer variables. Relaxations of MOIPs provide bounds on their solutions and play an important role in solution algorithms. In this talk, we present a Lagrangian relaxation of an MOIP and discuss its properties. Specifically, we show that there exist sparse Lagrangian multipliers that satisfy a complementary slackness property and generate tight relaxations at supported solutions. At unsupported solutions, where the convex hull relaxation is not tight either, we show that a Lagrangian relaxation improves upon the convex hull bounds. We use the Lagrangian relaxation to define a Lagrangian dual of an MOIP and establish weak duality. The dual is strong at supported solutions under certain conditions on the primal feasible region which are identical to the single-objective case.

On the Number of Degenerate Simplex Pivots – KIRILL KUKHARENKO, OTTO VON GUERICKE UNIVERSITY MAGDEBURG – Laura Sanità

The simplex algorithm is one the most popular algorithms to solve linear programs (LP). Starting at an extreme point solution of an LP, it performs a sequence of basis exchanges (called pivots) that allows one to move to a better extreme point along an improving edge-direction of the underlying polyhedron. A key issue in the simplex algorithm's performance is degeneracy, which may lead to a (potentially long) sequence of basis exchanges which do not change the current extreme point solution. In this paper, we prove that it is always possible to limit the number of consecutive degenerate pivots that the simplex algorithm performs to $n-m-1$, where n is the number of variables and m is the number of equality constraints of a given LP in standard equality form.

A better-than- $k/2$ -approximation for weighted k -Set Packing – MEIKE NEUWOHNER, LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE –

The weighted k -Set Packing problem is defined as follows: As input, we are given a collection S of sets, each of cardinality at most k , and a positive weight for each set in S . The task is to compute a subcollection A of S such that the sets in A are pairwise disjoint, and the total weight of A is maximum. The case $k=2$ is equivalent to the Maximum Weight Matching problem. For $k \geq 3$, already the special case of unit weights, the unweighted k -Set Packing problem, is NP-hard.

The technique that has proven most successful in designing approximation algorithms for both the unweighted and the weighted k -Set Packing problem is local search. For the unweighted problem, the state-of-the-art is a polynomial-time $(k+1)/3 + \epsilon$ -approximation by Fürer and Yu that takes into account certain local improvements of up to logarithmic size. For general weights, we recently showed that by considering a special class of local improvements of logarithmically bounded size, we can obtain a polynomial-time $(k + \epsilon \log k)/2$ -approximation, where $\epsilon \log k$ converges to 0 as k approaches infinity. We further proved that this result is asymptotically best possible: For general weights, one cannot achieve a better guarantee than $k/2$ by only considering local improvements of logarithmically bounded size.

At a first glance, this seems to conclude the story of local improvement algorithms for the weighted k -Set Packing problem. However, in this talk we show how to employ a black box algorithm for the unweighted k -Set Packing problem to gen-

erate local improvements of super-logarithmic size and obtain a polynomial-time $0.4986 \cdot k + 0.5194$ -approximation algorithm for weighted k -Set Packing. Our techniques further allow us to establish a general connection between the approximation guarantees achievable for unit and general weights.

This result has been published in the proceedings of SODA 2023.

2a - ThB161 - 513D - **Advances in ILP: Branch and cut** Chair: ALINSON SANTOS XAVIER

Super-Polynomial Lower Bounds for Branch-and-Bound with General Disjunctions via Interpolation – MARC PFETSCH, TU DARMSTADT – Max Gläser

We investigate linear programming based branch-and-bound using general disjunctions and derive the first super-polynomial exponential lower bound (in the encoding length of the integer program) for the size of a general branch-and-bound tree for a particular class of integer programs. This is achieved by showing that general branch-and-bound admits quasi-feasible monotone real interpolation. Specialized to a certain integer program expressing that no graph can have both a k -coloring and a $(k+1)$ -clique, this property states that any branch-and-bound tree for this program can be turned into a monotone real circuit of similar size, distinguishing between graphs with a k -coloring and graphs with a $(k+1)$ -clique. This allows to utilize known lower-bounds for such circuits. Moreover, using the recently introduced notion of infeasibility certificates of Hrubes and Pudlák, we can show that it is exponentially hard to prove infeasibility of certain random CNFs by general branch-and-bound with high probability via the same technique.

Compressing Branch-and-Bound Trees – JOSEPH PAAT, UNIVERSITY OF BRITISH COLUMBIA – Gonzalo Muñoz, Alinson Santos Xavier

A branch-and-bound (BB) tree certifies a dual bound on the value of an integer program. In this work, we introduce the tree compression problem (TCP): Given a BB tree T that certifies a dual bound, can we obtain a smaller tree with the same (or stronger) bound by either (1) applying a different disjunction at some node in T or (2) removing leaves from T ? We believe such post-hoc analysis of BB trees may assist in identifying helpful general disjunctions in BB algorithms. We initiate our study by considering computational complexity and limitations of TCP. We then conduct experiments to evaluate the compressibility of realistic branch-and-bound trees generated by commonly-used branching strategies, using both an exact and a heuristic compression algorithm.

Polyhedrality made easy: Using general cut operators to determine when cut closures are polyhedral – ANTONIA CHMIELA, ZUSE INSTITUTE BERLIN – Gábor Braun, Sebastian Pokutta

Cutting planes are a crucial tool in mixed-integer programming. When measuring the strength of a specific class of cutting planes, a key question is whether the corresponding cut closure is polyhedral, meaning that finitely many cuts are sufficient to obtain the closure. The vast amount of literature investigating polyhedrality for different kinds of cutting planes indicates that the field has not yet settled on a universal solution to this problem. In this talk, we will look at cutting plane

generation through a unifying framework and establish conditions that ensure polyhedrality of cut closures. This enables us to prove polyhedrality for a broad range of cutting plane families more easily.

Learning to Generate Cutting Planes with MIPLearn – ALINSON SANTOS XAVIER, ARGONNE NATIONAL LABORATORY – Berkay Becu, Santanu Dey, Feng Qiu

While many classes of cutting planes described in the literature have been effective in terms of gap closure, they are often computationally demanding to generate and offer only modest improvements in running times. Taking into account that many discrete optimization problems are solved repeatedly with only slight variations in input data, in this talk we investigate the usage of machine learning (ML) methods to accelerate the selection and generation of such cuts. We start by introducing MIPLearn, an extensible open-source framework which uses machine learning (ML) to enhance the performance of state-of-the-art MIP solvers. The framework is compatible with multiple MIP solvers (e.g. Gurobi, CPLEX, SCIP, HiGHS), multiple modeling languages (JuMP, Pyomo, gurobipy) and supports user-provided ML models. We then present computational experiments on using ML to accelerate the generation of rank-1 Gomory mixed-integer (GMI) cuts from multiple tableau bases.

2a - ThB918 - 513B - **Mixed Integer Linear Programming 4** Chair: VENKAT AKHIL ANKEM 

Strong Cutting Planes for the Feature Selection Problem – RENAUD CHICOISNE, UNIVERSITÉ CLERMONT AUVERGNE / LIMOS – Pierre Latouche, Rui Ma

The feature selection problem (FS) consists of determining a subset of features representing a dataset, that concentrate only the most relevant information the dataset contains. FS can be cast as a nonlinear, mixed-integer mathematical program with on-off constraints. Because of the presence of so-called big-M constraints, the continuous relaxation of FS is usually very weak, making the Branch & Bound process remarkably slow on real-world instances. In order to strengthen the model, we present new relaxations for FS that keep the integrality of a reasonably small subset of variables, and show an efficient algorithmic framework to separate a fractional incumbent from them. Preliminary results on small instances are presented to show the efficiency of the cutting planes generated.

The undirected circular facility layout problem – LOUISA SCHROEDER, TU DORTMUND UNIVERSITY – Anja Fischer

In the directed circular facility layout problem (DCFLP) one has given a set of departments, their lengths and pairwise weights between the departments. In contrast to the single row facility layout problem (SRFLP), where departments are arranged along one side of a path without overlapping, the DCFLP looks for an arrangement on a circle such that the sum of the weighted center-to-center distances measured in clockwise direction is minimized.

In this talk, we present a new facility layout problem, the undirected circular facility layout problem (UDCFLP). For

each pair of departments we consider here the shortest distance along the circle, either travelling in clockwise or in counter-clockwise direction.

We present two new mixed-integer linear programming models for the UDCFLP, whereby one model is an adaptation of a well-known model for the DCFLP and the other model is based on the well-known betweenness model for the SRFLP. We develop symmetry-breaking constraints and study the structure of the associated polytopes. Our models, which are strengthened by our new cutting planes, allow us to solve instances with up to 13 departments to optimality within a time limit of one hour. For larger instances we derive strong lower bounds exploiting the connections to some parallel machine scheduling problem.

Inexact Flow Decomposition in Quasispecies Network via Integer Linear Programming – FERNANDO DIAS, AALTO UNIVERSITY –

Flow Decomposition is the problem of decomposing a flow network into a set of paths that can perfectly replicate it. As a minimization problem (Minimal Flow Decomposition—MFD), it consists of finding the smallest set of paths that justify such conditions. Such a problem is very common in Computer Science and Bioinformatics, where many of its variants are present in multi-assembly tools. This problem, even in its most primordial form, is an NP-hard problem, hindering larger-scale applications due to the lack of an efficient polynomial solver. In practical applications, several heuristics and approximations have been attempted, as well as different ILP formulations, all with different levels of success.

In simple terms, multiple genomic sequences are reconstructed from a set of reads, and a weighted graph (known as a splicing graph) is generated. Then, the multi-assembly can be described as the problem where a set of weighted paths, each path equivalent to a genomic sequence in a sequencing sample, whose superposition perfectly explains the weight of the graph. Because such graphs are obtained from real data, their weights do not always form a flow.

In this work, we provide an ILP formulation for IMFD (Irregular Minimum Flow Decomposition). In this problem, we assume that each edge has, besides its weight, a confidence rate relative to its presence in the splicing graph, such that a higher confidence rate means that such an edge is more likely to have been correctly predicted as part of such a graph. Our goal is to provide an efficient flow decomposition method that can simultaneously obtain the minimal set of paths with their respective weights and the highest combined confidence rate among selected edges. In this way, the decomposition is minimal and most accurate based on the initial conditions.

Our approach has shown potential for solving instances with up to 200 nodes within 30 seconds on real flow and simulated data. We also show that our ILP formulation is flexible and adaptable enough to incorporate many practical variants (such as subpath constraints). We hope our results can contribute to better methods to mitigate the complexity of the multi-assembly problem and improve the efficiency of future practical tools.

Production scheduling of open pit mines using new cutting planes – VENKAT AKHIL ANKEM, POLYTECHNIQUE MONTREAL, GERAD – Guy De-

saulniers, Michel Gamache, Vincent Raymond

In this article, we address the open pit mine production scheduling problem (OPMPSP) by considering the discretized ore body representation known as the block model. The objective is to strategically determine the block sequence, extraction timing, quantity, and processing decisions over a predefined planning horizon while adhering to the operational constraints and maximizing the net present value of the profit. We explore OPMPSP, accounting for capacity, blending, production, and material flow constraints. It is generally modeled with mixed-integer linear programming, but the sheer volume of blocks, and complexity results in high computation time for large instances. Hence, this article proposes incorporating new cutting planes into the model to strengthen the formulation through the addition of user-defined cuts developed through a combination of production capacity constraints and precedence relationships amongst the blocks. The inclusion of these cuts substantially reduces the computational demands for achieving the optimal integer solution resulting in lower computation time. The effectiveness of the developed new cuts is evaluated over real mine problem datasets. These computational results are compared with cuts present in the existing academic literature for comparative analysis. The combination of cuts is later integrated into the model supplemented with a sliding window heuristic to assess the overall impact. Thus the article contributes valuable insights and practical approaches to address the complexities inherent in large-scale open-pit mining production scheduling problems.

2a - ThC180 - 513B - **Advances in Mixed Integer Programming Theory** Chair: DANIEL DADUSH

The Subspace Flatness Conjecture and Faster Integer Programming – THOMAS ROTHVOSS, UNIVERSITY OF WASHINGTON – Victor Reis

In a seminal paper, Kannan and Lovász (1988) considered a quantity $\mu_{KL}(\Lambda, K)$ which denotes the best volume-based lower bound on the *covering radius* $\mu(\Lambda, K)$ of a convex body K with respect to a lattice Λ . Kannan and Lovász proved that $\mu(\Lambda, K) \leq n \cdot \mu_{KL}(\Lambda, K)$ and the Subspace Flatness Conjecture by Dadush (2012) claims a $O(\log n)$ factor suffices, which would match the lower bound from the work of Kannan and Lovász. We settle this conjecture up to a constant in the exponent by proving that $\mu(\Lambda, K) \leq O(\log^3(n)) \cdot \mu_{KL}(\Lambda, K)$. Our proof is based on the Reverse Minkowski Theorem due to Regev and Stephens-Davidowitz (2017). Following the work of Dadush (2012, 2019), we obtain a $(\log n)^{O(n)}$ -time randomized algorithm to solve integer programs in n variables. Another implication of our main result is a near-optimal *flatness constant* of $O(n \log^3(n))$.

Integer programs with nearly totally unimodular matrices: proximity – STEFAN WELTGE, TECHNICAL UNIVERSITY OF MUNICH – Manuel Aprile, Samuel Fiorini, Gwenaël Joret, Stefan Kober, Michal Seweryn, Yelena Yuditsky

It is a notorious open question whether integer programs (IPs), with an integer coefficient matrix M whose subdeterminants are all bounded by a constant in absolute value, can be solved in polynomial time. We answer this question in the affirmative if we further require that, by removing a constant number of

rows and columns from M , one obtains the transpose of a network matrix. We achieve our result in two main steps, the first related to the theory of IPs and the second related to graph minor theory. In this talk, we focus on the first part: We derive a new proximity result for the case where M is a general totally unimodular matrix and show how it can be used algorithmically in our context.

Integrality Gaps for Random Integer Programs via Discrepancy – DANIEL DADUSH, CWI AND UTRECHT UNIVERSITY – Dan Mikulincer, Sander Borst

We prove new bounds on the additive gap between the value of a random integer program $\max c^T x, Ax \leq b, x \in \{0, 1\}^n$ with m constraints and that of its linear programming relaxation for a wide range of distributions on (A, b, c) . We are motivated by the work of Dey, Dubey, and Molinaro (SODA '21), who gave a framework for relating the size of Branch-and-Bound (B&B) trees to additive integrality gaps. Dyer and Frieze (MOR '89) and Borst et al. (Mathematical Programming '22), respectively, showed that for certain random packing and Gaussian IPs, where the entries of A, c are independently distributed according to either the uniform distribution on $[0, 1]$ or the Gaussian distribution $N(0, 1)$, the integrality gap is bounded by $O_m(\log^2(n)/n)$ with probability at least $1 - 1/n - e^{-\Omega(m)}$. In this paper, we generalize these results to the case where the entries of A are uniformly distributed on an integer interval (e.g., entries in $\{-1, 0, 1\}$), and where the columns of A are distributed according to an isotropic logconcave distribution. Second, we substantially improve the success probability to $1 - 1/\text{poly}(n)$, compared to constant probability in prior works (depending on m). Leveraging the connection to Branch-and-Bound, our gap results imply that for these IPs B&B trees have size $n^{\text{poly}(m)}$ with high probability (i.e., polynomial for fixed m), which significantly extends the class of IPs for which B&B is known to be polynomial. Our main technical contribution is a new linear discrepancy theorem for random matrices. Our theorem gives general conditions under which a target vector is equal to or very close to a $\{0, 1\}$ combination of the columns of a random matrix A . The proof uses a Fourier analytic approach, building on work of Hoberg and Rothvoss (SODA '19) and Franks and Saks (RSA '20).

2a - FA228 - 513B - **Integer programming: structure and algorithms** Chair: MANUEL APRILE

Parameterized Algorithms for Block-Structured Integer Programs with Large Entries – ALEXANDRA ANNA LASSOTA, EINDHOVEN UNIVERSITY OF TECHNOLOGY –

This talk focuses on block-structured integer programs, in particular on n -fold IPs (where the constraint matrix has small, independent blocks linked together by few global constraints) and 2-stage stochastic IPs (where the constraint matrix has small, independent blocks linked together by few global variables). I will introduce this field and give an overview on the recent developments, focusing on our latest results: the parameterized tractability results for two-stage stochastic and n -fold programs persist even when one allows large entries in the global part of the program (SODA 24). This is joint work

with Jana Cslovjcek, Martin Koutecky, Alexandra Lassota, Michal Pilipczuk, Adam Polak.

Cooperative Integer Programming Games – HYUNWOO LEE, VIRGINIA TECH – Robert Hildebrand

This paper explores the development of a novel framework for identifying multiple cores in cooperative integer programming games, significantly extending traditional core-finding approaches that typically focus on a single core based on the grand coalition. By integrating a general framework with the practical application of the cooperative knapsack game (CKG), we demonstrate the model's ability to handle various scenarios where traditional methods may deem the core to be empty. Our methodology leverages a new separation oracle and a cutting-plane algorithm designed specifically for finding multiple cores, effectively addressing the limitations of existing axiomatic and rule-based approaches. The computational experiments underscore the capability of our approach to efficiently resolve instances with varying numbers of knapsacks and items, illustrating potential scalability.

Total Matching and Subdeterminants – LUCA FERRARINI, ÉCOLE DES PONTS PARISTECH – Samuel Fiorini, Yelena Yuditsky, Stefan Kober

A total matching of a graph G is a subset of G such that its elements, i.e. vertices and edges, are pairwise not adjacent. In this context, the Total Matching Problem (TMP) calls for a total matching of maximum size. This problem generalizes both the Stable Set Problem, where we look for a stable set of maximum size, and the Matching Problem, where instead we look for a matching of maximum size. In this talk, we consider the natural integer programming formulation of the TMP and study it through the prism of subdeterminants. First, we focus on the case of the forests. We give an alternative characterization of the maximum subdeterminant of the constraint matrix, as well as an approximation of it. Then, we switch to general graphs and give an FPT algorithm for computing the maximum subdeterminant. Finally, we show that the total matching problem can be solved in strongly polynomial time when the maximum subdeterminant is bounded by a constant.

Integer programs with nearly totally unimodular matrices: the cographic case – STEFAN KOBER, UNIVERSITÉ LIBRE DE BRUXELLES – Manuel Aprile, Samuel Fiorini, Gwenaël Joret, Michal Seweryn, Stefan Weltge, Yelena Yuditsky

It is a notorious open question whether integer programs (IPs), with an integer coefficient matrix M whose subdeterminants are all bounded by a constant in absolute value, can be solved in polynomial time. We answer this question in the affirmative if we further require that, by removing a constant number of rows and columns from M , one obtains the transpose of a network matrix. We achieve our result in two main steps, the first related to the theory of IPs and the second related to graph minor theory. In this talk, we focus on the second part: We reformulate the problem as a maximum constrained integer potential problem on a minor-closed class of rooted graphs. We leverage this to obtain a tree-decomposition into highly structured graphs for which we can solve the problem locally, and combine the local solutions via dynamic programming.

2a - FA262 - 513D - Aspects of Symmetry and Fairness in Integer Programming and Combinatorial Optimization Chair: CHRISTOPHER HOJNY

FPT Algorithms using Minimal Parameters for a Generalized Version of Maximin Shares

– ADRIAN VETTA, MCGILL UNIVERSITY –

We study the computational complexity of fairly allocating indivisible, mixed-manna items. For basic measures of fairness, this problem is hard in general. Ergo, research has flourished concerning input classes where efficient algorithms exist, both for the purpose of establishing theoretical boundaries and for the purpose of designing practical algorithms for real-world instances. Notably, the paradigm of fixed-parameter tractability (FPT) has led to new insights and improved algorithms for a variety of fair allocation problems; see, for example, Bleim et al. (IJCAI'16), Aziz et al. (AAAI'17), Bredereck et al. (EC'19) and Kulkarni et al. (EC'21).

Our focus is the fairness measure *maximin shares* (MMS). Motivated by the general non-existence of MMS allocations, Aziz et al. (AAAI'17) studied *optimal MMS allocations*, namely solutions that achieve the best alpha-approximation for the maximin share value of every agent. These allocations are guaranteed to exist, prompting the important open question of whether optimal MMS allocations can be computed efficiently. We answer this question affirmatively by providing FPT algorithms to output optimal MMS allocations. Furthermore, our techniques extend to find allocations that optimize alternative objectives, such as minimizing the additive approximation, and maximizing some variants of global welfare.

In fact, all our algorithms are designed for a more general MMS problem in machine scheduling. Here, each mixed-manna item (job) must be assigned to an agent (machine) and has a processing time and a deadline. This generalization is of independent interest as it induces models for important applications such as ride-hailing services and shift scheduling. We develop efficient algorithms running in time FPT. Formally, we present polynomial time algorithms w.r.t. the input size multiplied by some function dependent on the parameters that yield optimal maximin-value approximations (among others) when parameterized by (i) the number of items, (ii) the maximum number of items in any group (which will be assigned independently), the number of agents, and the maximum absolute utility value, (iii) the number of agents, the number of different deadlines, the maximum processing time, and the maximum absolute utility value.

To obtain these results, we utilize machinery from integer programming. In particular, we employ the framework of n -fold integer programs – these are integer programs with a specific structure that allows them to be solved in time FPT. Our results are tight. We prove that any stronger parameterization is NP-hard even if the parameters only have constant value. Hence, under common complexity theoretical assumptions, there exists no FPT time algorithm. Consequently, this work gives a complete complexity theoretical picture and efficient algorithms for fairly allocating indivisible, mixed-manna items under the notion of maximin share fairness.

Handling symmetries in stable set problems

and perfect mixed graphs – ANNIKA JAEGER, TU DARMSTADT – Christopher Hojny, Marc Pfetsch, José Verschae

We consider a generalization of the stable set problem on mixed graphs. This adaptation enables the consideration of specific additional node relations which are expressed through directed arcs in the graph. Building upon previous research on the generalized stable set problem on mixed graphs, we investigate stable set problems on symmetric graphs. One common approach to handle symmetry in order to improve solution performance is via symmetry handling inequalities (SHIs). Choosing Schreier-Sims (SST) cuts as SHIs, we obtain inequalities that agree with the interpretation of directed edges in mixed graphs for the stable set problem. In the special case of undirected graphs with additional corresponding SST cuts, we demonstrate the simplification of the transformation of mixed graphs into closed mixed graphs. A central question that arises is how the addition of SST cuts influences the complexity of the stable set problem. In this context, we examine how SST cuts should be added to a trivially perfect graph such that the resulting mixed graph remains perfect and discuss different conditions.

Fair Schedules for Single Round Robin Tournaments with Ranked Participants – STEN WESSEL, EINDHOVEN UNIVERSITY OF TECHNOLOGY – Cor Hurkens, Frits Spieksma

We consider schedules for single round-robin tournaments on an even number of players, where each of the players plays every other player exactly once, either at home or away. We assume that the players are sorted on a predetermined ranking, where the first player is the strongest, and the last player is the weakest player of the competition. Such a ranking is sometimes directly available, such as with Elo rating scores for chess players, or can be derived from the final standings of previous editions of the tournament. With only one match between any pair of opponents, one team will have the asymmetric advantage of playing home (or in the case of chess, playing with white), which we consider as having positive effect on the match outcome for that player.

Designing schedules with desirable properties has received a lot of attention in the literature, such as minimizing the number of home and away breaks. Breaks are occurrences in the schedule where a player plays two home or two away matches in consecutive rounds. For each player, the home and away matches ordered by the round in which they are played, define a home/away pattern (HAP), which we refer to as a regular HAP.

Given a ranking of the players, we now also define a ranking HAP for each player, where the home and away matches for that player are ordered by the rank of their opponents. Intuitively, we want for a player to have the number of home and away matches balanced between opponents of a similar rank. For example, we consider it unfair for a player to play many home matches against top-ranked opponents, or to play many home matches against low-ranked opponents. We refer to this as ranking fairness. In this work, we define a ranking-fairness measure and show under which conditions a ranking-fair schedule, i.e., schedules that are optimal for the measure, can be obtained. Ranking-fair schedules exhibit several symmetries, which we exploit in integer programming formulations for finding such schedules, as well as in combinatorial arguments to

prove or disprove existence for specific cases. In particular, we show when the number of players is a multiple of four, there exists a ranking-fair schedule that also has a minimum number of breaks.

Detecting and Handling Reflection Symmetries in MINLP – CHRISTOPHER HOJNY, TU EINDHOVEN –

A symmetry of a mixed-integer nonlinear program (MINLP) is a bijective map that transforms feasible solutions into feasible solutions while preserving the objective value. Already for linear problems it is well known that symmetries can negatively impact the performance of branch-and-bound methods, since symmetric regions of the search space will be repeatedly explored without providing new information. Most of the existing approaches for detecting symmetries focus on permutation symmetries that exchange the order of entries in a solution vector. For many problems classes, however, permutation symmetries do not capture all symmetries. For example, when packing objects into rectangular containers, also reflection symmetries of the container can be taken into account.

In this talk, we present a novel mechanism for computing, next to permutation symmetries, also reflection and translation symmetries of MINLPs. As existing approaches for permutation symmetries, we introduce a suitable graph whose automorphisms correspond to symmetries of a MINLP. Since reflection symmetries may nontrivially interact with nonlinear functions of a MINLP, however, our graph is much more involved than the previously known symmetry detection graphs. We also briefly discuss generalizations of state-of-the-art symmetry handling methods to reflection symmetries. The talk is concluded by numerical experiments showing the effect of handling reflection symmetries in the MINLP solver SCIP.

2a - FB151 - 513D - **Recent Developments in MIP Theory and Applications** Chair: BURAK KOCUK

Proximity results in conic mixed-integer programming – BURAK KOCUK, SABANCI UNIVERSITY – Diego Alejandro Moran Ramirez

We study the distance between the optimal solutions (resp. optimal values) of mixed-integer conic programs and the optimal solutions (resp. optimal values) of their continuous relaxations. We show that these values can be upper bounded in terms of the recession cone of the feasible region of the continuous relaxation when the recession cone is regular. We then specialize our analysis to mixed-integer second order cone programs (SOCPs). In the case of SOCPs whose feasible regions are defined by a single Lorentz cone constraint, we show that the distance can be upper bounded in terms of the data of the problem (the objective function vector, the matrix defining the SOCP, and the right-hand side). We also study when this bound can be made independent of the right-hand side, akin to the mixed-integer programming case. Finally, in the case the feasible region of the SOCP is defined by two or more Lorentz cone constraints, then we show that, in general, the distance cannot be bounded independently of the right-hand side defining the SOCP, not even when its feasible region is compact or one of the Lorentz cones constraints defines a polyhedral set.

On formulations of the close-enough TSP

– GUSTAVO ANGULO, PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE –

We consider the traveling salesman problem (TSP) with neighborhoods, a generalization of the TSP where we have the flexibility of visiting a neighborhood of each node instead of visiting the node itself. To solve this problem, two formulations based on second-order conic programming are proposed along decomposition schemes. In addition, linear approximations arbitrarily close to the initial formulations are proposed. Finally, computational experiments are performed to test the effectiveness of the proposed formulations.

A Convexification-Based Outer-Approximation Method for Convex and Nonconvex MINLP

– DAVID BERNAL NEIRA, PURDUE UNIVERSITY – Zedong Peng, Kevin Furman, Can Li, Ignacio Grossmann

The advancement of domain reduction techniques, such as bound tightening, convexification, and redundant variables elimination, has significantly enhanced the performance of Branch-and-Bound-based solvers in mathematical programming. However, decomposition-based MINLP solvers have not yet fully harnessed the potential of domain reduction techniques. This work delves into the impact of integrating convexification and domain reduction techniques within the Outer-Approximation method. We propose a refined convexification-based Outer-Approximation method alongside a Branch-and-Bound method for both convex and nonconvex Mixed-Integer Nonlinear Programming problems. In both methods, these domain reduction techniques can be employed both in the pre-solve phase and during solving integer-fixed NLP subproblems. However, this work focuses on the impact of domain reduction methods at the method's initialization stage, including feasibility-based and optimality-based bound tightening, the auxiliary variable method, and other convex reformulations. The proposed methods have been developed and incorporated into the open-source Mixed-Integer Nonlinear Decomposition Toolbox for Pyomo-MindtPy. Comprehensive benchmark tests were conducted, validating the effectiveness and reliability of our proposed algorithms. These tests highlight the improvements achieved by incorporating convexification and domain reduction techniques into the Outer-Approximation and Branch-and-Bound methods for MINLP problems.

M-Decomposable Sets in Integer Programming – DIEGO ALEJANDRO MORAN RAMIREZ, RENSSELAER POLYTECHNIC INSTITUTE –

An M-decomposable set is a closed convex set which is the sum of a compact convex set and a closed convex cone. We present properties of these sets in the context of integer programming. In particular, we present properties of their integer hull, some cutting plane closures and subadditive duality.

2a - FB919 - 513B - **Mixed Integer Linear Programming 5** Chair: JAKOB NORDSTRÖM

Generalized Cuts and Grothendieck Covers: Algorithmic Performance – NATHAN BENEDETTO PROENÇA, UNIVERSITY OF WATERLOO – Marcel K. De Carli Silva, Cristiane M. Sato, Levent Tuncel

In recent work, we have exploited the Antiblocking Duality (as introduced by Fulkerson in the 1970s) between the maximum

cut and (fractional) cut-covering problems to develop approximation algorithms. Even more recently, we strengthened and extended this theory as well as the underlying algorithms to a wide class of generalized cuts and generalized fractional covers. Problems like Boolean 2-Constraint Satisfaction are included in this wide class. These algorithms all obtain fractional covers by repeated sampling from a distribution obtained via semidefinite programming (SDP), following the framework of the seminal work of Goemans and Williamson. We now explore the practical performance of this family of algorithms, experimentally studying the behavior of the approximation ratio with respect to the samples taken and the quality of the SDP solution, as well as other aspects of theoretical and practical interest.

Combinatorial solving with provably correct results – JAKOB NORDSTRÖM, UNIVERSITY COPENHAGEN –

Since the turn of the millennium there have been dramatic improvements in algorithms for combinatorial solving and optimization. The flipside of this is that as solvers get more and more sophisticated, it becomes harder to avoid bugs sneaking in during algorithm design and implementation, and it is well documented that even the most mature tools sometimes return incorrect results. Software testing, while important, has not been sufficient to resolve this problem, and formal verification methods are far from being able to scale to the level of complexity in modern combinatorial solvers. During the last ten years the Boolean satisfiability (SAT) community has instead successfully introduced proof logging, meaning that algorithms have to output, along the answer to a problem, a machine-verifiable proof that this answer is correct, but this success has not been replicated in stronger combinatorial optimization paradigms.

In the last few years, this has started to change with the introduction of so-called pseudo-Boolean proof logging using 0-1 integer linear constraints. In this talk, we will illustrate how this approach seems to hit a sweet spot between simplicity and ease of verification on the one hand and expressive power on the other. We will try to provide a glimpse how pseudo-Boolean reasoning has enabled proof logging not only for previously unsupported advanced SAT techniques (including symmetry breaking) but also for SAT-based optimization (MaxSAT solving), subgraph solving, constraint programming, and most recently presolving techniques in 0-1 integer linear programming (ILP). Perhaps the most striking feature of this method is that proofs can be written in a simple proof format referred to as "cutting planes with strengthening rules" that knows nothing about symmetries, groups, graphs, integer-valued variables, or global constraints.

2b - MB12 - 514C - **Binary Quadratic Programming** Chair: SVEN MALLACH

Matrix generation for binary quadratic programming – LUCAS LÉTOCART, LIPN, UNIVERSITY SORBONNE PARIS NORD, CNRS – Enrico Bettiol, Immanuel Bomze, Francesco Rinaldi, Emiliano Traversi

We consider binary quadratically constrained quadratic problems and propose a new approach based on decomposition to generate strong bounds. This new relaxation is based on the

Boolean Quadric Polytope and is solved via a Dantzig-Wolfe Reformulation in matrix space. For block-decomposable problems, we extend the relaxation and analyze the theoretical properties of this new decomposition. If overlapping size of blocks is at most two, we establish equivalence to the one based on the Boolean Quadric Polytope. We prove that this equivalence does not hold if the sparsity graph is not chordal and we conjecture that equivalence holds for any block structure with a chordal sparsity graph. Numerical results show that the proposed approach yields very good bounds in reasonable time.

Exact and Approximate Solution Methods for Convex Binary Quadratic Optimization – BORZOU ROSTAMI, UNIVERSITY OF ALBERTA – Guy Desaulniers, Fausto Errico, Andrea Lodi

This paper provides exact and approximate solution methods for convex binary quadratic programs with linear constraints (CBQP). We reformulate the problem using spectral decomposition and apply principal component analysis (PCA) to derive approximate solutions. The quality of the resulting solutions is assessed by determining worst-case optimality gaps. Furthermore, we show how to extend this approximate method to an exact strategy by reformulating it into a mixed-integer quadratic program and developing a Benders-type decomposition method to derive optimal solutions for the original CBQP. We apply this methodology to routing optimization with stochastic and correlated travel times, demonstrating its effectiveness across problem instances with different structures.

Binary Quadratic Optimization: A Polyhedral Characterization of Linearizable Instances – LUCAS WADDELL, BUCKNELL UNIVERSITY –

We introduce a new polyhedral framework for characterizing instances of binary quadratic programs that are linearizable, which means that the quadratic objective can be equivalently rewritten as a linear function in such a manner that preserves the objective function value at all feasible binary solutions. Of particular interest to our study are problems that are challenging when the objective function is quadratic but are readily solvable in the linear case, such as the quadratic assignment problem (QAP) and the quadratic minimum spanning tree problem (QMSTP). Our main result is that an instance of such a problem is linearizable if and only if the quadratic cost coefficients can be used to create a linear equation in a lifted variable space that is valid for the affine hull of a specially constructed discrete set. In this talk, we will discuss how viewing the concept of linearizability through this new polyhedral lens has interesting implications for both the QAP and QMSTP, leading to the resolution of open questions in each case. Furthermore, we will show how our framework for studying the linearizable property can provide insights into the identification and characterization of other families of problem instances that, while not linearizable, can still be solved in polynomial time.

Inductive Linearization for Binary Quadratic Programs – SVEN MALLACH, UNIVERSITY OF BONN –

Given binary products in the objective function or constraints, the inductive linearization technique may serve as a computationally attractive compromise between the (often weak) "standard" linearization and the (often large) result of applying

the reformulation linearization technique (RLT). In the best case, inductive linearizations may be more constraint-side compact than the standard linearization, and still provide a continuous relaxation that is at least as tight or even close to the one of a complete first-level RLT. Prominent combinatorial optimization problems where this applies are for instance the Quadratic Assignment, the Quadratic Matching, and the Quadratic Traveling Salesman Problem. We shed light on some properties of a program that appear particularly suited to experience computational advantage when applying the inductive linearization technique in combination with a MIP solver. doi: 10.1007/s10288-023-00537-5

2b - MC286 - 513F - MINLP Applications Chair:
AKSHAY GUPTA

The lattice of dijoins – OLHA SILINA, CARNEGIE MELLON UNIVERSITY –

In a digraph, a dicut is a cut with all arcs crossing in one direction. A dijoin is a subset of arcs that intersects each dicut at least once. It was known that the minimum size of a dijoin equals the maximum number of disjoint dicuts, however, the reverse statement i.e. whether the minimum size of a dicut equals the maximum number of disjoint dijoins, remains open (Woodall's conjecture). This problem can be viewed as following: given a digraph with minimum dicut of size T , can its arcs be split into T dijoins? We study a relaxation of this problem, where we look to express the arcs as a linear combination of dijoins. In the spirit of Lovász's theory of the matching lattices, we prove that the lattice of tight dijoins is equal to the set of integer points in its linear hull, which, in particular, implies the relaxation of Woodall's conjecture mentioned above. This is a joint work with Ahmad Abdi, Gérard Cornuéjols, and Siyue Liu.

Mixed integer linear programming formulations and column generation algorithms for the Minimum Normalized Cuts problem – JUSTO PUERTO, UNIVERSITY OF SEVILLE – Ponce Diego, Francisco Temprano

This paper deals with the k -way normalized cut problem in complex networks. It presents a methodology that uses mathematical optimization to provide mixed-integer linear programming formulations for the problem. The paper also develops a branch-and-price algorithm for the above-mentioned problem which scales better than the compact formulations. Additionally, a heuristic algorithm which is able to approximate large-scale image problems in those cases where the exact methods are not applicable is presented. Extensive computational experiments assess the usefulness of these methods to solve the k -way normalized cut problem. Finally, we have applied the minimum normalized cut objective function to the segmentation of actual images, showing the applicability of the introduced methodology.

Scaling-up exact methods for solving the discrete bilevel network design problem under user equilibrium – DAVID REY, SKEMA BUSINESS SCHOOL – Michael Levin

We address the bilevel discrete network design problem (DNBP) in transportation. Given a transportation network and a budget, the DNBP consists of determining the optimal subset of links to be added to minimize congestion effects.

Congestion is measured using traffic assignment where link travel times are modeled as convex flow-dependent functions and network users are represented as utility-maximizing agents making selfish route choice decisions. The collective choice of users can be formulated as a Nash equilibrium problem known as a Wardropian equilibrium. The DNDP admits a natural bilevel optimization where the leader represents the network manager and the follower is a parameterized traffic assignment problem (TAP). The TAP is a convex optimization problem whose optimality conditions require that, for each commodity, all paths used at user equilibrium (UE) have minimum and equal travel time. Embedding UE conditions in a bilevel optimization formulation via path-based variables requires handling path set enumeration. While column generation techniques for (integer-) linear programs are well-developed, this is not the case for nonlinear programs. This has led researchers to develop link-based formulations amenable to mixed-integer nonlinear programming approaches. However, computational studies have found that even after approximating link travel time functions via piecewise linear functions, exact algorithms struggle to solve the DNDP to global optimality on medium-size networks with more than 20 candidate links. This study introduces a novel path-based approach to solve the DNDP based on Lagrangian relaxation that overcomes path enumeration challenges by leveraging the scalability of TAP algorithms. We propose to solve the DNDP by considering its high point relaxation and reformulating it using Lagrangian relaxation. Relaxing and penalizing the leader-follower linking constraints leads to a separable, system optimum (SO) formulation of the DNDP that is solved by iteratively combining the solutions of parameterized SO-TAPs and binary knapsack problems. We develop customized Lagrangian multiplier update rules and use this Lagrangian relaxation to provide lower bounds. Upper bounds are found by solving UE-TAPs search. We embed this bounding procedure in exact algorithms for discrete-continuous bilevel optimization problems such as branch and bound, and cut generation methods to solve the DNDP to optimality. We discuss algorithmic considerations and present numerical results on benchmark problem instances for the DNDP.

2b - TA152 - 513F - Solving ACOFF problems
Chair: DANIEL BIENSTOCK

AC-Unit Commitment: Capturing Losses – HASSAN HIJAZI, GUROBI –

This talk will present some recent work on solving the nonlinear AC-Unit Commitment problem as defined in the ARPA-E Grid Optimization Challenge 3 (<https://gocompetition.energy.gov/challenges/challenge-3>).

Computational results will be shown on the competition datasets. We will also discuss plans on improving Gurobi's capabilities to handle these types of problems.

New Advances in Solving Large-Scale Unit-Commitment-AC-OPF Problem – ANDY SUN, MASSACHUSETTS INSTITUTE OF TECHNOLOGY –

In this talk, we will present some recent advances in algorithms and computation techniques for solving large-scale grid optimization problems involving unit commitment and ACOFF. These problems are at the core of power grid operations and planning. We develop new methods for temporal and spatial decomposition, convex relaxations, and parallel computation. The new algorithmic framework can solving UC-AC-OPF prob-

lems on power grids with up to 23,000 buses, 12 time periods, within 10 min. The solutions obtained are highly feasible (within 10^{-8} tolerance) and on average have less than 1% global optimality gap. This progress won top performance in the ARPA-E Grid Optimization Challenge 3 and has significant implications for practical applications in power grid operations.

Fusing AI and Optimization for Security-Constrained Economic Dispatch – PASCAL VAN HENTENRYCK, GEORGIA TECH –

This talk shows how the fusion of AI and Optimization can lead to orders of magnitude improvements for Security-Constrained Economic Dispatch optimizations. It presents an end-to-end self-supervised learning framework that is guaranteed to find high-quality solutions in milliseconds, significantly outperforming the state of the art solutions. The key technical contributions is the use of differential layers and primal-dual learning to restore feasibility and handle contingencies.

Accurate and Warm-Startable Linear Cutting-Plane Relaxations for ACOFF – MATIAS VILLAGRA, COLUMBIA UNIVERSITY – Daniel Bienstock

We present a linear cutting-plane relaxation approach that rapidly proves tight lower bounds for the Alternating Current Optimal Power Flow Problem (ACOPF). Our method leverages outer-envelope linear cuts for well-known second-order cone relaxations for ACOFF along with modern cut management techniques. These techniques prove effective on a broad family of ACOFF instances, including the largest ones publicly available, quickly and robustly yielding sharp bounds. Our primary focus concerns the (frequent) case where an ACOFF instance is considered following a small or moderate change in problem data, e.g., load changes and generator or branch shut-offs. We provide significant computational evidence that the cuts computed on the prior instance provide an effective warm-start for our algorithm.

2b - TB195 - 513F - Optimization over integers, cones and polynomials Chair: LUZE XU

Spurious local minima in nonconvex sum-of-square optimization – SHIXUAN ZHANG, TEXAS A&M UNIVERSITY – Grigoriy Blekherman, Rainer Sinn, Mauricio Velasco

We study spurious local minima in a nonconvex low-rank formulation of sum-of-squares optimization on a real variety X . We reformulate the problem of finding a spurious local minimum in terms of syzygies of the underlying linear series, and also bring in topological tools to study this problem. When the variety X is of minimal degree, there exist spurious local minima if and only if both the dimension and the codimension of the variety are greater than one, answering a question by Legat, Yuan, and Parrilo. Moreover, for surfaces of minimal degree, we provide sufficient conditions to exclude points from being spurious local minima. In particular, all spurious local minima on the Veronese surface, corresponding to ternary quartics, lie on the boundary and can be written as a binary quartic, up to a linear change of coordinates, complementing work by Scheiderer on decompositions of ternary quartics as a sum of three squares. For general varieties of higher degree, we give examples and characterizations of spurious local minima

in the interior, which can potentially lead to efficient low-rank sum-of-squares algorithms.

Solving Nonconvex Optimization Problems using Outer Approximations of the Set-Copositive Cone – KURT ANSTREICHER, UNIVERSITY OF IOWA – Markus Gabl

We consider the solution of nonconvex quadratic optimization problems using an outer approximation of the set-copositive cone that is iteratively strengthened with conic constraints and cutting planes. Our methodology utilizes an MILP-based oracle for a generalization of the copositive cone that considers additional linear equality constraints. In numerical testing we evaluate our algorithm on a variety of different nonconvex quadratic problems.

Integer Points in Arbitrary Convex Cones: The Case of the PSD and SOC Cones – LUZE XU, UNIVERSITY OF CALIFORNIA, DAVIS – Jesus De Loera, Brittney Marsters, Shixuan Zhang

We investigate the semigroup of integer points inside a convex cone. We extend classical results in integer linear programming to conic integer programming. We show that the semigroup associated with nonpolyhedral cones can sometimes have a notion of finite generating set. We show this is true for the cone of positive semidefinite matrices (PSD) and the second order cone (SOC). Both cones have a finite generating set of integer points, similar in spirit to Hilbert bases, but require the action of a finitely generated group. We also extend notions of total dual integrality, Gomory-Chvátal closure, and Carathéodory rank to integer points in arbitrary cones.

On Polynomial-Time Solvability of Combinatorial Markov Random Fields – SHAONING HAN, NATIONAL UNIVERSITY OF SINGAPORE – Andres Gomez, Jong Shi Pang

The problem of inferring Markov random fields (MRFs) with a sparsity or robustness prior can be naturally modeled as a mixed-integer program. This motivates us to study solution methods for a general class of convex submodular optimization problems with indicator variables, which we show to be polynomially solvable in this paper. The key insight is that, possibly after a suitable reformulation, indicator constraints preserve submodularity. Fast computations of the associated Lovasz extensions are also discussed under certain smoothness conditions, and can be implemented using only linear-algebraic operations in the case of quadratic objectives.

2b - TC920 - 513F - Mixed Integer Nonlinear Programming 1 Chair: ANDREA GHEZZI 

Periodic Capacity Development in Nonlinear Closed-loop Logistics Planning: an advanced Continuous Genetic Algorithm – EHSAN YADEGARI, POSTDOC FELLOW – Thierry Garaix

Logistics Network Design includes network configuration decisions having long-standing influences on other tactical and operational decisions. Recently, regarding environmental issues and customer awareness, Closed-loop logistics network (CLLN) design is taken into consideration. This paper presents a nonlinear mixed integer programming for a dynamic multi-period multi-products forward/reverse logistics network with seven layers: suppliers, Plants, distribution centers (DCs), Customer

zones, and inspection and disposal centers. The possibility of opening and closing of each facility at each period is one of the primary characteristics of the proposed model. Moreover, operational decisions such as inventory cost and capacity development in each period as well as the optimum assignment flow between consecutive layers of the proposed network are considered in CLLN modeling. To cope with such an NP-hard problem an advanced continuous Genetic Algorithm (cGA) has been applied. The primary focus of cGA lies in its innovative representation method, which allows it to avoid the time-consuming repair mechanisms due to fixing the infeasible solutions. Furthermore, through the implementation of this representation, additional operators can be seamlessly integrated into the algorithm. Numerical results show the efficiency of the proposed method in large scale problems.

A First-Relax-then-Reformulate Approach for Bilevel Problems with Nonconvex Mixed-Integer Nonlinear Lower Levels – HENRI LEFEBVRE, TRIER UNIVERSITY – Martin Schmidt

In this talk, we consider bilevel optimization problems in which the follower solves a nonconvex mixed-integer nonlinear problem and all leader decisions are binary. For this challenging class of problems, we introduce a first-relax-then-reformulate approach to compute feasible points of good quality. Our algorithm is based on approximating the value function of the follower's problem by using a tailored penalty approach. In the most general setting, we show that this approximation can be made arbitrarily close to the original value function. We further show that, for mixed-integer linear problems, our approximation is exact, i.e., it coincides with the original value function. For both cases, we then derive a new column-and-constraint generation algorithm, which makes use of the novel approximation to guide the search towards feasible or even optimal points. Finally, we report computational results on instances taken from the literature that show the applicability of the approach.

A Sequential Benders-based Mixed-Integer Quadratic Programming Algorithm – ANDREA GHEZZI, UNIVERSITY OF FREIBURG – Wim Van Roy, Sebastian Sager, Moritz Diehl

In this talk we present a sequential mixed-integer quadratic programming (MIQP) algorithm for solving mixed-integer nonlinear problems (MINLPs). The algorithm employs a three-step method in each iterate: First, the algorithm linearizes the MINLP at a given iterate. Second, an MIQP with its feasible set restricted to a specific region around the current linearization point is formulated and solved. Third, the integer variables obtained from the MIQP solution are fixed, and only an NLP in the continuous variables is solved. The outcome of the third step is compared to previous iterates, and the best iterate so far is used as a linearization point in the next iterate. The polyhedral region in the second step is computed based on objectives and derivatives from all previous iterates, and build on concepts from generalized Benders' decomposition. Although the presented MINLP algorithm is a heuristic method without any global optimality guarantee, it converges to the exact integer solution when applied to convex MINLP with a linear outer structure. The conducted numerical experiments on existing benchmarks demonstrate that the proposed algorithm is competitive with other open-source solvers for MINLP.

2b - WA8 - 513F - **Mixed Integer Quadratic/Conic Optimization** Chair: ATAM-TURK ALPER

An Outer Approximation Method for Solving Mixed-Integer Convex Quadratic Programs with Indicators – LINCHUAN WEO, NORTHWESTERN UNIVERSITY – Simge Kucukyavuz

Mixed-integer convex quadratic programs with indicator variables (MIQP) encompass a wide range of applications, from statistical learning to energy, finance, and logistics. The outer approximation (OA) algorithm has been proven efficient in solving MIQP, and the key to the success of an OA algorithm is the strength of the cutting planes employed. In this paper, we propose a new technique for deriving cutting planes for MIQP from various convex relaxations, and, as a result, we develop new OA algorithms for solving MIQP at scale. The contributions of our work are two-fold: (1) we bridge the work on the convexification of MIQP and the algorithm design to solve large-scale problems, and (2) we demonstrate through a computational study on the sparse portfolio selection problem that our algorithms give rise to significant speedups compared with the state-of-the-art methods in the literature.

New Perspectives on Deriving Compact Extended Formulations for Optimization Problems with Indicator Variables – FATMA KILINC KARZAN, CARNEGIE MELLON UNIVERSITY – Soroosh Shafieezadeh Abadeh

Applications in statistics and data sciences require modeling an inherit sparsity as well as structural relationships among variables and often lead to NP-Hard nonconvex problems. Sparsity and structural relations in such problems are usually modeled by introducing indicator (binary) variables associated with the original continuous variables and enforcing complementarity relations between them. We consider optimization problems with convex objective functions of the continuous variables and arbitrary constraints on their associated indicators, and study the resulting conic mixed-binary sets with complementarity restrictions. For such sets, by studying the recessive directions and also utilizing perspective functions, we provide a practical framework to derive compact, ideal, and conic-representable extended formulations. The size of our extended formulation depends on the "rank" of the function and the presence of sign restrictions. Moreover, our techniques highlight that the complexity of the convex hull characterizations of these conic mixed-binary sets with complementarity restrictions hinges solely on the complexity of the convex hull characterization of a set defined by only the indicator variables. This then enables us to take advantage of the extensive research on convex hull descriptions of binary sets and related sophisticated optimization software. In addition, our results unify and generalize previous results established for special cases, e.g., perspective reformulation for separable functions, non-separable rank-one functions, or low-rank quadratic functions optimized over domains with combinatorial constraints on indicator variables and possibly sign constraints on continuous variables. On the computational side, we illustrate the effectiveness of our approach on sparse structured logistic regression problems.

Convexification and solution of mixed-integer quadratic programs using decision diagrams –

ANDRES GOMEZ, UNIVERSITY OF SOUTHERN CALIFORNIA –

Decision diagrams have emerged as a new technique to solve mixed-integer problems, based on using dynamic programming. Moreover, decision diagrams can also be used to convexify the feasible regions of the problems under consideration, and thus the decision diagram techniques can be combined with standard branch-and-bound methods. However, most of results concerning decision diagrams pertain to pure integer sets, as the convexifications obtained, based on the path polytope, are inherently polyhedral. In this talk we discuss how to use decision diagrams to solve and convexify mixed-integer quadratic optimization problems, a class of problems with non-polyhedral ideal formulations.

Sensitivity Analysis for Mixed Binary Quadratic Programs – SANTANU DEY, GEORGIA INSTITUTE OF TECHNOLOGY – Cifuentes Diego, Jingye Xu

We consider sensitivity analysis for Mixed Binary Quadratic Programs (MBQPs) with respect to changing right-hand-sides (rhs). We show that even if the optimal solution of a given MBQP is known, it is NP-hard to approximate the change in objective function value with respect to changes in rhs. Next, we study algorithmic approaches to obtaining dual bounds for MBQP with changing rhs. We leverage Burer's completely-positive (CPP) reformulation of MBQPs. Its dual is an instance of co-positive programming (COP), and can be used to obtain sensitivity bounds. We prove that strong duality between the CPP and COP problems holds if the feasible region is bounded or if the objective function is convex, while the duality gap can be strictly positive if neither condition is met. We also show that the COP dual has multiple optimal solutions, and the choice of the dual solution affects the quality of the bounds with rhs changes. We finally provide a method for finding good nearly optimal dual solutions, and we present preliminary computational results on sensitivity analysis for MBQPs.

2b - WB7 - 513F - **Algorithms for Global Mixed Integer Nonlinear Optimization** Chair: PIETRO BELOTTI

Solving MINLPs to Global optimality with the FICO Xpress Solver – IMRE POLIK, FICO XPRESS –

This talk introduces the global optimization engine in FICO Xpress, allowing to solve general mixed-integer nonlinear problems to global optimality. We will discuss how existing features of the MILP and local NLP solvers are used and explain several different extensions for, e.g., spatial branching and convexification cuts, and their performance implications. Recent improvements and their impacts on the performance of the solver are also discussed.

RAPOSa: A free global solver for mixed integer polynomial optimization problems – JULIO GONZÁLEZ DÍAZ, UNIVERSITY OF SANTIAGO DE COMPOSTELA – Brais González Rodríguez, Ignacio Gómez Casares, Iria Rodríguez Acevedo

This talk aims to showcase the latest developments in RAPOSa, a global solver for polynomial optimization problems, with a primary focus on its extension to address mixed-integer

problems. The talk will delve into various facets of the key challenges encountered in this extension, along with some numerical experiments designed to assess the performance of different approaches, both relative to one another and also relative to other state-of-the-art global solvers for MINLP problems.

Regarding the extension to mixed integer problems, we will discuss several approaches, targeted to improve the quality of the final bounds returned by the algorithm. In particular, for the upper bounds we will discuss the integration of the calls to auxiliary local solvers (discrete and continuous) and the trade-offs that arise when trying to do it efficiently.

Finally, we will also briefly touch on other recent developments in RAPOSa, which will primarily revolve around the use of (nonlinear) conic constraints, such as SOCP and SDP constraints, inside RAPOSa's branch-and-bound scheme. In particular, we will discuss the effect of using these conic constraints to tighten, at different places of the algorithm, the linear RLT relaxations on which RAPOSa is built upon.

Techniques and advances for solving MINLPs with Gurobi – ROBERT LUCE, GUROBI –

We consider solving mixed-integer nonlinear optimization problems to global optimality. Our solver is based on the common branch-and-bound paradigm, but includes a number of specialized techniques to deal with nonconvex constraints and nonconvex objective functions. In this talk we will outline a few of these components from a theoretical and computational point of view, with some emphasis on relaxation schemes and cutting planes. Further we will preview some features of the next Gurobi version that enhance MINLP solving and modeling capabilities.

Spatial branching for convex MIQO problems – PIETRO BELOTTI, POLITECNICO DI MILANO –

Branch-and-bound algorithms are the standard in mixed integer optimization, both with convex and non-convex continuous relaxations. Spatial branching, i.e., branching on continuous variables, is a necessary tool of global nonlinear solvers for tackling non-convexities arising in the objective function and/or the constraints. Spatial branching is harder than branching on integer variables and it has been comparably less studied. Similar to other techniques that were originally intended for MINLP but are now common in the MIP world, we show that it can be successfully used on a special class of convex mixed integer quadratic optimization problems with application in classification, where it is far more effective than the standard integer branching machinery of a commercial MIP solver.

2b - WC346 - 513F - MINLP meets Machine Learning
Chair: EMILIANO TRAVERSI

Outlier Detection in Regression: Conic Quadratic Formulations – JOSÉ NETO, TÉLÉCOM SUDPARIS, INSTITUT POLYTECHNIQUE DE PARIS – ANDRÉS GÓMEZ

In many applications, when building regression models, it is important to account for the presence of outliers, i.e., corrupted input data points resulting, e.g., from recording errors. Problems consisting in simultaneously performing linear regression and discarding outliers may be formulated as mixed-integer optimization problems which involve cubic terms, each given by the product of a binary variable and a quadratic term of the

continuous variables. This can be notably illustrated with the least trimmed squares problem, which consists in minimizing some given number of squared residuals, and is known to be both NP-hard and hard to approximate. Existing approaches in the literature typically rely on the linearization of the cubic terms using big-M constraints. However the latter lead to weak relaxations and poor performance in practice. In this work, we derive stronger second-order conic relaxations not involving big-M constraints. Our computational experiments indicate that the proposed formulations are several orders of magnitude faster than existing big-M formulations in the literature for this problem.

Optimization over trained graph neural networks: mixed-integer formulations, symmetry breaking, and optimal molecular design – SHIQIANG ZHANG, IMPERIAL COLLEGE LONDON – JUAN S. CAMPOS, CHRISTIAN FELDMANN, DAVID WALZ, FREDERIK SANDFORT, MIRIAM MATHEA, CALVIN TSAY, RUTH MISENER
Optimization over trained machine learning models has applications including verification, minimizing neural acquisition functions, and integrating a trained surrogate into a larger decision-making problem. We formulate and solve optimization problems constrained by trained graph neural networks (GNNs). For the classical GNN architectures considered in this talk, optimizing over a GNN with a fixed graph is equivalent to optimizing over a dense neural network. Thus, we study the case where the input graph is not fixed, implying that each edge is a decision variable, and develop two mixed-integer optimization formulations. To circumvent the symmetry issue caused by graph isomorphism, we propose two types of symmetry-breaking constraints, and provide a proof that adding these constraints will not remove all symmetric solutions. As a byproduct, a graph indexing algorithm is constructed to yield an indexing satisfying these constraints. To test our symmetry-breaking strategies and optimization formulations, we consider an application in optimal molecular design.

ML-based Clique Merging Algorithms to Solve Semidefinite Relaxations of Optimal Power Flow Problems – EMILIANO TRAVERSI, UNIVERSITÉ DE MONTPELLIER – ALIZADEH CHARLY, MIGUEL ANJOS, LUCAS LÉTOCAR

Semidefinite Programming (SDP) is a powerful technique to compute tight lower bounds for Optimal Power Flow (OPF) problems. Even using clique decomposition techniques, semidefinite relaxations are still computationally demanding. However, there are many different clique decompositions for the same SDP problem and they are not equivalent in terms of computation time. In this paper, we propose a new strategy to compute efficient clique decompositions with a clique merging heuristic. This heuristic is based on two different estimates of the computational burden of an SDP problem: the size of the problem and an estimation of a per-iteration cost for a state-of-the-art interior-point algorithm. We compare our strategy with other algorithms on MATPOWER instances and we show a significant decrease in solver time. In the last part of the talk we present recent developments on how to incorporate machine learning techniques to automatically identify effective clique decompositions.

2b - ThA77 - 513F - **Polynomial Optimization Algorithms and Applications** Chair: BISSAN GHADDAR

Parameter estimation in ODEs: assessing the potential of local and global solvers – ANGEL MANUEL GONZÁLEZ RUEDA, UNIVERSITY OF SANTIAGO DE COMPOSTELA – Manuel Fernández De Dios, Julio R. Banga, Julio González Díaz, David R. Penas

In this work, we consider the problem of parameter estimation in models described by nonlinear ordinary differential equations (ODEs). In particular, a direct transcription approach is considered to solve the problem, in such a way that the controls and state profiles are discretized. Furthermore, the original ODE system is transformed into algebraic equations for the optimization model after applying a finite difference scheme. This leads to the formulation of a standard nonlinear programming (NLP) problem that can be solved with mathematical programming techniques. The main goal of this work is to analyze the potential of mathematical modeling and state-of-the-art solvers to handle these problems. An extensive numerical study over a set of eight challenging parameter estimation problems of nonlinear dynamic biological systems is carried out. In this study different modeling aspects such as the chosen discretization scheme or different variants of the formulations are analyzed. Besides, thanks to the use of global and local state-of-the-art solvers, our approach provides insight into the challenges posed by local optimality and identifiability. The results obtained with this approach are very promising, outperforming the capacity to solve these problems that had been acknowledged previously in the literature.

Domain reduction techniques in polynomial optimization – IGNACIO GÓMEZ CASARES, UNIVERSITY OF SANTIAGO DE COMPOSTELA – Brais González Rodríguez, Julio González Díaz, Pablo Rodríguez Fernández

The design and implementation of global optimization algorithms for general MINLP problems is a very active field of research, and the number of available solvers has been steadily increasing over the past years. Recently, the companies behind state-of-the-art MILP solvers such as Xpress and Gurobi have announced the release of new versions capable of solving general MINLP to certified global optimality. Convexifications and domain reduction techniques are probably the two most important elements behind the efficiency of the spatial branching required to handle the nonconvexities. The focus of this talk is precisely on the joint assessment of the individual impact of different aspects of domain reduction on the performance of a global optimization algorithm. In particular, we will take a look at the RLT technique (designed to solve polynomial optimization problems) and several enhancements that can be added to it, such as improvements to the selection of the branching variable and the branching point, and bound-tightening techniques such as FBBT and OBBT. The aim is to check which of these techniques can have a noticeable impact on the performance of this particular technique (and, because we are working only on improving the spatial branching performance, on the performance of general non-linear solvers) and therefore guide the developers of this solvers towards implementing the most promising enhancement.

Learning to Accelerate the Global Optimization of Quadratically-Constrained Quadratic Programs – ROHIT KANNAN, VIRGINIA TECH – Harsha Nagarajan, Deepjyoti Deka

We use machine learning to predict optimal instance-specific heuristics and accelerate the guaranteed global minimization of families of nonconvex quadratically-constrained quadratic programs (QCQPs). Specifically, we consider partitioning-based convex mixed-integer programming (MIP) relaxations for nonconvex QCQPs and propose the novel concept of strong partitioning (SP) to optimally partition variable domains without sacrificing global optimality. We design a local optimization method for solving this challenging max-min SP problem and learn a boosted regression tree approximation of SP for homogeneous families of QCQPs. Furthermore, we propose an end-to-end neural network model trained to directly predict partitioning points for a QCQP instance, thereby eliminating the need for computationally expensive SP training data.

We present a detailed computational study on randomly generated families of QCQPs, including instances of the pooling problem, using the open-source global solver Alpine. Our numerical experiments demonstrate significant reductions in solution time achieved by strong partitioning, its boosted regression tree approximation, and the end-to-end neural network model. On average, these techniques reduce Alpine's solution time by factors ranging from 3.5 to 16.5, 2 to 4.5, and 6 to 9, respectively, across various QCQP families. Moreover, we observe maximum reduction factors of 15 to 700, 10 to 200, and 15 to 100 for the three techniques, respectively.

Strong and fast SDP bounds for water network problems – BISSAN GHADDAR, IVEY BUSINESS SCHOOL –

In this work we exploit two types of sparsity in polynomial optimization for the valve setting problem in water networks. The valve setting problem consists of many small subproblems involving few variables and monomials and connected with sparse linking constraints. The problem exhibits correlative sparsity (subsets of variables appearing separately in constraints) and term sparsity (few monomials present in the problem). We analyze the performance of several existing sparse SDP relaxations and suggest a new simple SDP relaxation that uses both correlative and term sparsity to reach a trade-off between the bound quality and running times. We report numerical results on four water networks ranging in size from 4 to about 2000 nodes. The approach extends to electricity and gas network problems.

2b - ThB34 - 513F - **Advances on Multi-Objective and Separable Mixed Integer Nonlinear Programming** Chair: MARIANNA DE SANTIS

Pareto LEAP: An Algorithm for Biobjective Mixed Integer Nonlinear Programming – PHILIP DE CASTRO, CLEMSON UNIVERSITY – Margaret Wiecek

Biobjective mixed integer programs (BOMIPs) have recently received much attention due to their broad applications. A

class of methods which use tabu or "no-good" constraints have been developed to solve biobjective mixed integer linear programs. In this work, we propose an extension of tabu methods to the case of nonlinear problems. We use the structure of the outcome space of nonlinear BOMIPs to "leap" between slices of the Pareto set, skipping over dominated slices. We discuss theoretical results and propose an algorithm, Pareto LEAP, which may be flexibly implemented in any language, using any suitable solver. We also demonstrate Pareto LEAP on an example problem from the literature.

Pseudo enclosures for multi-objective mixed-integer nonconvex optimization – MORITZ LINK, UNIVERSITÄT KONSTANZ – Gabriele Eichfelder, Stefan Volkwein, Leo Warnow

In this talk, we present a concrete approach for solving multi-objective mixed-integer nonconvex optimization problems. In fact, our algorithm uses information from the image space to build adaptively individualized piecewise linear relaxations of nonlinear terms appearing in the original problem. This procedure allows us to compute a so-called pseudo enclosure of the nondominated set of some predetermined accuracy relying on MILPs and NLPs. Within the successive refinement framework, the introduction of adaptivity and therefore individualization with respect to the current region of interest serves to avoid unnecessary complexity of the MILP relaxations. We validate the proposed scheme with numerical experiments.

Detecting the integer efficient assignments for multiobjective mixed-integer convex quadratic problems – DANIELE PATRIA, SAPIENZA, UNIVERSITY OF ROME – Marianna De Santis, Gabriele Eichfelder, Leo Warnow

An efficient integer assignment for multiobjective mixed-integer problems is a fixing of the integer variables so that there exists an efficient point with exactly that fixing. Depending on the instance, solution algorithms for multiobjective mixed-integer problems may need to detect a large number of efficient integer assignments, possibly all integer feasible assignments, so that a full enumeration cannot be avoided. This shows a big difference with respect to the single objective case and a big challenge in terms of algorithms development. In this talk, we present a branch-and-bound method for multiobjective mixed-integer convex quadratic programs that computes a superset of efficient integer assignments and a coverage of the nondominated set. The method relies on outer approximations of the upper image set of continuous relaxations. These outer approximations are obtained addressing the dual formulations of specific subproblems where the values of certain integer variables are fixed. The devised pruning conditions and a tailored preprocessing phase allow a fast enumeration of the nodes. Despite we do not require any boundedness of the feasible set, we are able to prove that the method stops after having explored a finite number of nodes. Numerical experiments on a broad set of instances with two, three, and four objectives are presented.

MWU 2.0 with approximation guarantee for non-convex separable problems – LUCA MENCARELLI, UNIVERSITÀ DI PISA –

In this talk, we introduce a novel approximation guaranteed algorithm for (feasibility) mixed integer nonlinear programming (MINLP) by straightforwardly extending the Plotkin,

Shmoys and Tardos algorithm for fractional packing and covering problems and the multiplicative weights update (MWU) framework. The approximation guaranteed algorithm aims to find a epsilon-feasible solution with respect to a feasibility region described by possibly non-convex nonlinear constraints, by solving an opportune surrogate relaxation of the given problem, calling an oracle for a polynomial number of times, and adapting the weights of the surrogate relaxation by means of the MWU algorithm. We implement our new method in order to heuristically solve (non-convex) separable problems (SPs), i.e., MINLP problems in which each occurring function is representable as the sum of non-convex univariate functions. We can assume without loss of generality that the objective function of SPs is linear. In particular, we combine plain bisection approach with the novel approximation guaranteed method in order to deal with optimality and feasibility issues at the same time. In this latter particular case, since the original problem is separable, the oracle is reasonably easy to implement, by computing the roots of the first derivative of the univariate functions and its value on the boundary and eventually rounding the computed values. We present preliminary computational results on random generated instances, showing the proposed method is competitive against the state-of-the-art exact solvers for non-convex MINLP, namely Couenne and Scip.

2b - ThC159 - 513F - Discrete structures in non-convex problems Chair: SONIA CAFIERI [↔](#)

On constrained mixed-integer DR-submodular minimization – KIM YU, UNIVERSITÉ DE MONTRÉAL – Simge Kucukyavuz

Submodular set functions play an important role in integer programming and combinatorial optimization. Increasingly, applications in resource allocation and machine learning have led to generalized submodular functions defined over general integer or continuous domains. In our study, we focus on the notion of Diminishing Returns (DR) submodularity and the problem of minimizing DR-submodular functions over mixed-integer feasible sets defined by box constraints and monotonicity constraints. We propose multiple classes of valid inequalities and further show that these inequalities, along with trivial constraints, fully describe the convex hull of the epigraph of DR-submodular functions under the aforementioned constraints. Our results hold in the original decision space, which avoids pseudo-polynomial expansions that are used to handle general integer variables in the existing literature.

Exploiting the polyhedral geometry of stochastic linear bilevel programming – GONZALO MUÑOZ, UNIVERSIDAD DE O'HIGGINS – David Salas, Anton Svensson

We study linear bilevel programming problems whose lower-level objective is given by a random cost vector with known distribution. We consider the case where this distribution is nonatomic, allowing to reformulate the problem of the leader using the Bayesian approach in the sense of Salas and Svensson (2023) with a vertex-supported belief. We prove that this formulation is piecewise affine over the so-called chamber complex of the feasible set of the high point relaxation. We propose two algorithmic approaches to solve general problems enjoying this last property. The first one is based on enumerating the vertices of the chamber complex. The second one is a Monte-Carlo

approximation scheme based on the fact that randomly drawn points of the domain lie, with probability 1, in the interior of full-dimensional chambers, where the problem (restricted to this chamber) can be reduced to a linear program. Finally, we evaluate these methods through computational experiments showing both approaches' advantages and challenges.

On computing upper bounds in nonlinear problems involving disjunctive constraints

– SONIA CAFIERI, ENAC, UNIVERSITÉ DE TOULOUSE – Marcel Mongeau

The recently introduced continuous quadrant penalty formulation of logical disjunctive constraints constitutes a continuous-optimization alternative to the classical formulations (such as the bigM formulation) based on the introduction of binary variables. We focus on nonlinear problems whose only combinatorial aspect comes from their logical constraints, and build on the continuous quadrant penalty, that yields to a continuous nonconvex problem, to derive an efficient computation of upper bounds to be used within Branch-and-Bound-based multi-step approaches. We show on two problems, respectively from discrete geometry and arising in air traffic management optimization, that our approach is effective at speeding up the computational convergence.

2b - FA348 - 513F - **Convexification Techniques for MINLP** Chair: AMÉLIE LAMBERT ↩

Exact Augmented Lagrangian Duality for Mixed Integer Convex Optimization

– ABHISHEK PATHAPATI PATHAPATI, INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY – Avinash Avinash Bhardwaj, Vishnu Narayanan

Augmented Lagrangian dual augments the classical Lagrangian dual with a non-negative non-linear penalty function of the violation of the relaxed/dualized constraints in order to reduce the duality gap. We investigate the cases in which mixed integer convex optimization problems have an exact penalty representation using sharp augmenting functions (norms as augmenting penalty functions). We present a generalizable constructive proof technique for proving existence of exact penalty representations for mixed integer convex programs under specific conditions using the associated value functions. This generalizes the recent results for MILP (Feizollahi, Ahmed and Sun, 2017) and MIQP (Gu, Ahmed and Dey 2020) whilst also providing an alternative proof for the aforementioned along with quantification of the finite penalty parameter in these cases.

Using quadratic cuts to iteratively strengthen convexifications of box quadratic programs

– AMÉLIE LAMBERT, CEDRIC CNAM – Daniel Porumbel

We seek the global minimum of a quadratic function f with box constrained variables. For this goal, we underestimate f by a convex piecewise-quadratic function defined as the maximum of $p \geq 1$ convex quadratic functions. We show that when $p \rightarrow \infty$ the optimal solution of this relaxation converges to an optimal solution of the strong "Shor plus RLT" semi-definite relaxation of the initial problem. To compute the new relaxation, we introduce an iterative algorithm that adds convex quadratic cuts one by one in cutting plane fashion. The resulting convexification is tighter than the one produced by previous related methods that use $p = 1$. Its integration into a spatial

branch-and-bound algorithm brings a second advantage: compared to previous work, we can refine the lower bound at each node of the branching tree. Numerical results show that even a small value of p can often be enough to reduce the branching tree size by half compared to sticking to $p = 1$. The resulting algorithm is also competitive in terms of CPU time compared to well-established solvers that rely on other techniques.

The if-then Polytope: Conditional Relations over Multiple Sets of Binary Variables

– TOBIAS KUEN, FRAUNHOFER IIS – Robert Burlacu, Patrick Gemander

I would like to present a comprehensive study on a special case of the bipartite boolean quadric polytope, which we, the authors of a recently submitted paper to this topic, namely me, Dr. Robert Burlacu and Dr. Patrick Gemander, termed the if-then polytope. Our research introduces a novel class of valid inequalities, that are sufficient to fully characterize the if-then polytope. Moreover, we present a linear program that is able to separate these inequalities in polynomial time. The presentation will highlight the practical relevance of our findings through a comprehensive computational study. We were able to demonstrate the usefulness of separating the inequalities found in the state-of-the-art branch-and-cut solver Gurobi for a real-world timetabling application for an underground train system and instances of the quadratic assignment problem. By providing a complete description of the if-then polytope, our work contributes to a deeper understanding of binary quadratic problems with multiple-choice constraints. We believe that our research will have significant impact on multiple application fields: among others, fixed recourse stochastic problems, the quadratic assignment problem and piecewise linear relaxations of nonlinear terms.

Cutting planes for multilinear optimization

– MATTHIAS WALTER, UNIVERSITY OF TWENTE – Alberto Del Pia, Emily Schutte

Multilinear constraints require that some variable is the product of other variables. In the talk I will explain different classes of cutting planes that can be applied if many such constraints are present in a MINLP. First I will discuss recent results about so-called flower inequalities in comparison to recursive McCormick linearizations. Second, separation algorithms for different classes of cutting planes will be analyzed in theory and in practice. This is joint work with Alberto Del Pia and Emily Schutte.

2b - FB9 - 513F - **Theory of Mixed Integer Non-linear Programming** Chair: AIDA KHAJAVIRAD ↩

Complexity of Mixed Integer Nonlinear Programming with Nonconvexities

– ROBERT HILDEBRAND, VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY –

We establish some complexity results for mixed integer programming in cases where there are nonconvexities describing the feasible region. This departs from most complexity results in integer programming that focus on linear or convex constraints. Nonconvexities can easily result in NP-Hard cases, even in fixed dimension 2. Our goal is to expand the class of problems that we can solve in polynomial time when the di-

mension is fixed, ideally in a fixed parameter tractable way.

A Knowledge Compilation Take on Binary Polynomial Optimization – SILVIA DI GREGORIO, UNIVERSITÉ SORBONNE PARIS NORD –

The Binary Polynomial Optimization (BPO) problem is defined as the problem of minimizing a given polynomial function over all binary points. The main contribution of this paper is to draw a novel connection between BPO and the problem of performing model counting over the $(\max, +)$ semiring on Boolean functions. This connection allows us to give a strongly polynomial algorithm that solves BPO with a hypergraph that is either beta-acyclic or with bounded incidence treewidth. This result unifies and significantly extends the known tractable classes of BPO. The generality of our technique allows us to deal also with extensions of BPO, where we enforce extended cardinality constraints on the set of binary points, and where we seek k best feasible solutions. We also extend our results to the significantly more general problem where variables are replaced by literals. Preliminary computational results show that the resulting algorithms can be significantly faster than current state-of-the-art.

On the power of linear programming for data clustering – AIDA KHAJAVIRAD, LEHIGH UNIVERSITY –

We propose a linear programming (LP) relaxation for K-means clustering. We derive sufficient conditions under which this LP relaxation is tight and subsequently obtain recovery guarantees under a popular stochastic model for the input data. We present extensive computational experiments on real data sets. Finally, we present extensions that incorporate various notations of group and individual fairness measures.

Sample Complexity and Computational Results for Branch and Cut Using Neural Networks – BARBARA ANNA FIEDOROWICZ, JOHNS HOPKINS UNIVERSITY – Hongyu Cheng, Sammy Khalife, Amitabh Basu

Data-driven algorithm design is a paradigm that uses statistical and machine learning techniques to select from a class of algorithms for a computational problem an algorithm that has the best expected performance with respect to some (unknown) distribution on the instances of the problem. We build upon recent work in this line of research by introducing the idea where, instead of selecting a single algorithm that has the best performance, we allow the possibility of selecting an algorithm based on the instance to be solved. In particular, given a representative sample of instances, we learn a neural network that maps an instance of the problem to the most appropriate algorithm *for that instance*. We formalize this idea and derive rigorous sample complexity bounds for this learning problem, in the spirit of recent work in data-driven algorithm design. We then apply this approach to the problem of making good decisions in the branch-and-cut framework for mixed-integer optimization (e.g., which cut to add?). In other words, the neural network will take as input a mixed-integer optimization instance and output a decision that will result in a small branch-and-cut tree for that instance. Our computational results provide evidence that our particular way of using neural networks for cut selection can make a significant impact in reducing branch-and-cut tree sizes, compared to previous

data-driven approaches.

2c - MB223 - 511C - Complexity and algorithmic aspects of structured families of graphs

1 Chair: KATHIE CAMERON



Local Certification of Geometric Graph Classes – AURÉLIE LAGOUTTE, UNIVERSITY GRENOBLE ALPES, FRANCE – Oscar Defrain, Louis

Esperet, Pat Morin, Jean Florent Raymond

The goal of local certification is to locally convince the vertices of a graph G that G satisfies a given property. A prover assigns short certificates to the vertices of the graph, then the vertices are allowed to check their certificates and the certificates of their neighbors, and based only on this local view, they must decide whether G satisfies the given property. If the graph indeed satisfies the property, all vertices must accept the instance, and otherwise at least one vertex must reject the instance (for any possible assignment of certificates). The goal is to minimize the size of the certificates. In this talk, we study the local certification of geometric and topological graph classes. While it is known that in n -vertex graphs, planarity can be certified locally with certificates of size $O(\log n)$, we show that several closely related graph classes require certificates of linear size. This includes penny graphs, unit-distance graphs, (induced) subgraphs of the square grid, 1-planar graphs, and unit-square graphs. These bounds are tight up to a constant factor and give the first known examples of hereditary (and even monotone) graph classes for which the certificates must have linear size. For unit-disk graphs we obtain an almost linear lower bound on the size of the certificates, and an upper bound of $O(n \log n)$.

Topological tools for H-recoloring and H-mixing – MORITZ MÜHLENTHALER, G SCOP, GRENOBLE INP, UNIVERSITÉ GRENOBLE ALPES –

The H-coloring problem asks whether there is a graph homomorphism from a given graph G to a fixed template graph H . The complexity of H-coloring is well understood in the sense that there is a classification that tells us (even for digraphs), for which graphs H the problem is in P or NP-complete. Less is known for the so-called H-recoloring problem: given two homomorphisms from G to H , can one be transformed into the other step-by-step, changing the image of a single vertex of G at each step and keeping a homomorphism from G to H throughout? Similarly, H-mixing asks whether such a transformation exists for any two homomorphisms from G to H . We present a unified view of recent complexity results for H-recoloring and H-mixing and the topological tools that are used.

Hereditary properties defined by 2-edge-colourings avoiding a set of 2-edge-coloured graphs – CÉSAR HERNÁNDEZ CRUZ, UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO –

Given a set F of 2-edge-coloured graphs and a hereditary property of graphs P , we say that F expresses P if a graph G has the property P if and only if it admits a 2-edge-colouring not having any graph in F as an induced 2-edge-coloured subgraph. In this talk, we consider the following problems. Given a hereditary property of graphs, determine if it is possible to express it by a finite set of 2-edge-coloured graphs. Given a finite set of 2-edge-coloured graphs, find a useful characterization of the hereditary property it expresses. Given a finite set of 2-edge-

coloured graphs, determine the complexity of recognizing the hereditary property it expresses.

Splitting-off in Hypergraphs – KARTHEKEYAN CHANDRASEKARAN, UNIVERSITY OF ILLINOIS, URBANA CHAMPAIGN –

The splitting-off operation in undirected graphs is a fundamental reduction operation that detaches all edges incident to a given vertex and adds new edges between the neighbors of that vertex while preserving their degrees. Lovász (1974) and Mader (1978) showed the existence of this operation while preserving global and local connectivities respectively in graphs under mild conditions. In this talk, I will introduce a splitting-off operation in hypergraphs. Main results are that there exists a local connectivity preserving complete splitting-off in hypergraphs and give a strongly polynomial-time algorithm to compute it in weighted hypergraphs. If time permits, I will illustrate the usefulness of our splitting-off operation in hypergraphs by showing two applications: (1) constructive characterization of k -hyperedge-connected hypergraphs and (2) alternate proof of an approximate min-max relation for max Steiner rooted-connected orientation of graphs and hypergraphs (due to Király and Lau (Journal of Combinatorial Theory, 2008)). Our proof of the approximate min-max relation for graphs circumvents the Nash-Williams' strong orientation theorem and uses tools developed for hypergraphs.

Based on joint work with Kristof Berczi, Tamas Kiraly, and Shubhang Kulkarni.

2c - MB921 - 511B - **Combinatorial Optimization and Graph Theory 1 Chair:** TASUKU SOMA

Weighted graph burning: Complexity and MIP formulations – STEPHAN MARNACH, RWTH AACHEN UNIVERSITY – Arie M.C.A. Koster

Graph burning is a discrete-time process mimicking the spread of fire in graphs and was introduced to model the spread of influence in networks. In every step existing fires spread to neighbouring vertices and a new vertex is set ablaze. Its associated graph parameter, the burning number, denotes the minimal number of steps needed to fully burn a graph.

We introduce a weighted variant of graph burning, where laying a fire carries a vertex dependent cost. The aim is to burn a graph within a bounded number of steps for minimal cost. We focus on a special case that amounts to only laying fire to vertices from a preselected set and show that it remains NP-hard even when restricted to path graphs.

Further, we introduce a neighbourhood based MIP-formulation for graph burning and adopt it to weighted graph burning. Preliminary computational experiments suggest that the new formulation outperforms earlier formulations by a magnitude.

Computing cuts in graphs: Homory-Hu tree and the Sparsest Cut problem – VLADIMIR KOLMOGOROV, INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (ISTA) –

I will consider two classical problems in combinatorial optimization for undirected weighted graphs. The first one is to compute a Gomory-Hu tree, which is a data structure representing minimum s-t cuts for all pairs s,t. I will present a new

algorithmic approach based on a reduction to the problem that I call "OrderedCuts". I will discuss its theoretical properties and show some experimental results, which indicate that the new approach is among current state-of-the-art GH tree construction algorithms.

The second problem is to compute the Sparsest Cut, i.e. cut (S,T) that minimizes the ratio $\text{cost}(S,T) / \min(-S-, -T-)$. The best known algorithm for this NP-hard problem is due to [Sherman'09], which computes $O(\sqrt{(\log n)/\varepsilon})$ -approximation using $O(n^\varepsilon * \text{polylog}(n))$ maxflow computations. I will present an alternative approach which simplifies Sherman's algorithm and also allows parallelization: it gives the same approximation using $O(\text{polylog}(n))$ maxflows on $O(n^\varepsilon)$ processors.

On the Polynomial Solvability of the Two-Stage Robust Perfect b-Matching Problem

– JENNY SEGSCHEIDER, RWTH AACHEN – Arie M.C.A. Koster

The perfect b-matching problem is a well-known variant of the famous matching problem with many applications. Given an undirected graph, each vertex v can be matched exactly $b(v)$ times and edges can be used multiple times. The problem can be solved in polynomial time. In theory, the capacity $b(v)$ is fixed and known for each vertex, but in practice, the capacity is often subject to uncertainties. In this talk, we present a Two-Stage variant of the perfect b-matching problem under uncertain capacity ensuring robustness, called the directed robust (perfect) b-matching problem. Here, we consider a directed graph. In the first stage, for each vertex, we set a pre-matching which fixes the value of the matching over all outgoing arcs. Then, depending on the pre-matching, the worst-case scenario is chosen from a given set of scenarios. Finally, a perfect matching is set in the second stage minimizing the costs while respecting the pre-matching as well as the capacities given by the scenario. The problem turns out to be weakly NP-hard even on bipartite graphs. However, we show that it can be solved in polynomial time on graphs without directed alternating cycle.

Spectral sparsification of hypergraphs – TASUKU SOMA, INSTITUTE OF STATISTICAL MATHEMATICS –

Spectral sparsification is a technique to approximate dense objects (e.g., weighted graphs and matrices) with sparse objects while preserving the original spectral information (e.g., eigenvalues and energy functions). This concept has been the cornerstone for developing faster algorithms for continuous/combinatorial optimization and numerical linear algebra. In this talk, I will discuss recent results of spectral sparsification for hypergraphs. Since hypergraph energy functions are no longer quadratic, spectral sparsification presents additional challenges compared to the existing spectral sparsification and requires more advanced tools. I will present recent advances in simple edge-sampling sparsification algorithms and the analysis.

2c - MC250 - 511B - **Graph Theory Chair:** FRÉDÉRIC MEUNIER

Obtaining bounds on topological indices of chemical graphs – HADRIEN MÉLOT, UNIVERSITÉ DE MONS –

In chemical graph theory, a topological index is a value which characterizes a graph representing a molecule. This value is correlated with some properties of said molecule. A widely studied class of such indices is degree-based invariants. A well-known example of this is the Randic index. Many others were introduced subsequently. In this presentation, we present a general technique which potentially allows us to obtain bounds on such chemical graph invariants.

Extremal Chemical Graphs for the Arithmetic-Geometric Index – ALAIN HERTZ, POLYTECHNIQUE MONTREAL –

The arithmetic-geometric index is a newly proposed degree-based graph invariant in mathematical chemistry. We give a sharp upper bound on the value of this invariant for connected chemical graphs of given order and size and characterize the connected chemical graphs that reach the bound. We also prove that the removal of the constraint that extremal chemical graphs must be connected does not allow to increase the upper bound.

Antimagic labelling of (k,2)-bipartite biregular graphs – GRÉGOIRE BEAUDOIRE, CNAM PARIS –

A graph $G=(V,E)$ is antimagic if there is a bijection $f : E \rightarrow \{1, 2, \dots, |E|\}$ such that all vertex sums are pairwise distinct, where the vertex sum is the sum of labels of all edges incident to a given vertex. It has been proven that regular graphs are antimagic, and more recently that a subclass of bipartite biregular graphs is antimagic as well. We show a new labelling algorithm that allows to extend the result to (k,2)-bipartite biregular graphs, as well as connected (k,l)-bipartite biregular graphs.

2c - TA263 - 511C - Graph Coloring Chair: ECKHARD STEFFEN

Column Generation for Generalized Min-Cost Flows – JONAS ALKER, TU DARMSTADT – Marc Pfetsch

The generalized min-cost flow problem is a variation of the classical min-cost flow problem with losses on the arcs. In contrast to the classical problem, the flow amount is multiplied with the arc loss when traversing an arc. The formulation is motivated by the modelling of energy systems where losses arise naturally. When applying a column generation approach on a path-based formulation, the pricing problem has the structure of a shortest path problem with losses. Due to the multiplicative behavior of these losses, a classical shortest path algorithm cannot be used to generate paths. In particular, the costs on arcs depend on previously selected arcs and their respective losses. The pricing problem can be solved via a dynamic programming scheme. Based on the corresponding Bellman equation, we formulate conditions on the input data under which the dynamic programming algorithm correctly computes the shortest paths. Additionally, we compare the performance of the column generation approach with other solution methods in practice.

Pairwise disjoint perfect matchings in r-graphs – ISAAC H. WOLF, PADERBORN UNIVERSITY – Yulai Ma, Davide Mattiolo, Eckhard Steffen

For each integer $r \geq 2$, an r -graph is an r -regular graph in which every odd set of vertices is connected to its complement by at least r edges. Being an r -graph is a necessary condition for an r -regular graph to be r -edge-colorable. On the other hand, for every $r \geq 3$ there exist r -graphs in which every two perfect matchings intersect. We show that for every $1 < k < r$ it is NP-complete to decide whether a given r -graph has k -pairwise disjoint perfect matchings. Furthermore, some results concerning the relation between the edge-connectivity and the number of pairwise disjoint perfect matchings in r -graphs will be discussed.

Injective Edge Colouring Some Classes of Sparse Graphs – ALEXANDER CLOW, SIMON FRASER UNIVERSITY –

We consider injective edge colouring problem, a theoretical model for a packet radio network problem, in which the goal is to assign communication frequencies to network node pairs in a way that eliminates secondary interference. Given a graph G an edge colouring c is injective if for all edges e and e' both incident to a third edge e'' , $c(e)$ is not equal to $c(e')$. The injective chromatic index of G is the least k such that G admits an injective k -edge colouring. We prove that the injective chromatic index of a graph with maximum degree D and degeneracy d is at most big-O of d cubed times $\log(D)$, and that the injective chromatic index of a graph of Euler genus g is at most $(3+o(1))g$. The upper bound for Euler genus is tight up to the choice of lower order term when G is a clique. We will spend any remaining time exploring how we use the ideas from these proofs to bound the oriented chromatic number above by a polynomial in Euler genus.

Joint work with Peter Bradshaw (University of Illinois Urbana Champaign), and Jingwei Xu (University of Illinois Urbana Champaign).

Sets of r -graphs that color all r -graphs – ECKHARD STEFFEN, PADERBORN UNIVERSITY –

An r -regular graph is an r -graph, if every odd set of vertices is connected to its complement by at least r edges. Let G and H be r -graphs. An H -coloring of G is a mapping $f: E(G) \rightarrow E(H)$ such that each r adjacent edges of G are mapped to r adjacent edges of H . For every $r \geq 3$, let $\mathcal{H}_r$ be an inclusion-wise minimal set of connected r -graphs, such that for every connected r -graph G there is an $H \in \mathcal{H}_r$ which colors G .

We show that $\mathcal{H}_r$ is unique and characterize $\mathcal{H}_r$ by showing that $G \in \mathcal{H}_r$ if and only if the only connected r -graph coloring G is G itself.

The Petersen Coloring Conjecture states that the Petersen graph P colors every bridgeless cubic graph. We show that if true, this is a very exclusive situation. Indeed, either $\mathcal{H}_3 = \{P\}$ or $\mathcal{H}_3$ is an infinite set and if $r \geq 4$, then $\mathcal{H}_r$ is an infinite set. Similar results hold for the restriction on simple r -graphs. Moreover, we determine the set of smallest r -graphs and show that it is a subset of $\mathcal{H}_r$.

This is joint work with Yulai Ma, Davide Mattiolo and Isaac H. Wolf.

Preference-based Constrained Optimization – MALEK MOUHOUB, UNIVERSITY OF REGINA – Sultan Ahmed

The Conditional Preference Network (CP-net) graphically represents the user's qualitative and conditional preference statements under the ceteris paribus interpretation. The constrained CP-net extends the CP-net to a set of hard constraints. The constrained CP-net can help represent and solve real-world applications where the goal is to select one or more scenarios that are feasible according to constraints while maximizing user's preferences.

Existing algorithms for solving the constrained CP-net require expensive dominance testing operations. To overcome this challenge in practice, we propose several efficient solving approaches. In our first approach, we alter the constrained CP-net by eliciting additional relative importance statements between variables to have a total order over the outcomes. This new model is the constrained Relative Importance Network (constrained CPR-net). Consequently, we show that the constrained CPR-net has one optimal outcome (assuming the constrained CPR-net is consistent) that we can obtain without dominance testing. In our second solution, we extend the Lexicographic Preference Tree (LP-tree) to a set of constraints. Then, we propose a recursive backtrack search algorithm to find the most preferable outcome. We show that the first feasible outcome returned by Search-LP (without dominance testing) is also preferable to any other feasible outcome. Finally, in our third solution, we preserve the semantics of the CP-net and propose a variant of the branch-and-bound algorithm improved with look-ahead strategies and variable ordering heuristics. A divide-and-conquer variant of the branch-and-bound algorithm is also explored as an alternative method.

Exact approaches for solving interdependent multi-stage interdiction games: Applications in Communication Problems – JOSE WALTEROS, UNIVERSITY AT BUFFALO – Rafal Muszalski

We study a general class of interdiction games that model an adversarial, finite, sequential, and alternating interaction between two decision-makers—i.e., an attacker and a defender. A prototypical example of these types of games consists of a three-stage defender-attacker-defender problem in which the defender conducts a set of fortification strategies during the first stage to protect its operation from the attacker's course of actions; in the second stage, the attacker deploys a set of disruptive attacks, parameterized by the fortification strategies, to optimally deteriorate the defender's operational objective; finally, in the third stage, the defender optimizes their operation, considering the effects of the previous stages.

We focus on a general variation in which there is a direct (and possibly disruptive) interaction between the actions taken by non-consecutive rounds. For example, in the three-stage game described before, we consider instances in which actions made by the defender in the first stage may also turn solutions for the defender's operation in the third stage infeasible. General applications of these games arise in prob-

lems with an underlying network structure, where the defender can alter the composition of the network in the first stage by adding/removing edges and vertices. In this presentation, we develop an exact solution framework for these problems based on a combinatorial branch-and-bound search. We provide some examples of applications arising from communications and network design problems.

Sparse factorization of the square all-ones matrix of arbitrary order – EDWARD DUC HIEN NGUYEN, RICE UNIVERSITY – Xin Jiang, Cesar A Uribe, Bicheng Ying

In this presentation, we study sparse factorization of the (scaled) square all-ones matrix J of arbitrary order. We introduce the concept of hierarchically banded matrices and propose two types of hierarchically banded factorization of J : the reduced hierarchically banded (RHB) factorization and the doubly stochastic hierarchically banded (DSHB) factorization. Based on the DSHB factorization, we propose the sequential doubly stochastic (SDS) factorization, in which J is decomposed as a product of sparse, doubly stochastic matrices. Finally, we discuss the application of the proposed sparse factorizations to the decentralized average consensus problem and decentralized optimization.

The Edge-based Contiguous p-median Problem with Connections to Territorial Districting – ADOLFO ESCOBEDO, NORTH CAROLINA STATE UNIVERSITY – Zeyad Kassem

This talk introduces the edge-based contiguous p-median (ECpM) problem to partition the roads in a network into a given number of compact and contiguous territories. Two binary programming models are introduced, both of which incorporate path-based node-to-edge distances to reflect road travel. The first model requires an exponential number of cut set-based constraints to model contiguity; it is paired with a separation scheme that usually generates only a small number of these constraints, namely, a branch-and-cut (B&C) algorithm. The second model utilizes a polynomial number of shortest-path constraints to model contiguity and can be solved with off-the-shelf solvers. The two solution approaches associated with the proposed models are tested on road networks with over 2,700 nodes and close to 3,400 edges, yielding models with over 9.6 million binary variables. Solving the model based on shortest path contiguity (SPC) constraints via standard branch and bound attains speedups in computational time of up to 52x relative to the cut set-based B&C implementation. In addition, the SPC constraints are demonstrated to be non-valid logic cuts of the edge-based p-median (EpM) model (i.e., for which contiguity is not explicitly required), meaning that they cut off integer feasible but non-optimal solutions of this simpler problem. Finally, the talk explores structural insights and connections between ECpM and the edge-based districting (EBD) problem, which enforces an additional planning criterion (work balance). An existing model that utilizes cut set-based contiguity constraints was unable to find a feasible solution within 12 hours for all instances, while the new EBD model with SPC constraints was able to solve most of the feasible instances to optimality.

Graphs with well-structured eigenbases – HERMIE MONTERDE, UNIVERSITY OF MANITOBA –

Let $S \subseteq \mathbb{R}$. A matrix M is said to be S -diagonalizable if there exists a matrix P having entries from S such that $M = P\Delta P^{-1}$ for some diagonal matrix Δ . A graph X is said to be S -diagonalizable if its Laplacian matrix $L := D - A$ is S -diagonalizable (Johnston and Plosker, 2023). Here, D is the diagonal matrix of vertex degrees of X and A is the adjacency matrix of X . In this talk, we survey results on S -diagonalizable graphs, where $S = \{-1, 0, 1\}$ or $S = \{-1, 1\}$. In particular, we will focus on the case where P is a Hadamard (that is, P has entries from $\{-1, 1\}$ and the columns of P are pairwise orthogonal) or a weak Hadamard matrix (that is, P has entries from $\{-1, 0, 1\}$ and the nonconsecutive columns of P are pairwise orthogonal). Our main motivation for considering these graphs is its potential application to theory of continuous-time quantum walks, where knowledge about eigenvectors of certain matrices associated to graphs (more specifically, their orthogonal projection matrices) is important in determining which types of quantum state transfer occur in a graph. We also present open problems concerning S -diagonalizable graphs, if time allows. Some of the material presented in this talk is joint work with Darian McLaren and Sarah Plosker.

On the max k-cut and the eigenvalue of a graph – LEONARDO LIMA, UNIVERSIDADE FEDERAL DO PARANÁ –

Let G be a graph. The maximum k -cut problem of a graph G is the maximum number of edges in a k -partite subgraph of G . The max k -cut problem is a fundamental NP-hard combinatorial optimization problem, and it is related to many applications such as data clustering, scheduling, power networks, and others. More recently, upper bounds to the max k -cut have been developed regarding the eigenvalues of a matrix related to graph G . In this talk, we present an extension of the graph parameter max k -cut to square matrices and prove a general sharp upper bound, which implies upper bounds on the max k -cut of a graph using the smallest signless Laplacian eigenvalue, the smallest adjacency eigenvalue, and the largest Laplacian eigenvalue of the graph. In addition, we show that our bounds are the best possible by constructing infinite families of extremal graphs for the obtained upper bounds.

Main Signless Laplacian Eigenvalues of Quasi-Threshold Graphs – CYBELE VINAGRE, UNIVERSIDADE FEDERAL FLUMINENSE –

This talk is about main eigenvalues of the signless Laplacian matrix of graphs. For a simple graph G with vertices $v_1, \dots, v_n$, its signless Laplacian matrix $Q(G) = [q_{ij}]$ is the square matrix of order n where q_{ii} is equal to the degree of vertex v_i and, for $i \neq j$, $q_{ij} = 1$ if v_i and v_j are adjacent and 0, otherwise. An eigenvalue λ is a main eigenvalue of $Q(G)$ (or a main Q -eigenvalue) if the eigenspace $E(\lambda)$ of λ is non-orthogonal to the all ones vector $\mathbf{j}$, that is, if $E(\lambda)$ contains some eigenvector for which the sum of its entries is non-zero. In a previous work, we gave an exact characterization of the threshold graphs having k main Q -eigenvalues, for each $k \geq 2$. Here we give a step forward to generalize this characterization for quasi-threshold graphs. We may recall that threshold graphs form a subclass of cographs characterized as the class of $\{P_4, C_4, 2K_2\}$ -free graphs, while quasi-threshold graphs constitute the subclass of $\{C_4, P_4\}$ -free graphs. The main result of this work is a structural characterization of connected quasi-threshold graphs

having $k \geq 2$ main Q -eigenvalues. In addition, we determine precisely what are the connected quasi-threshold graphs having two main Q -eigenvalues. We also provide a full characterization of a subclass of quasi-threshold graphs: more precisely, for $k = 2, 3, \dots$, we determine what are the generalized core-satellite graphs having k main Q -eigenvalues. This is a joint work with Átila Jones and Vilmar Trevisan.

Zero forcing and eigenvalue multiplicities – SHAHLA NASSERASR, ROCHESTER INSTITUTE OF TECHNOLOGY –

Zero forcing game starts with an initial set of blue and white vertices of a graph, and the goal is to color all vertices blue under some color change rule at minimum cost. The rule for the original zero forcing game is that if all of the neighbors of a blue vertex are blue except one, then the white neighbor is forced to be blue. The zero forcing process applies this rule in discrete time-steps until no more changes are possible. The zero forcing number of a graph G is the minimum number $Z(G)$ of vertices so that there exists an initial blue coloring with $Z(G)$ vertices that is transformed to an all-blue coloring of G by the zero forcing process. The zero forcing was introduced to study the maximum multiplicity of eigenvalues of matrices associated with a graph. In this talk, some properties and applications of the zero forcing number will be discussed.

2c - TC359 - 511B - Advances in Combinatorial Optimization Chair: TAMON STEPHEN

An Introduction to Circuit Walks – STEFFEN BORWARDT, UNIVERSITY OF COLORADO DENVER –

We provide an introduction to the studies of circuit walks and review the past decade of research on circuit diameters and augmentation. We connect the field to classical topics such as conformal sums, combinatorial diameters, and vertex enumeration, discuss the efficient computation of improving circuits, and end with some open questions.

A Yannakakis-type theorem for affine semigroups – AMY WIEBE, UNIVERSITY OF BRITISH COLUMBIA – João Gouveia

The seminal result on lifts of polyhedral cones is the theorem of Yannakakis that relates lifts of cones with nonnegative factorizations of their slack matrices. In this talk we generalize the notion of slack matrices to affine semigroups - a discrete analog of polyhedral cones - and present the corresponding result relating lifts of affine semigroups to nonnegative integer factorizations of their slack matrices. We use these generalizations to present new results on nonnegative integer rank of integer matrices.

Hypergraph Transversal Pairs Near the Fredman-Khachiyan Bound – TAMON STEPHEN, SIMON FRASER UNIVERSITY –

The Fredman-Khachiyan algorithm generates the transversal of a hypergraph in incremental quasi-polynomial time. It is a recursive algorithm which focuses on the most frequent vertex in terms of the number of edges in either the hypergraph or its transversal. Hypergraphs where this maximum frequency is low are therefore of interest. This gives a group of related optimization questions for hypergraph-transversal pairs. Here we present some preliminary extremal results.

Besides using classic combinatorial constructions, we adapt Wagner's deep reinforcement learning scheme from graphs to hypergraphs to help find some unexpected examples.

Joint work with Parsa Salimi.

2c - WA169 - 511C - Complexity and algorithmic aspects of structured families of graphs

2 Chair: CÉSAR HERNÁNDEZ CRUZ

Reconfiguration of Vertex Colourings in Hereditary Classes of Graphs – KATHIE CAMERON, WILFRID LAURIER UNIVERSITY –

A k -colouring of a graph is an assignment of one of k colours to each vertex so that the ends of each edge get different colours.

The reconfiguration graph of the k -colourings of graph G , denoted $Rk(G)$, is the graph whose vertices are the k -colourings of G , and two colourings are adjacent in $Rk(G)$ if they differ in colour on exactly one vertex. We are interested in whether the reconfiguration graph is connected, and if so, its diameter. Being connected means we can obtain any k -colouring from any other by changing the colour of one vertex at a time, while always having a k -colouring.

We call a graph recolourable if $Rk(G)$ is connected for every k bigger than the chromatic number of G . We have characterized the graphs H such that all graphs G which don't have H as an induced subgraph are recolourable. We have done the same for other hereditary graph classes including when two 4-vertex graphs are excluded as induced subgraphs.

We also explore graph decompositions applied to recolouring.

This is joint work with Manoj Belavadi, Owen Merkel and Ni Luh Dewi Sintari.

Gallai-like characterization of strong cocomparability graphs – JING HUANG, UNIVERSITY OF VICTORIA –

Strong cocomparability graphs are the reflexive graphs whose adjacency matrix can be rearranged by a simultaneous row and column permutation to avoid the submatrix with rows 01, 10. Strong cocomparability graphs form a subclass of cocomparability graphs (i.e., the complements of comparability graphs) and can be recognized in polynomial time. In his seminal paper, Gallai characterized cocomparability graphs in terms of a forbidden structure called asteroids. Gallai proved that cocomparability graphs are precisely those reflexive graphs which do not contain asteroids.

In this talk, I will present a characterization of strong cocomparability graphs which is analogous to Gallai's characterization for cocomparability graphs. More precisely, I will show that strong cocomparability graphs are precisely those reflexive graphs which do not contain weak edge-asteroids (a weaker version of asteroids). The characterization also leads to a polynomial time recognition algorithm for strong cocomparability graphs.

On Mixed Cages – M. GABRIELA ARAUJO PARDO, MATH INSTITUTE, UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO –

I will give a review of the state of the art on mixed cages since we introduced them in a joint work with C. Hernández, and

J.J. Montellano-Ballesteros, in 2019 (see [1]). A mixed regular graph is graph $G = (V; E \cup A)$ where $V(G)$ is the set of vertices, $E(G)$ is the set of edges, and $A(G)$ is the set of arcs where each vertex has z in and out arcs, and r edges. A mixed cycle is a cycle with edges and arcs, and all the arcs are traversed in the same direction, then the girth in a mixed graph is the length of the minimal mixed cycle. A $[z, r; g]$ -a mixed graph of girth g is a mixed regular graph of girth g . Moreover, a $[z, r; g]$ -mixed cage of girth g is a $[z, r; g]$ -mixed graph of minimum order. In [1] we give simple results about this and a nice lower bound, denoted by $n[1, r; g]$, for $[1, r; g]$ -graphs attained for $r = 2$ and any girth $g \geq 3$. Recently, in 2023, Exoo in [4] proved that this bound, called now as the AHM-Bound, is attained also for $r = 3$ and $g = 6$ finding a $[1, 3, 6]$ -cage with 30 vertices. I will show the graphs attained from Exoo in [4] using searching computational techniques, and mixed graphs of girth 5, obtained by G. Araujo, C. De la Cruz, and D. González-Moreno ([2]), using finite geometries. Moreover, in that paper, we give results on the structure of mixed cages related to connectivity and monotonicity. Also, I will talk about a recent work with L. Mendoza (see [3]), on mixed cages of girth 6 and new lower bounds for any three values $z \geq 1, r \geq 2$ and $g \geq 3$. In this work, we describe the $[1, 3, 6]$ -cage given by Exoo in [4] as a subgraph with some "added arcs" of a $(4, 6)$ -cage or the incidence graph of a projective plane of order 3. Finally, I will give other results given by Jacay and Jacayovna in two recent conferences related to $[z, r; g]$ -mixed graphs with few vertices using the $CD(n, q)$ -graphs introduced by Lazebnik, Ustimenko, and Woldar in [5].

[1] Graphs and Combinatorics, 35 (5):989–999, (2019). [2] Discrete Mathematics, 345 (5):112792, (2022). [3] (submitted). [4] Discrete Mathematics and Theoretical Computer Science, vol. 25:2, (2023). [5] J. Comb. Theory, Ser. B 61 (1) 111–117 (1994).

Complexity dichotomies for list homomorphism problems of signed graphs – JAN BOK, CHARLES UNIVERSITY – Brewster Richard C., Feder Tomás, Pavol Hell, Jedlicková Nikola, Rafiey Arash

We investigate the computational complexity of list homomorphism problems for signed graphs. The CSP dichotomy conjecture has been recently established, but several other dichotomy questions remain open, including the dichotomy classification of list homomorphism problems for signed graphs. Compared to the classical homomorphism problems, we additionally get lists of admissible targets for each vertex of the input graph in the list version.

Usually, a dichotomy classification is easier to obtain for list homomorphisms than for homomorphisms. However, this is not the case for signed graphs. The complexity classification of the list homomorphism problem proves to be challenging, and as of now, no structural classification conjecture has been proposed. I will show a structural classification in the special case of weakly balanced irreflexive and reflexive signed graphs, a conjecture previously made by Kim and Siggers. This builds up on our previous results for signed trees and separable signed graphs, which I will briefly outline as well. The proof for the irreflexive case relies on establishing an extension theorem for min orderings of (unsigned) bipartite graphs, which is interesting on its own.



The betweenness model for the single-row facility layout problem and beyond – ANJA FISCHER, TU DORTMUND UNIVERSITY – Frank Fischer
 Given a set of departments, their lengths and pairwise weights the single-row facility layout problem (SRFLP) asks for an arrangement of the departments along one side of a path such that the weighted sum of the center-to-center distances is minimized. In 2009 Amaral presented the well-known betweenness model for the SRFLP. We will strengthen his model by new cutting planes which are separated heuristically in our tests. The combination with some new construction heuristic allows us to solve instances which up to 60 departments exactly. In the second part of this talk we extend our setting of the standard SRFLP since in practice, especially in job shop production systems, there are often several departments which are of the same or a similar type and could do the same job steps. So the paths of the products and the exact transport amounts between each pair of departments, which equal the weights in the SRFLP, are not known beforehand. This motivates the simultaneous assignment of the tasks to the departments without exceeding the capacity of each of them as well as the determination of an SRFLP layout in order to minimize the overall transportation costs. We present two new models for this extended problem. Computational results show that a pure 0-1 modelling approach based on the betweenness model combined with symmetry-breaking constraints works best.

Envy-free Pricing of Seats in a Planetarium – FREIJA VAN LENT, MAASTRICHT UNIVERSITY – Julian Golak, Alexander Grigoriev

In this paper, we explore envy-free pricing problems with unit-demand that naturally arise when considering revenue maximization strategies for venues like planetaria. We are given a set of customers, each with a preference for one specific seat within the planetarium's layout. For any seat other than their preferred one, a customer's valuation decreases with the distance from their preferred seat. The planetarium operator decides on the prices for all seats within the layout and aims to maximize revenue by pricing seats and allocating them to customers in an envy-free and price non-discriminatory manner. This means that every assigned customer receives a seat that provides the highest possible utility for given non-discriminatory prices, while all unassigned customers do not desire any available seat. We examine a continuous variant where customers can be placed arbitrarily within the Planetarium's layout, without being limited by seat size constraints. In the discrete counterpart, seats layout correspond to typical seat configurations in cinemas and theaters. For the special case of the continuous variant of the problem, wherein all customers have a mutual preferred seat, we introduce a dynamic programming algorithm solving the problem in polynomial time. For the discrete variant of the problem, we extend the dynamic programming approach, solving the problem in pseudo-polynomial time. In the final part of our study, we unravel an intriguing connection between the arrangement of seats and principles of lattice theory.

Packing Hypertrees and the k-cut problem in hypergraphs – FRANCISCO BARAHONA, IBM RE-

SEARCH – Mourad Baiou

We present a combinatorial algorithm for determining a maximum packing of hypertrees in a capacitated hypergraph. This extends algorithms for packing spanning trees in graphs. It gives an algorithmic proof of a theorem by Frank et al. This allows the extension of several algorithms developed for graphs to hypergraphs, for the k-cut problem.

Finding Maximum Edge-Disjoint Paths Between Multiple Terminals – SATORU IWATA, UNIVERSITY OF TOKYO – Yu Yokoi

Let $G = (V, E)$ be a multigraph with a set T of terminals. A path in G is called a T -path if its ends are distinct vertices in T and no internal vertices belong to T . In 1978, Mader showed a characterization of the maximum number of edge-disjoint T -paths.

In this talk, we provide a combinatorial, deterministic algorithm for finding the maximum number of edge-disjoint T -paths. The algorithm adopts an augmenting path approach. More specifically, we utilize a new concept of short augmenting walks in auxiliary labeled graphs to capture a possible augmentation of the number of edge-disjoint T -paths. To design a search procedure for a short augmenting walk, we introduce blossoms analogously to the matching algorithm of Edmonds (1965). When the search procedure terminates without finding a short augmenting walk, the algorithm provides a certificate for the optimality of the current edge-disjoint T -paths. From this certificate, one can obtain the Edmonds-Gallai type decomposition introduced by Sebő and Szegő (2004). The algorithm runs in $O(|E|^2)$ time, which is much faster than the best known deterministic algorithm based on a reduction to linear matroid parity.

We also present a strongly polynomial algorithm for the maximum integer free multiflow problem, which asks for a non-negative integer combination of T -paths maximizing the sum of the coefficients subject to capacity constraints on the edges.



Good and Fast Row-Sparse ah-Symmetric Reflexive Generalized Inverses – GABRIEL PONTE, UNIVERSITY OF MICHIGAN / FEDERAL UNIVERSITY OF RIO DE JANEIRO – Marcia Fampa, Jon Lee, Luze Xu

We construct sparse and structured generalized inverses, with application to the efficient computation of least-squares solutions of inconsistent systems of linear equations. Leveraging our earlier formulations to minimize the 1- and 2,1- norms of generalized inverses that satisfy important properties of the Moore-Penrose pseudoinverse, we develop efficient and scalable ADMM algorithms to address these norms minimization, and to limit the number of non-zero rows in the solution. We prove a 2,1-norm approximation result for a local-search procedure that we originally used for 1-norm minimization, and compare the ADMM algorithms with the local-search procedure and optimization solvers.

A column-generation based heuristic for integrating machine scheduling and personnel allocation – DANIELA LUBKE, UNIVERSITY OF WATTERLOO – Ricardo Fukasawa, Luis Ricardez-Sandoval

This work investigates the integration of machine scheduling and personnel allocation problems. In the machine scheduling problem one aims to find an optimal assignment of jobs to tasks in a given time horizon while only considering processing times and machine capacities in a time-discretized formulation. The goal of personnel allocation problems is to find the best employee allocation, considering business, regulatory, or satisfaction constraints (e.g., set the maximum time an employee can work in a specific activity, how long an employee can be in one particular activity without a period of rest, or which are the preferred days for work). Those two problems are intrinsically related and should ideally be considered together in a single monolithic formulation, though doing so is computationally challenging. This work presents a column generation-based algorithm (CG) used to find good quality solutions for an industrial-scale analytical services facility where integrated machine scheduling and personnel allocation decisions must be taken daily. Our computational results show that the proposed approach is able to consistently find high-quality solutions even when the size of the problem grows, and it outperforms other proposed approaches.

Semialgebraic characterization of triangulation – YUAN ZHOU, UNIVERSITY OF KENTUCKY

The study of triangulations of finite point configurations is a rich topic at the intersection of combinatorics, geometry, and optimization. It is well known that two point configurations A and B with identical oriented matroids (encoded by the set of circuits or cocircuits, equivalently) have the same set of triangulations. Given a triangulation T of A , we consider the coordinates of the points in A as parameters, and ask: which set of polynomial inequalities on the parameters ensures that T is a triangulation of A ? That is, we are interested in the conditions under which A and B share a particular triangulation but not necessarily all triangulations. We show that, in low dimensions, a restricted set of cocircuits corresponding to hyperplanes spanned by facets of the simplices in T suffices. This gives convenient parametric volume formulas with few side conditions, with numerous obvious applications in cutting planes, computational biology and optimal control.

On disjunction convex hulls by lifting – YUSHAN QU, UNIVERSITY OF MICHIGAN – Jon Lee

In the study of disjunctive programming, a critical aspect is the exploration of feasible regions formed by the union of multiple convex sets. The ‘big-M’ method is well-established and often used (and abused) for formulating multiple-choice constraints. It is well-known that such formulations can be rather weak and also suffer from numerical instability, even when the big-M coefficients are chosen optimally. So, MILP/MINLP solvers often struggle with these formulations, while they are ubiquitous in MILP/MINLP applications. On the other hand, sometimes these formulations are best possible; that is, the inequalities can give a complete description of the convex-hull of the mixed-integer solutions, and we considered broad and useful situations where this is indeed the case. In such situations, the ‘big-M’ inequalities can be derived by optimally lifting facet-describing inequalities of the disjunctive pieces.

Specifically, we study the natural extended-variable formulation for the disjunction of $n + 1$ polytopes in $\mathbb{R}^d$. We demonstrate that the convex hull D in the natural extended-

variable space $\mathbb{R}^{d+n}$ is given by full optimal big-M lifting (i) when $d \leq 2$ (and that it is not generally true for $d \geq 3$), and also (ii) when the polytopes are all axis-aligned hyperrectangles. We note that both of these cases are very natural for applications and within spatial branch-and-bound, the main algorithmic paradigm for mixed-integer nonlinear optimization. Additionally, we give further results on the polyhedral structure of D , giving very general conditions under which it is full dimensional and when the full optimal big-M liftings of the facet-describing inequalities for the individual polytopes describe facets of D .

2c - WB185 - 511C - Complexity and algorithmic aspects of structured families of graphs 3 Chair: CHÍNH HOÀNG

H-Chromatic Symmetric Functions – ANGELE FOLEY, WILFRID LAURIER UNIVERSITY – Nancy Mae Eagles, Alice Huang, Elene Karangozishvili, Annan Yu

We recently introduced H-chromatic symmetric functions which use the H-coloring of a graph G to define a generalization of Stanley’s chromatic symmetric functions. In this talk we take a tour of these new symmetric functions, considering equivalence questions and basis questions. We also include conjectures and open problems.

Coloring graphs by forbidden induced subgraphs – CHÍNH HOÀNG, WILFRID LAURIER UNIVERSITY –

Given a set L of graphs, a graph G is L -free if G does not contain any graph in L as an induced subgraph. There has been keen interest in coloring graphs whose forbidden list L contains graphs with four vertices. Recent results in the literature identify three outstanding classes: $L=(4K1, \text{claw})$, $L=(4K1, \text{claw, co-diamond})$, and $L=(4K1, C4)$. In this talk, we will discuss recent advances on these three problems.

Intersection of chordal graphs and some related partition problems – R. SRITHARAN, UNIVERSITY OF DAYTON – Atif Abueida, Arthur Busch

The chordality of a graph is the minimum number of chordal graphs whose intersection is the graph. A result of Yannakakis’ from 1982 can be used to infer that for every fixed k at least 3, deciding whether the chordality of a graph is at most k is NP-complete. We consider the problem of deciding whether the chordality of a graph is 2 and present results for a version of the problem.

Efficient certifying algorithms for k-colouring subfamilies of P5-free graphs – BEN CAMERON, UNIVERSITY OF PRINCE EDWARD ISLAND –

The family of P_5 -free graphs is of particular interest when studying graph colouring as P_5 is the largest connected graph that can be forbidden as an induced subgraph among input graphs that results in polynomial-time algorithms to decide k -colourability for all k (assuming that P is not equal to NP). These algorithms can be extended to certifying ones for subfamilies of P_5 -free graphs if there are only finitely many $(k+1)$ -critical graphs among the subfamily, where a critical graph is one whose chromatic number decreases with the deletion of any vertex. In this talk, we will explore some of the subfamilies where these efficient certifying algorithms have been developed

along with a brief overview of the proof techniques involved. We will then present some of our new results and highlight the new, simpler proof techniques used to obtain them.

This talk contains joint work with Chính Hoàng and Thaler Knodel.

2c - WC353 - 511B - **Combinatorial Optimization** Chair: FABIO FURINI 

Recent heuristic development for the quadratic knapsack problem – FRANKLIN DJEU-MOU FOMENI, UQAM –

The Quadratic Knapsack Problem (QKP) is a well-known combinatorial optimization problem which has many applications in finance, logistics, telecommunications, facility location problems, etc. The QKP is NP-hard in the strong sense and the existing state-of-the-art algorithms can only handle problems of small and moderate sizes. This presentation is mainly aimed at discussing some recent contributions in the development of heuristic solution methods for large scale and challenging QKP instances. The strengths as well as the weaknesses of each of those methods will be discussed.

Critical infrastructures location via submodular maximization – FABIO FURINI, SAPIENZA UNIVERSITY OF ROME – Stefano Coniglio, Ivana Ljubic

We study a family of discrete optimization problems asking for the maximization of the expected value of a concave, strictly increasing, and differentiable function composed with a set-union operator. The expected value is computed concerning a set of coefficients taking values from a discrete set of scenarios. The function models the utility function of the decision maker, while the set-union operator models a covering relationship between two ground sets, a set of items, and a set of metaitems. This problem and it can be modeled as a mixed integer nonlinear program involving binary decision variables associated with the items and metaitems. Its goal is to find a subset of metaitems that maximizes the total utility corresponding to the items it covers. It has applications to, among others, maximal covering, location of critical infrastructures and influence maximization problems.

Processual programming - Using business process models to formulate mathematical programs – ASVIN GOEL, KÜHNE LOGISTICS UNIVERSITY –

Despite wide application domains, significant optimisation potential, and tremendous advancements in solving mathematical programs over the last decades, it appears that mathematical programming still faces difficulties with regard to wider adoption in many organisations. Moreover, its importance in the perception of the general public appears to fall short in comparison with other approaches, e.g. those subsumed under the term artificial intelligence. One obstacle to wider adoption of mathematical programming is the widespread lack of mathematical proficiency in many organisations. Consequently, developing and maintaining mathematical models of decision problems is a time- and cost-consuming challenge with a large upfront investment and unknown return. On the other hand, many organisations have already developed substantial modelling skills in terms of business process modelling due to

the immediate benefit that can be obtained even without optimisation. With BPMN 2.0 an industry standard for business process modelling has evolved with a growing toolset supporting the standard. Leveraging the growing software ecosystems and modelling capabilities concerning BPMN 2.0, this presentation proposes a new class of mathematical programs allowing domain experts to formulate mathematical programs without the need to learn how to model a linear program, mixed integer program, constraint program and many others. Instead, the new class of mathematical programs allows modellers to specify so-called processual programs by a set of process models with additional data relevant for process execution. Examples show how classical optimisation problems, e.g., the assignment problem, the knapsack problem, the bin packing problem, job shop scheduling problems, vehicle routing problems, etc. can be formulated as processual programs. With the graphical representation of business problems, processual programs can build an important communication tool between domain experts and optimisation experts. Moreover, general-purpose solvers for processual programs can be developed, allowing domain experts to model and solve optimisation problems without the need to consult optimisation experts as long as no special-purpose approach exploiting the specific characteristics of the application case is necessary. The development of powerful general-purpose processual programming solvers, however, is still a challenge that requires future research. To facilitate such research, a modular open source framework allowing to simulate and optimise processual programs has been developed. The framework allows to easily replace modules responsible for making decisions during process execution and, thus, allows module developers to focus on the decision-making approaches without the need to implement process execution logic.

2c - ThA924 - 511B - **Combinatorial Optimization and Graph Theory 4** Chair: LUCAS MOSCHEN 

Totally equimodular matrices – ROLAND GRAPPE, UNIVERSITY PARIS DAUPHINE, PSL LAMSADE –

Totally unimodular matrices play a fundamental role in combinatorial optimization, as they offer combinatorial min-max theorems such as the MaxFlow-MinCut theorem and König's theorem about matchings in bipartite graphs. Recently, a generalization called totally equimodular matrices has emerged within the context of box-totally dual integral polyhedra, which are the polyhedra hidden underneath "strong" min-max relations. This talk aims to explore some connections between totally unimodular and totally equimodular matrices, offering insights into promising avenues while also addressing one notable drawback.

A 2/3-approximation algorithm for the maximum weight spanning star forest problem – VIET HUNG NGUYEN, LIMOS CLERMONT AUVERGNE UNIVERSITY –

Given an undirected graph $G = (V, E)$ where the edges in E is weighted positively, a star is a subgraph of G of diameter at most 2. A star forest is a collection of (vertex) disjoint stars in G . In particular, an isolated vertex is also considered as a star. A spanning star forest is a star forest whose ver-

text set is V . In this paper, we consider the Maximum Weight Spanning Forest Problem (MWSFP) which consists in finding a spanning forest of maximum weight in G . The MWSFP is defined in a paper by Nguyen et al. (SODA 2005) where it is proved to be NP-hard. In the same paper, the authors have given of a $\frac{1}{2}$ -approximation algorithm for MWSFP and also a 0.64-approximation algorithm for the unweighted version. Since then several improvements have been done for the unweighted version of MWSFP to achieve an approximation ratio 0.804 (Athanasopoulos et al. MFCS 2009) but the approximation ratio $\frac{1}{2}$ obtained by Nguyen et al. remains the best so far for the MWSFP. In this paper, we improve the approximation ratio for MWSFP to $\frac{2}{3}$. Our algorithm is mainly an iterated rounding approximation algorithm based on a simple integer formulation for MWSFP.

On the Congruency-Constrained Matroid Base – SIYUE LIU, CARNEGIE MELLON UNIVERSITY – Chao Xu

Consider a matroid where all elements are labeled with an integer. We are interested in finding a base where the sum of the labels is congruent to $g \bmod m$. We show that this problem is fixed parameter tractable if we parametrize by m when m is either the product of two primes or a prime power. The algorithm can be generalized to all moduli and, in fact, to all abelian groups if a classic additive combinatorics conjecture by Schrijver and Seymour holds true. We also discuss the optimization version of the problem.

Combinatorial algorithms for large household allocation problems – LUCAS MOSCHEN, TRIER UNIVERSITY – Ulf Friedrich, Ralf Münnich, Martin Schmidt

Microsimulation models are efficient tools for predicting changes in certain aspects of society, such as demography and economy. A dynamic spatial microsimulation model is being developed for the entire German population in the project MikroSim. A fundamental stage of this model is the use of available geographic statistical information to create a high-quality allocation of households in dwellings at the municipal level. In this work, the step of household-dwelling allocation is treated as a maximum weight matching problem with side constraints, which in general is too large to be solved by the available optimization methods. We present two different solution algorithms for the problem. The first is a case-specific decomposition algorithm, which generates an allocation with the property that it cannot be extended by any assignment without becoming infeasible. The second is a penalty approach where approximation algorithms are applied to obtain solutions for the penalized models, guaranteeing a certain optimality level for the final allocation. Numerical experiments show that both methods find high-quality allocations using less time and memory than standard algorithms, which allows them to handle instances with more than one billion variables.

2c - ThB352 - 511B - **Integer/Linear Optimisation** Chair: DANIELE CATANZARO 

Enumerating full binary trees in polynomial delay – YASUKO MATSUI, TOKAI UNIVERSITY – Hirotaka Ono, Dominik Koepl

A binary tree is a rooted tree in which each node has at most

two children. A full binary tree is one in which each node has exactly two children or no child. A node is called a leaf if it has no child. For a given rooted tree and a node v , the number of edges from the root to v is called the depth of v . For a non-negative integer vector $w = (w_1, w_2, \dots, w_d)$, if there is a full binary tree T such that each w_i corresponds to the number of leaves at depth i in T , where w is said to be a leaf vector of full binary trees. In this talk, we present an algorithm to enumerate all full binary trees for a given leaf vector. Our algorithm runs in constant delay and generates each full binary tree in $O(\max w_i)$ time as the difference from the previous output. The algorithm is extended to enumerate all binary trees in polynomial delay for a given leaf vector.

Optimizing over Path-Length Matrices of Unrooted Binary Trees – DANIELE CATANZARO, CORE UNIVERSITÉ CATHOLIQUE DE LOUVAIN –

Investigating the set of Path-Length Matrices (PLMs) induced by unrooted binary trees with n leaves remains an elusive challenge in combinatorial optimization and graph theory, persisting since its introduction in the 1970s. This characterization plays a pivotal role in exploring the optimization facets of diverse applications, spanning from epidemiology to data compression and information theory. Here, we address this longstanding problem by presenting the first known characterization for the PLM set, accompanied by a collection of valid inequalities and polyhedral results that contribute to understanding its convex hull.

Leveraging these findings, we delve into the Balanced Minimum Evolution Problem (BMEP), a NP-hard nonlinear optimization problem arising from the literature on computational phylogenetics. The study of the polyhedral combinatorics of the BMEP reveals strong connections with the symmetric Traveling Salesman polytope, and sheds light on the aspects related to the dyadic nature of its solutions.

By working in an extended space, our characterization leads to the smallest integer linear programming formulation known to date for the BMEP. By further embodying the facets and the strengthening valid inequalities derived earlier we obtain a branch-and-cut algorithm that significantly outperforms the best exact solution algorithm currently available in the literature.

Breaking the quadratic gap for strongly polynomial solvers to combinatorial linear programs – BENTO NATURA, GEORGIA TECH –

Recent years have seen tremendous progress in high-accuracy solvers for Maximum Flow, Minimum-Cost Flow and general Linear Programs (LP). Progress on strongly polynomial solvers for combinatorial LP on the other hand has stalled. For combinatorial LP beyond directed graphs this gap between exact and high-accuracy solvers is currently quadratic. We finally break the quadratic gap and design a strongly polynomial interior-point-method for combinatorial LP, which reduces the gap to only a linear factor.

A strongly polynomial algorithm for the minimum cost generalized flow problem – ZHUAN KHYE KOH, CENTRUM WISKUNDE & INFORMATICA –

We give a strongly polynomial algorithm for minimum cost generalized flow, and as a consequence, for all linear programs

with at most two variables per inequality. Previously, strongly polynomial algorithms were only known for the primal and dual feasibility problems. Our approach is to show that the path-following interior point method of Allamigeon et al. '22 terminates in a strongly polynomial number of iterations for minimum cost generalized flow. We achieve this by bounding the 'straight line complexity' of the central path, which is the minimum number of pieces required by a piecewise affine curve to multiplicatively approximate the central path.

Based on joint work with Daniel Dadush, Bento Natura, Neil Olver and László Végh.

2c - ThC355 - 511B - **Kernels in Graphs** Chair: FRÉDÉRIC MEUNIER ↔

Kernels and quasi-kernels – HÉLÈNE LANGLOIS, LAMA –

Kernels are defined as subsets of vertices which are independent and “absorbing” (every vertex outside the subset has at least one outneighbor in the subset). Initially introduced by von Neumann and Morgenstern for game theory, they have also found applications in economics and logic. The existence of kernels in a digraph is not systematic and is even NP-complete to decide. The first part of this presentation aims to explore the challenges and theoretical aspects of kernels. The notion of quasi-kernels, which was introduced by Chvátal and Lovász in 1974, derives from a slight modification in the definition of kernels. A quasi-kernel is a subset of vertices that is independent and such that there is a directed path of length at most two from every vertex to that subset. Any digraph has a quasi-kernel that can be found in polynomial time. The research about quasi-kernels focuses on a conjecture called the “small quasi-kernel conjecture.” It suggests that a sink-free digraph (where every vertex has a positive outdegree) has a quasi-kernel of size at most half of the vertex set. However, this conjecture is only confirmed for specific classes of digraphs. The second half of the presentation focuses on quasi-kernels, on their algorithmic aspect and strategies of proofs for the small quasi-kernel conjecture.

Extensions of classical results on kernels – FRÉDÉRIC MEUNIER, ÉCOLE DE PONTS – Hélène Langlois

In a directed graph, a kernel is a subset of vertices that is independent and absorbing. Kernels form a fundamental topic of the theory of directed graphs, and appear in a variety of contexts: logic, game theory, economics, and others. Not all graphs possess a kernel and deciding the existence of a kernel is NP-complete. Among the most fundamental results ensuring existence of a kernel are the famous theorems of Sands, Sauer, and Woodrow (1982) and of Galeana-Sánchez and Neumann-Lara (1984). In this talk, we will discuss extensions of these theorems and some challenges they pose.

Quasi-kernels in finite and countable infinite (?) directed graphs – PÉTER ERDOS, A. RÉNYI INSTITUTE OF MATHEMATICS, HUN REN –

A subset Q of the vertices of a directed graph is a quasi-kernel (QK for short) if it is independent and the shortest path to every vertex from Q is of length at most two. Chvátal and Lovász proved in 1974 that all digraphs in fact possess QK. Maybe it is characteristic, that almost any raised problem (like finding

a minimum QK) proved to be NP-hard. In this lecture I will survey in short the literature about QK.

Along the way I will talk about the small QK conjecture, which states that every source-free digraph has a QK of size at most half of its size (Erdős and Székely, 1976). There is a growing list of partial results, but the end seems to be far away. I also will mention the large quasi-kernel conjecture of Sam Spiro, 2024.

It is easy to see that the Chvátal-Lovász's theorem does not hold for countable infinite digraphs. Therefore, if the time's limit allows, I will talk about infinite quasi-kernel conjecture (Erdős-Soukup, 2008).

2c - FA351 - 511C - **Combinatorial Optimisation and Graph Theory** Chair: MARTIN DREES ↔

Rounding the Lovasz Theta Function with a Value Function Approximation – RUI GONG, GEORGIA INSTITUTE OF TECHNOLOGY – Cifuentes Diego, Alejandro Toriello

The Lovasz theta-function is a semidefinite programming (SDP) relaxation for the maximum stable set problem. In this paper, we develop a novel rounding scheme for the Lovasz theta-function, i.e., a procedure that extracts a stable set from the solution of the SDP. The main idea behind our scheme is to construct a value function approximation from the SDP solution, and then retrieve the stable set by using ideas from dynamic programming. We prove that our method recovers an optimal stable set in almost all perfect graphs. To the best of our knowledge, this is the only known rounding strategy for the Lovasz theta function that provably works on a large class of perfect graphs. Our rounding scheme relies on simple linear algebra computations; a single SDP is solved to construct the value function approximation. In contrast, previous methods for finding an optimal stable set in perfect graphs require solving multiple SDPs. Preliminary computational experiments show that our rounding strategy produces very good solutions for the stable set problem, even on non-perfect graphs.

Intersecting and Dense Restrictions of Clutters in Polynomial Time – MARTIN DREES, RESEARCH INSTITUTE OF DISCRETE MATHEMATICS, BONN –

A clutter is a family of sets, called members, such that no member contains another. It is called intersecting if every two members intersect, but not all members have a common element. Dense clutters additionally do not have a fractional packing of value 2. We are looking at certain substructures of clutters, namely minors and restrictions.

For a family of clutters we introduce a general sufficient condition such that for every clutter we can decide whether the clutter has a restriction in that set in polynomial time. It is known that the sets of intersecting and dense clutters satisfy this condition. For intersecting clutters we generalize the statement to k -wise intersecting clutters using a much simpler proof.

We also give a simplified proof that a dense clutter with no proper dense minor is either a delta or the blocker of an extended odd hole. This simplification reduces the running time of the algorithm for finding a delta or the blocker of an

extended odd hole minor from previously $O(n^4)$ to $O(n^3)$ filter oracle calls.

Optimizing Mediated Graphs: Applications in Conic Optimization and Representations of Circuits as Sums-Of-Squares – VÍCTOR BLANCO, UNIVERSIDAD DE GRANADA – Martínez Antón Miguel

We present here mathematical optimization based approaches to optimize the size of different types of mediated graphs. On the one hand, mediated graphs have been proven to be a fundamental tool to represent general p-order cones as second order cones. Given the parameters of a weighted geometric cone, minimizing the cardinality of a (continuous) mediated set for those parameters results in the minimal extended representation of the cone. In this case, we provide a novel Integer Linear Optimization formulation for this problem and provide some insights on the advantages of using our approach against the enumerative approaches that have been previously proposed. On the other hand, (discrete) mediated sets are useful in the determination of minimal representations of circuits by sos polynomials. For this problem, we derive a different and more challenging Integer Linear Optimization model for the problem and analyze different strategies to solve it.

Packing cycles in planar and bounded-genus graphs – NIKLAS SCHLOMBERG, UNIVERSITY OF BONN – Hanjo Thiele, Jens Vygen

We devise constant-factor approximation algorithms for finding as many disjoint cycles as possible from a certain family of cycles in a given planar or bounded-genus graph. The family of cycles under consideration must satisfy two properties: it must be uncrossable and allow for a certain oracle access. Our setting generalizes many problems that were studied separately in the past. For example, three families that satisfy the above properties are (i) all cycles in a directed or undirected graph, (ii) all odd cycles in an undirected graph, and (iii) all cycles in an undirected graph that contain precisely one demand edge, where the demand edges form a subset of the edge set. The latter family (iii) corresponds to the classical disjoint paths problem in fully planar and bounded-genus instances. Furthermore, we obtain approximate min-max theorems of the Erdos-Pósa type.

2c - FA925 - 511B - **Combinatorial Optimization and Graph Theory 5** Chair: AKIYOSHI SHIOURA

Labeling Methods for Partially Ordered Paths – RICARDO EULER, ZUSE INSTITUTE BERLIN – Pedro Maristany De Las Casas

The multi-objective shortest path problem (MOSP) is a well-studied combinatorial optimization problem that has recently seen significant methodological advances. It can easily be reinterpreted as optimizing over a partial order on the paths, namely the d-fold product order of the canonical total order on the real numbers. This raises the question whether labeling methods for MOSP can be generalized to find sets of efficient paths with regard to different partial orders. Previous attempts to answer this question were either limited to acyclic graphs or to specific applications, usually solved with modifications of Martins' algorithm. The structure that these

modifications exploit is often closely related but clouded by application-specific language. In this talk, we formalize the partial order shortest path problem (POSP), introduce generalized labeling methods for POSP and identify properties of POSP instances that guarantee the correctness of these methods. By basing our general label-setting method on the recently introduced Multi-Objective Dijkstra Algorithm, we improve on the best achieved theoretical complexity for all POSP variants that were previously solved using Martins' algorithm without additional effort. Finally, we provide a taxonomy that allows the reader to quickly discern whether their specific application of POSP is suitable for generalized labeling methods and show that several classical shortest path problems fit well within our framework.

Measuring Constituency Based On Minimum Campaign Distance And Application To Evaluate The Merger Of Electoral Districts – YUICHIRO MIYAMOTO, SOPHIA UNIVERSITY – Toshio Nemoto

The area and population of an electoral district are main indicators of its constituency size. They are not, however, indicators that can be used to limit the merger of electoral districts for Japan's upper house (House of Councilors). This study proposes the minimum campaign distance (MCD) as a new indicator for measuring constituency size and examines relevant issues based on actual measurements, with the aim of using this indicator to evaluate the merger of electoral districts. With characteristics different from those of area and population, the MCD is an indicator that can be used to limit the merger of electoral districts. Here, the campaign duration is evaluated based on the MCD required to cover the entire constituency. Also, the inverse relationship between the number of constituents to cover and the MCD is quantified based on the MCD required to cover a portion of the constituency. The relationship varies across prefectures, which could explain why campaign strategies differ among regions. Lastly, the MCD is used to evaluate the merger of electoral districts: efficiency in terms of travel distance resulting from a merger of electoral districts, the effect of a merger on the smaller district (prefecture), and the feasibility of a merger are formulated as attributes calculated based on the MCD in order to demonstrate a new method of quantitatively evaluating a merger. We also present the result of the computational experiments. Test instances are generated from the Japanese national census.

The Colorful Components Problems – MARTINA CERULLI, UNIVERSITY OF SALERNO –

We study the colorful components framework, historically motivated by comparative genomics in bioinformatics. From a graph-theoretic perspective, we aim to partition a vertex-colored graph into colorful components, namely, connected subgraphs whose vertices have different colors from others in the same component. We address this problem with three different objectives, according to which we: (i) minimize the number of removed edges (MOP problem); (ii) minimize the number of colorful components (MCC problem); or (iii) maximize the number of edges in the transitive closure of the output graph (MEC problem).

In social networks, where nodes represent individuals and edges represent connections (e.g., friendships, interactions), all three problems can be used to determine the most influential

connections linking distinct (colorful) and cohesive (connected) communities. By guaranteeing the colorfulness of the communities through the removal of specific edges, as in the MOP problem, a social network aims to mitigate the risk of echo chambers, where users predominantly interact with similar individuals. Maximizing the number of edges in the transitive closure in the MEC problem means promoting the propagation of information through the community, ensuring that a node can be reached by others through a sequence of edge. Instead, by minimizing the number of colorful components, the MCC problem helps prevent the isolation of users within small disconnected groups. When considering cyberspace networks of devices with various types of connections (e.g., permissions, data flow), identifying a subset of edges to remove while ensuring that the network remains in connected colorful components helps optimize the network's resilience to cyber threats.

We formulate MOP, MEC, and MCC problems as integer nonlinear problems. We linearize these formulations and enhance them with valid inequalities. Furthermore, we provide bounds on the maximum number of colorful components at optimality and exploit them to reduce the size of the models. The formulations are computationally evaluated to demonstrate the efficiency of the proposed approaches.

On the monotonicity in steepest descent algorithm for M-convex functions – AKIYOSHI SHIOURA, TOKYO INSTITUTE OF TECHNOLOGY –

M-convex function is a class of discrete convex function defined on the integer lattice points, and plays a primary role in the theory of discrete convex analysis. We consider the minimization of an M-convex function. It is known that a minimizer of an M-convex function can be found by a steepest descent algorithm in a finite number of iterations. In this talk, we present some monotonicity properties of the steepest descent algorithm for the class of M-convex functions and a more general class of discrete convex functions called jump M-convex functions. We show that the algorithm finds the “nearest” optimal solution to a given initial solution. We also analyze the trajectory of the solutions generated by the algorithm, and show that it is a “shortest” path from the initial solution to the “nearest” optimal solution. We also investigate the monotonicity in the slope of steepest descent directions.

2c - FB926 - 511B - **Combinatorial Optimization and Graph Theory 6** Chair: JOSEPH POREMBA

Bounding the Proper Orientation Number by Tree-depth – HIROTAKA ONO, NAGOYA UNIVERSITY – Yasuko Matsui, Takeaki Uno

A proper orientation of a simple undirected graph $G = (V, E)$ is an orientation D of E such that $d_D^-(u) \neq d_D^-(v)$ for every two adjacent vertices u and v , where $d_D^-(u)$ is the number of incoming arcs to u on D . The proper orientation number $\text{pon}(G)$ of G is the minimum of the maximum indegrees over all its proper orientations. It is known that a graph always has a proper orientation and that the proper orientation number is at most the maximum degree. The proper orientation number of a tree is known to be at most 4, and for a cactus, at most 7. On the other hand, it is known that computing the proper orientation number is NP-hard even for 4-regular graphs and planar graphs. This paper focuses on tree-depth $td(G)$ as a

graph structural parameter and shows that the proper orientation number is above bounded by $2(td - 1)$. We then give a $2^{O(td(G)^2)}n$ -time algorithm, where n is the number of vertices in G .

Beam Search to Minimise Cutting Patterns with Maximum Utilisation – CLAUDIA O LÓPEZ SOTO, NATIONAL AUTONOMOUS UNIVERSITY OF MEXICO (UNAM) – Marta Cabo

In many cutting and packing problems the sole objective is to maximise the utilisation of the cutting board, while satisfying the demand of pieces to be cut. When one focuses on optimising the use of raw material, it is very likely that for each board we will have to enter different cutting coordinates. In operations that involve a large number of pieces many boards will be required, so the cutting process can be highly affected when changing boards. When dealing with a heterogeneous mix of pieces, this problem cannot be avoided. However, when the set of pieces to be cut present large repetitions of the same pieces, it would seem reasonable to pay attention to the number of changes in the machine settings, reducing the cutting process time.

In this problem we introduce a new objective in the classic two-dimensional bin packing problem, where on top of maximising the total utilisation, we are also interested in minimising the number of different patterns needed to complete the demand. For this work, a pattern is defined as the coordinates of a set of pieces that will be cut from a bin. Minimising the number of patterns, will result in less changes on the settings of the cutting machine among bins, with a reduction on total cutting times. This variant of the problem has not been widely studied and so far we have only found the work by [?] which uses a column generation model to solve this problem.

To start tackling this problem, we expand the initial work in [?] and work with a similar setting based on the glass industry, with irregular pieces that can be freely rotated and reflected. We also start by considering the restriction of separating each piece by means of a guillotine cut, which reduces the solution space and somehow limit the placement of each piece. [?] present a beam search strategy where each node represents a complete bin, thus the branch with minimum depth is the one with less bins, or maximum utilisation. The same idea applies in this problem, interpreting the branch depth as the number of changes in machine settings.

We introduce some novelty approaches in terms of the sorting of pieces, considering not just the largest as the first ones to place, but also taking in consideration the number of repetitions of each piece. The fact that pieces are separated by guillotine cuts, translates into irregular shapes where to place pieces after a guillotine cut is performed, thus we also look at the similarity between the angles of the pieces and the angles of each section of the bin to decide which piece to place next. Preliminary test will be shown to demonstrate the efficiency of these new criteria for sorting pieces.

We also provide initial results that confirms that the strategy of dealing with both objectives simultaneously is more effective than using existing heuristics that are mainly design for heterogeneous mix of pieces.

3D Phase Unwrapping using 3D Surface minimization. – EL MEHDI OUDAOUD, POLYTECH-

NIQUE MONTREAL – Thibaut Vidal, Youssouf Emine
3D Phase Unwrapping is a pivotal step in analyzing MRI imagery, where traditional methods are often of a heuristic nature and suboptimal outcomes. Central to this challenge is the phenomenon of phase values wrapped between 0 and 2π , which can veil true phase differences and introduce inaccuracies, particularly in critical applications like medical imaging and materials science.

Our study introduces a novel approach that constructs a graph linking neighboring residues, enabling the rapid detection of "Singular Loops" by the use of the "Graph Untangling" method — surpassing the efficiency of existing state-of-the-art methods. We address these loops by identifying their minimal surfaces, and that by solving a discrete version of the "Plateau Problem". This involves the use of Integer Programming with constraint separation to compute these minimal surfaces. We here use cycle basis theories to optimize our constraints separation methods, aiming to optimize the unwrapping process.

Subsequently, we reconstruct the original phase while strategically avoiding these Singular Loops, thereby bypassing the identified optimal minimal surfaces. This methodology not only enhances the accuracy of phase reconstruction but also contributes to a deeper understanding of the complexities involved in 3D Phase Unwrapping. Our findings offer a significant advancement in the field, hopefully transforming practices in medical imaging analysis and beyond.

Uncrossable Multicommodity Flows – JOSEPH POREMBA, THE UNIVERSITY OF BRITISH COLUMBIA – Chandra Chekuri, Naves Guylain, Bruce Shepherd

We investigate uncrossed multiflows, multiflows on planar graphs with the property that the curves identified by their flow paths do not cross in the plane. We demonstrate that given a fractional such multiflow, it can be rounded to an integral multiflow of congestion 2, motivating them as an interesting combinatorial object. We explore the complexity of computing uncrossed multiflows. On the one hand we can prove that determining whether a fractional uncrossed multiflow exists is NP-hard. In contrast, an integral (uncrossed) multiflow can be computed efficiently when the number of faces with demands is bounded.

2d - MB248 - 513F - **Distance Optimization in Data Science** Chair: EMILIO CARRIZOSA

A Novel Column Generation Framework for One-for-many Counterfactual Explanations – JASONE RAMIREZ AYERBE, UNIVERSIDAD DE SEVILLA – Andrea Lodi

In this talk, we address the problem of generating a set of counterfactual explanations for a group of instances using the one-for-many allocation rule. We aim to minimize the number of explanations necessary to explain all instances while also considering sparsity by restricting the collective perturbed features within each explanation. We introduce a novel column generation methodology that is adaptable to diverse black-box classifiers, including neural networks. Through a comparison with a straightforward adaptation of a mixed-integer formulation from the literature, we demonstrate the superiority of our column generation approach in terms of scalability, computational efficiency, and solution quality.

Solving the Contextual Multiobjective Inverse Ideal Point Problem via Mathematical Optimization – NURIA GÓMEZ VARGAS, UNIVERSITY OF SEVILLE – Rafael Blanquero, Emilio Carrizosa

We propose a Mathematical Optimization model tailored to address the challenging Contextual Multiobjective Inverse Ideal Point Problem (CMIIPP). The CMIIPP involves yielding the underlying objective functions of a multicriteria decision-making model based solely on a set of pairs of concurrently observed context and decisions, particularly focusing on the ideal point scenario in multiobjective optimization. By formulating the CMIIPP as an optimization problem, our approach offers a framework to uncover the implicit preferences governing the decision-making process that led to a given ideal point solution preserving its consistency. Furthermore, considerations are made for incorporating domain-specific knowledge to enhance applicability and exploring other requirements such as interpretability.

Distance-Based Fairness in Classification and Regression – THOMAS HALSKOV, COPENHAGEN BUSINESS SCHOOL – Emilio Carrizosa, Dolores Romero Morales

In this presentation we introduce a novel approach to integrate fairness considerations into regression and classification models using mathematical programming techniques. We consider a setting with a sensitive and non-sensitive group defined by some attribute (e.g. race or gender), and wish to make the predictions from both groups similar through a distance. Our approach is applicable to well known methodologies such as LASSO or Support Vector Machines (SVM). In LASSO this is achieved by choosing a decision threshold on the predictions and considering the distance between the two groups above the chosen threshold. For SVM it is done by incorporating a penalty based on the Wasserstein Distance to measure the disparity in score distributions between the sensitive and non-sensitive groups. We show that this can be formulated as the original SVM problem with a fairer kernel. To solve the problem we implement an alternating algorithm. Both of these models are illustrated on real-world datasets, where we are able to significantly improve fairness while maintaining a high accuracy.

Sequential Counterfactual Decisions – EMILIO CARRIZOSA, UNIVERSIDAD DE SEVILLA – Jasone Ramirez Ayerbe, Dolores Romero Morales

Given a probabilistic classifier in a binary classification problem, a counterfactual solution to a record is a feasible solution, close to the record and having a high probability of being labeled in the positive class by the classifier. Finding a counterfactual solution for a given record amounts to finding a Pareto-optimal solution to the problem of simultaneous maximization of closeness as well as the probability of classification in the positive class.

In this talk, we address the problem where, instead of one counterfactual solution, we seek a counterfactual solutions plan, i.e., a sequence of T solutions with increasing probabilities of positive and minimizing (a function of) the overall distances. The so-obtained optimization problems will be analyzed for different choices of the classifier and metrics induced by norms, or, in their asymmetric version, gauges.

2d - MC298 - 513A - **Classification and regression via mathematical optimization** Chair: DOLORES ROMERO MORALES

Soft regression trees: a model variant and a decomposition training algorithm – ANTONIO CONSOLO, POLITECNICO DI MILANO – Edoardo Amaldi, Andrea Manno

Decision trees are popular supervised Machine Learning (ML) methods used for classification and regression tasks arising in a variety of application fields. Unlike most black-box ML models, they reveal the sequence of decisions leading to the tree response for any input vector. Interpretability is of particular importance in applications where the ML models support the decisions of the domain experts and justifiable predictions are required, such as for instance in medical diagnosis. In the seminal CART method [3] for building univariate classification and regression trees and the later variants C4.5 and ID3, a top-down and greedy approach is adopted where at each branch node a split is determined by minimizing a local impurity measure. Since training decision trees is known to be NP-hard, it is a challenging problem to globally optimize them over all the tree parameters. During the past decade, growing attention has been devoted to globally optimized (trained) decision trees with deterministic or soft splitting rules at branch nodes, leveraging on the remarkable advances in mixed integer optimization and nonlinear optimization. We propose a variant of soft multivariate regression trees (SRTs) where, for every input vector, a potential linear prediction is available at each leaf node (as a linear regression of the features) but the actual tree prediction is the one associated to the single leaf node obtained by routing the input vector from the root along the branches with the higher probabilities. Our nonlinear optimization formulation for training such soft trees is well-suited to decomposition and to impose fairness constraints. After showing a universal approximation result for SRTs, we present a node-based decomposition training algorithm which includes a clustering-based initialization procedure and a heuristic for reassigning the data points along the tree. Under mild assumptions, we also establish asymptotic convergence guarantees. Experiments on 15 data sets from the UCI and KEEL repositories indicate that our SRT model variant and decomposition algorithm yield improved testing accuracy compared with the soft regression trees discussed in [2], and significant speed-up in training time and similar testing accuracy compared with the mixed integer optimization approach described in [1]. Other experiments on the same datasets show that our SRT approach is significantly more robust than classical Hierarchical Mixture of Experts trained with the Expectation-Maximization algorithm and provides single soft trees whose testing accuracy is comparable to that of Random Forest.

References

- [1] Bertsimas, D. and Dunn, J.: Machine learning under a modern optimization lens. Dynamic Ideas LLC (2019).
- [2] Blanquero, R., Carrizosa, E., Molero-Río, C. and Romero Morales, D.: On sparse optimal regression trees. European Journal of Operational Research, 299(3), 1045-1054 (2022).
- [3] Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and regression trees. CRC press (1984).
- [4] Irsoy, O., Yildiz, O.T., Alpaydin, E.: Soft decision trees. Proceedings

of the 21st International Conference on Pattern Recognition (ICPR2012), 1819–1822 (2012).

Mixed-Integer Linear Optimization for Cardinality-Constrained Random Forests – MARIA EDUARDA PINHEIRO, TRIER UNIVERSITY – Jan Pablo Burgard, Martin Schmidt

Random forests are among the most famous algorithms for solving classification problems, mainly for large-scale data sets. Considering a set of labeled points and several decision trees, the method takes the majority vote to classify a new given point. In some scenarios, however, labels are only accessible for a proper subset of the given points. Moreover, this subset can be non-representative, e.g., due to collection bias. Semi-supervised learning considers the setting of labeled and unlabeled data and often improves the reliability of the results. In addition, it can be possible to obtain additional information about class sizes from undisclosed sources. We propose a mixed-integer linear optimization model for computing a semi-supervised random forest that covers the setting of labeled and unlabeled data points as well as the overall number of points in each class for a binary classification. Since the solution time rapidly grows as the number of variables increases, we present some preprocessing techniques and an intuitive branching priority rule to reduce the model's size. Our numerical results show that our approach leads to a better accuracy and a better Matthews correlation coefficient for biased samples compared to random forests by majority vote, even if only few labeled points are available.

On Enhancing the Explainability and Fairness of Tree Ensembles – DOLORES ROMERO MORALES, COPENHAGEN BUSINESS SCHOOL – Emilio Carrizosa, Kurishchenko Kseniia

Tree ensembles are one of the most powerful methodologies in Machine Learning. In this talk, we investigate how to make tree ensembles more flexible to incorporate by design explainability and fairness. While explainability helps the user understand the key features that play a role in the classification task, with fairness we ensure that the ensemble does not discriminate against a group of observations that share a sensitive attribute. We propose a Mixed Integer Linear Optimization formulation to train an ensemble of trees that apart from minimizing the misclassification error, controls for sparsity as well as the accuracy in the sensitive group. Our formulation is scalable in the number of observations since its number of binary decision variables is independent of the number of observations. In our numerical results, we show that for standard datasets used in the fairness literature, we can dramatically enhance the fairness of the benchmark, namely the popular Random Forest, while using only a few features, all without damaging the misclassification error.

2d - TA297 - 513A - **Optimization for Safe and Private Machine Learning** Chair: THIBAUT VIDAL

A Framework for Data-driven Explainability in Mathematical Optimization – FRAUKE LIERS, FRIEDRICH ALEXANDER UNIVERSITÄT ERLANGEN NÜRNBERG – Kevin Martin Aigner, Marc Goeigk, Michael Hartisch, Arthug Miehlich

Advancements in mathematical programming have made it possible to efficiently tackle large-scale real-world problems that were deemed intractable just a few decades ago. However, provably optimal solutions may not be accepted due to the perception of optimization software as a black box. Although well understood by scientists, this lacks easy accessibility for practitioners. Hence, we advocate for introducing the explainability of a solution as another evaluation criterion, next to its objective value, which enables us to find trade-off solutions between these two criteria. Explainability is attained by comparing against (not necessarily optimal) solutions that were implemented in similar situations in the past. Thus, solutions are preferred that exhibit similar features. Although we prove that already in simple cases the explainable model is NP-hard, we characterize relevant polynomially solvable cases such as the explainable shortest path problem. Our numerical experiments on both artificial as well as real-world road networks show the resulting Pareto front. It turns out that the cost of enforcing explainability can be very small.

Get Off The Hype Train: Why Using Neural Networks Can Be Very Risky – ROBERT BASSETT, NAVAL POSTGRADUATE SCHOOL – Anthony Austin, Austin Van Dellen

Adversarial manipulations show that many state-of-the-art neural network models suffer from a catastrophic lack of robustness. In this talk, we investigate how a capable adversary can exploit this weakness to disrupt autonomous acoustic sensors. We take the perspective of the adversary, formulating a partial differential equation (PDE) constrained optimization problem that constructs an acoustic signal which optimally disrupts a defender's classifier. The primary novelty in our formulation is the PDE constraint, which connects the acoustic signal emitted by the adversary to the signal received by the classifier. By incorporating this PDE constraint, our formulation models an adversary which can make changes that impact the ambient acoustic environment but cannot directly manipulate inputs into the classifier. After discussing the problem formulation, we introduce new methods to compute solutions at scale. We conclude by summarizing the implications of our results on applications relevant to the US Navy.

OR with a White Hat : Evidencing Privacy Vulnerabilities in ML Models – THIBAUT VIDAL, POLYTECHNIQUE MONTRÉAL – Julien Ferry, Ricardo Fukasawa, Timothée Pascal

We present an optimization-based reconstruction attack that can effectively recover a dataset used to train a random forest model. Our approach utilizes information readily provided by random forests trained with scikit-learn. We frame the dataset reconstruction problem as a combinatorial optimization problem with a maximum likelihood objective and rely on constraint programming to solve it. Our extensive experiments reveal that random forests trained with feature randomization but without bootstrap aggregation are highly vulnerable to complete dataset reconstruction, even with very few trees. When bootstrap aggregation is used, most of the data can still be reconstructed. These findings underscore a critical vulnerability inherent in widely adopted ensemble methods, warranting attention and mitigation.

SMLE: Safe Machine Learning via Embedded Overapproximation – MATTEO FRANCOBALDI,

UNIVERSITY OF BOLOGNA – Michele Lombardi

During the last decade Machine Learning systems, in particular Neural Networks, have witnessed a tremendous impact over a wide range of domains, becoming a pervasive technology in our society. Despite their success, however, they still lack formal guarantees on their behavior, representing crucial requirements for their adoption in safety-critical scenarios, such as autonomous driving or aircraft collision avoidance. This is why a wide range of frameworks have been proposed to formally verify safety properties in state-of-the-art AI systems. These methods essentially rely either on Optimization and Searching or on Reachability Analysis. The former perform exact verification in a declarative fashion, suited also for certifying hybrid symbolic and subsymbolic systems, but usually fail to scale to real-world use cases. The latter adopt algorithmic approaches able to offer increased tractability, but often at the price of sacrificed completeness. In this work, we contribute to bridge the gap between these two paradigms by introducing a novel framework, called SMLE (Safe Machine Learning via Embedded overapproximation), and consisting in equipping the learning model with an independent trainable overapproximator, which can be much simpler than the main model, given that its complexity is completely customizable by the user. Therefore, this auxiliary model can serve as a surrogate to facilitate the verification of safety properties, either in post-hoc phase, or immediately at training time via Projected Gradient Descent or Penalty Methods, hence enabling the construction of safe-by-design systems.

2d - TB199 - 513A - **Learning to Select Chair:**
VASSILIS DIGALAKIS 

Minimizing difference of submodular functions via DC programming – GEORGE ORFANIDES, MCGILL UNIVERSITY – Marwa El Halabi, Tim Hoheisel

Many machine learning problems can be naturally expressed as minimizing the difference of two submodular (DS) functions. Via the Lovász extension, it is well known that minimizing a DS function can be equivalently formulated as minimizing the difference of two convex (DC) functions. While DS minimization has developed theory and algorithms in its own right, it has yet to fully exploit the rich theory and effective algorithms offered by DC programming. To this end, we introduce variants of the well-known DCA (difference of convex functions algorithm) which can be applied to DS minimization. Our results using DCA match the theoretical guarantees satisfied by existing DS algorithms, while providing a more complete characterization of convergence properties. Moreover, numerical results show that our proposed algorithms outperform existing baselines on two applications: speech corpus selection and feature selection.

Slowly varying regression under sparsity – VASSILIS DIGALAKIS, HEC PARIS – Michael Li, Dimitris Bertsimas

I will present the framework of slowly varying regression under sparsity, allowing sparse regression models to exhibit slow and sparse variations, through an application in energy consumption prediction. First, I will formulate the problem of parameter estimation as a mixed-integer optimization problem; then, I will demonstrate that it can be precisely reformulated as a binary convex optimization problem through a

novel relaxation technique, convexifying the non-convex objective function while matching the original objective on all feasible binary points. I will develop a highly optimized implementation of a cutting plane-type algorithm, a fast regularization-based heuristic method that guarantees a feasible solution, and a practical hyperparameter tuning procedure relying on binary search that, under certain assumptions, is guaranteed to recover the true model parameters.

Learning to Cover: Online Learning and Optimization with Irreversible Decisions – MICHAEL LI, HARVARD BUSINESS SCHOOL – Alexandre Jacquillat

We define an online learning and optimization problem that aims to achieve a target coverage level by the end of a finite horizon. At each period, a decision-maker can open new facilities, receives information on the success of each one, and can then update a classification machine learning model to guide future facility location decisions. Irreversible facility opening decisions give rise to an exploration-exploitation trade-off with commitment. We derive an optimal algorithm and a tight lower bound in an asymptotic regime characterized by a large target number of facilities $m \rightarrow \infty$ but a fixed planning horizon $T \in \mathbb{Z}_+$. We find that the regret grows sub-linearly at a rate $\Theta\left(m^{\frac{1}{2} \cdot \frac{1}{1-2^{-T}}}\right)$, and therefore converges exponentially fast to $\Theta(\sqrt{m})$ as T increases. These results highlight the benefits of even a few rounds of online learning and optimization, as compared to a no-learning baseline that achieves linear regret. We establish the robustness of these insights in settings with offline data, with imperfect learning, with continuous uncertainty on the performance of each facility, and with heterogeneous facilities in a networked environment. We conclude with a case study in sustainable infrastructure planning using real-world data to show the benefits of the proposed online learning and optimization approach in practice.

Optimal Experimentation for Learning Personalized Policies Across Locations – STEFANOS POULIDIS, INSEAD – Georgina Hall, Spyros Zoumpoulis

Firms wish to learn personalized policies for customers in heterogeneous yet related locations or markets to maximize their monetary gains. To do this, they conduct experiments at each location to estimate the parameters of a customer response function. A crucial decision is which action to assign to each participant in the experiment, especially when a participant can only be assigned one action or there are budget constraints. The existing experimentation methodology considers locations and experiments individually. In this work, we leverage the relationship between locations in the experimentation problem to learn more profitable policies by proposing novel estimators and a semidefinite programming approach.

2d - TC289 - 513A - **On Machine Learning for Mathematical Optimization Chair:** ANDREA LODI

↩

A Data-Driven Approach to Cut Generating Function Selection – HONGYU CHENG, JOHNS HOPKINS UNIVERSITY – Amitabh Basu

Selecting an effective cutting plane to minimize the branch-and-cut tree size is a critical challenge in the branch-and-cut

algorithm. Recent advancements have employed a data-driven approach to select optimal cutting planes from a parameterized family, aimed at reducing the branch-and-bound tree size for a given distribution of integer programming instances. We extend this idea to the selection of the best cut generating function (CGF) from a parameterized family of facet-defining valid functions for the infinite group relaxation problem, with Gomory's mixed-integer (GMI) cuts as a particular instance. We provide rigorous sample complexity bounds for the selection of an effective CGF from these families for any specified distribution under some mild assumptions. Our empirical results show that the selected CGF can outperform the best GMI cut for certain synthetic distributions. Additionally, we explore the sample complexity of using neural networks for instance-dependent CGF selection.

Variable aggregation in a Benders decomposition for the p-median problem – RICK WILLEMSEN, ERASMUS UNIVERSITY ROTTERDAM – Daniel Aloise, Raf Jans

The p-median problem is a classical discrete location problem and is equivalent to the well-known k-medoids problem in the unsupervised clustering literature. The aim is to select p centers while minimizing the sum of distances from each customer to its nearest center. Recent advancements in solving the p-median and related problems have successfully leveraged Benders decomposition methods. In order to reduce the number of variables and possibly the number of Benders cuts in these models, it is possible to aggregate distance variables corresponding to customers. We propose to partially aggregate the distance variables based on an initial solution: aggregation occurs only when the corresponding customers are assigned to the same center in the initial solution. In addition, we propose a set of tailored valid inequalities for these aggregated variables. Initial experiments indicate that our model, post-initialization, provides a stronger lower bound, thereby accelerating the resolution of the root node. Furthermore, this approach seems to positively impact the branching procedure, leading to an overall faster Benders decomposition method.

Interactive learning of a knapsack constraint using a membership oracle – ROSARIO MESANA, UNIVERSITÀ DEGLI STUDI DI MILANO – Rui Chen, Andrea Lodi

We consider a variant of the knapsack problem in which the knapsack constraint is unknown and the decision maker only has access to a membership oracle that, given a solution, determines whether the solution is feasible or infeasible with absolute certainty. The goal of the decision maker is to find a nearly-optimal feasible solution and a knapsack constraint that, used in place of the unknown one in the knapsack model, yields such a solution. We assume that the oracle can be queried no more than a pre-defined number of times.

The problem is of interest both in the context of pure linear constraint interactive learning and for its applications in business and educational settings in which respecting a limit on the query budget and ensuring interpretability of the solution are critical concerns.

In our work we tackle the problem relying on an algorithmic structure inspired by SVM-based active learning. In particular, we show how to adapt from literature an algorithm that iteratively applies SVM for linear separation and Simple

Margin as a query strategy. We propose two alternative approaches for linear separation, the first one based on finding the Chebyshev center of the version space relative to the candidate separating hyperplanes and the other adapted from an algorithm for interactive convex optimization based on separation oracle. In addition, we introduce an alternative query strategy also based on the notion of version space.

We show how different separation methods and query strategies influence the quality of the result, in terms of objective value and constraint approximation. As a variant of the pure knapsack problem, we consider the composition of a college study plan in which the unknown constraint models time budgeting and the objective is to maximize the number of acquired credits. The student acts as the oracle and interacts with a system that proposes combinations of courses to attend that respect known dependencies, e.g. prerequisites. We conduct additional experiments to show how the proposed algorithms perform in this setting.

2d - WA927 - 513A - Machine Learning and Discrete Optimization 1 Chair: YASSINE YAAKOUBI ←

Gurobi Machine Learning – ROLAND WUNDERLING, GUROBI – Pierre Bonami

Gurobi-machinelearning is an open-source python package to formulate trained regression models in a gurobipy model that can then be solved with the Gurobi solver. We will present this package with a focus on trained neural networks. Representing such as constraints in a MIP leads to large models. Optimization based bound tightening (OBBT) can be used to improve performance for solving such models. We will discuss how Gurobi has been extended to perform OBBT tailored to NN models and present performance results.

Wasserstein Distributionally Robust Shallow Convex Neural Networks – JULIEN PALLAGE, POLYTECHNIQUE MONTRÉAL – Antoine Lesage Landry

Neural networks have been traditionally overlooked in critical sectors, e.g., energy. Even though they tend to be successful nonlinear predictors, their lack of interpretability, limited performance guarantees, complex training procedures as well as high susceptibility to data corruption and other adversarial attacks have labelled them as unreliable. Many solutions have been proposed in the past decade to guide and frame their training as much as possible, e.g., hyperparameter optimization, adversarial defense procedures, post-training verification frameworks, and tight MLOps lifecycle management. Still, these advances add up to training complexity and don't scale easily. In this work, we leverage recent results from distributionally robust optimization and convex learning to propose a new Wasserstein distributionally robust shallow convex neural network (WaDiRo-SCNN) with out-of-sample performance guarantees and simple low-stochasticity training. The training is formulated as a convex optimization problem efficiently solvable with open-source solvers. We first benchmark our model with a numerical study in a controlled synthetic environment. Lastly, we apply our approach to distribution feeder-specific impact prediction of commercial buildings enrolled in a demand response program

A machine learning approach for neighbor generation in metaheuristic search – DEFENG LIU, POLYTECHNIQUE MONTREAL – Andrea Lodi, Alain Hertz, Vincent Perreault

This paper presents a methodology for integrating machine learning techniques into metaheuristics for solving combinatorial optimization problems. Namely, we propose a general machine learning framework for neighbor generation in metaheuristic search. We first define an efficient neighborhood structure constructed by applying a transformation to a selected subset of variables from the current solution. Then, the key of the proposed methodology is to generate promising neighbors by selecting a proper subset of variables that contains a descent of the objective in the solution space. To learn a good variable selection strategy, we formulate the problem as a classification task that exploits structural information from the characteristics of the problem and from high-quality solutions. We validate our methodology on two metaheuristic applications: a Tabu Search scheme for solving a Wireless Network Optimization problem and a Large Neighborhood Search heuristic for solving Mixed-Integer Programs. The experimental results show that our approach is able to achieve a satisfactory trade-offs between the exploration of a larger solution space and the exploitation of high-quality solution regions on both applications.

Machine Learning for Distributionally Robust Warm-Starting in Mineral Supply/Value Chains – YASSINE YAAKOUBI, MCGILL UNIVERSITY – Roussos Dimitrakopoulos

Optimizing industrial mining complexes (mineral supply/value chains), from extraction to end-product delivery, presents significant challenges due to non-linearities and multiple sources of uncertainty, including in supply and demand. The two-stage stochastic integer program results in formulations with tens of millions of variables and nonlinear constraints, challenging the computational limits of state-of-the-art solvers. To address this computational complexity and streamline the optimization process, a novel distributionally robust warm-starting framework is introduced.

Leveraging historical solutions, a graph-based predictive model is trained to learn connectivity probabilities of mining blocks. Subsequently, resource-constrained min-cost max-flow problems are solved to establish initial production schedules that adhere to operational constraints and capture uncertainties related to both the optimization problem and the predictive model. An ambiguity set, constructed from historical connectivities, shifts the warm-start from a deterministic single point estimate -common in mining- to a distributionally robust optimization approach.

The framework captures the spatial distribution and inherent uncertainty of mining operations, yielding a diverse solution set for different scenarios and/or a singular risk-aware solution informed by block connectivity probabilities. Theoretical analysis and computational experiments demonstrate significant computational efficiencies and enhanced decision-making robustness. We discuss these findings, potential limitations, and future research directions for developing robust, reasoning, and responsible decision support systems for large-scale industrial applications under uncertainty.

2d - WB230 - 513A - **Combinatorial Optimization Augmented Machine Learning and Decision Aware Learning** Chair: AXEL PARMENTIER

Learning to solve large-scale districting problems – CHEIKH AHMED, ECOLE POLYTECHNIQUE MONTREAL –

Districting is a complex combinatorial problem that consists in partitioning a geographical area into small districts. In logistics, it is a major strategic decision determining operating costs for several years. Solving them using traditional methods is intractable even for small geographical areas and existing heuristics, while quick, often provide sub-optimal results. We present a structured learning approach to find high-quality solutions to real-world districting problems in a few minutes. It is based on integrating a combinatorial optimization layer, the capacitated minimum spanning tree problem, into a graph neural network architecture. To train this pipeline in a decision-aware fashion, we show how to construct target solutions embedded in a suitable space and learn from target solutions. Experiments show that our approach outperforms existing methods and can reduce costs by up to 15% on real-world cities.

Surrogate policies with generalization guarantees for combinatorial optimization problems – AXEL PARMENTIER, CERMICS, ECOLE DES PONTS – Pierre Cyril Aubin, Yohann De Castro, Alessandro Rudi

Consider a combinatorial optimization problem whose objective is given by a moderately expensive oracle that models complex phenomena. Operations Research practitioners generally ignore distributional information over the instances. To be practical, an algorithm must therefore scale well on any large instance of the problem. Practitioners usually solve a surrogate problem with an oversimplified objective that is amenable to classic combinatorial optimization techniques. We take instead a learning perspective. We assume that we have access to a training set of instances coming from an unknown distribution. And we restrict ourselves to algorithms that chain a neural network with a linear combinatorial optimization oracle. The latter is used as a tool to explore the combinatorially large set of feasible solutions efficiently. We use the training set to learn the parameters of the neural network that leads to the best performance. Minimizing the empirical risk on such a model class is difficult as the risk is piecewise constant. We therefore introduce a perturbation approach, that makes the empirical risk minimization problem tractable using exact algorithms. This allows us to derive an upper-bound on the non-optimality of the learned. This bound vanishes in the large training set regime. It is based on an ad-hoc statistical learning tools that exploit the geometry implied by the linear combinatorial optimization layer and the perturbation.

Neur2BiLO: Neural Bilevel Optimization – ESTHER JULIEN, TU DELFT –

Bilevel optimization deals with nested problems in which a leader takes the first decision to minimize their objective function while accounting for a follower best-response reaction. Constrained bilevel problems with integer variables are particularly notorious for their hardness. While exact solvers have been proposed for mixed-integer linear bilevel optimization,

they tend to scale poorly with problem size and are hard to generalize to the non-linear case. On the other hand, problem-specific algorithms (exact and heuristic) are limited in scope. Under a data-driven setting in which similar instances of a bilevel problem are solved routinely, our proposed framework, Neur2BiLO, embeds a neural network approximation of the leader's or follower's value function, trained via supervised regression, into an easy-to-solve mixed-integer program. Neur2BiLO serves as a heuristic that produces high-quality solutions extremely fast for a variety of bilevel problems spanning linear/non-linear, integer/mixed-integer knapsack problems, and four application domains.

From Large to Small Datasets: Size Generalization for Clustering Algorithm Selection – ELLEN VITERCIK, STANFORD –

In clustering algorithm selection, we are given a massive dataset and must efficiently select which clustering algorithm to use. We study this problem in a semi-supervised setting, with an unknown ground-truth clustering that we can only access through expensive oracle queries. Ideally, the clustering algorithm's output will be structurally close to the ground truth. We approach this problem by introducing a notion of size generalization for clustering algorithm accuracy. We identify conditions under which we can (1) subsample the massive clustering instance, (2) evaluate a set of candidate algorithms on the smaller instance, and (3) guarantee that the algorithm with the best accuracy on the small instance will have the best accuracy on the original big instance. We verify these findings both theoretically and empirically.

This is joint work with Vaggos Chatziafratis and Ishani Karmarkar.

2d - WC203 - 513A - **Machine Learning and Combinatorial Optimization** Chair: BAR-TOLOMEIO STELLATO

Valid Inequalities for Verification of Binarized Neural Networks – JIM LUEDTKE, UNIVERSITY OF WISCONSIN MADISON – Woojin Kim

Binarized neural networks (BNNs) are neural networks in which the weights are binary and the activation functions are the sign function. Verification of BNNs against input perturbation is one way to measure robustness of BNNs. BNN verification can be formulated as a mixed integer linear optimization problem. The natural formulation is often difficult to solve due to large integrality gap induced by big-M constraints. We present simple but effective techniques for generating valid inequalities for this formulation that exploit the recursive structure of the network. We find that our techniques enable verifying BNNs against a higher range of input perturbation than existing approaches.

Slimming down and beefing up: MILP Multi-Objective Optimization and Training for Few-Bit Neural Networks – SIMONE MILANESI, UNIVERSITY OF PAVIA – Ambrogio Maria Bernardelli, Stefano Gualandi, Neil Yorke Smith, Hoong Chuin Lau

Combinatorial optimization solvers have gained attention for training neural networks, particularly in scenarios with limited data availability. Employing sophisticated solvers such

as mixed integer linear programming enables precise network training without reliance on computationally intensive GPU-based methods. Our focus is on few-bit discrete-valued neural networks, encompassing architectures like Binarized Neural Networks (BNNs) and Integer Neural Networks (INNs). These lightweight models are distinguished by their capacity to function effectively on low-power devices, often implemented using boolean operations. Our proposed methodology entails training a network for each class pair within the classification problem and employing a majority voting scheme for output prediction. The optimization process adheres to principles of robustness and sparsity through a multi-objective function. We evaluate this approach, named BeMi, against existing solver-based and gradient-based techniques, particularly in the context of few-shot learning with BNNs. Additionally, we analyze trade-offs between INNs and BNNs. Empirical results on the MNIST dataset demonstrate that BeMi achieves higher accuracy (up to 81.8%) with a notable reduction in active weights, surpassing previous methodologies.

Differentiable Cutting-plane Layers for Mixed-integer Linear Optimization – BARTOLOMEO STELLATO, PRINCETON UNIVERSITY – Gabriele Dragotto, Stefan Clarke, Jaime Fernández Fisac We consider the problem of optimizing a sequence of mixed-integer linear optimization problems with varying parameters. By combining recent advances in cutting plane generation and differentiation through convex optimization problems, we construct a new differentiable architecture to predict the optimal cutting planes from the key parameters of each instance. During the offline phase, our method maximizes the cutting planes’ efficiency by evaluating the derivative of the solution of the continuous relaxations with respect to the cut-generating parameters. We show on preliminary computational results that, once trained, our architecture computes solutions with low infeasibility and suboptimality with fast and predictable execution times.

2d - ThA211 - 513A - Explainable Artificial Intelligence and Mathematical Programming
Chair: JANNIS KURTZ

ODTlearn: A Python Package for Learning Optimal Decision Trees for Prediction and Prescription – PATRICK VOSSLER, STANFORD UNIVERSITY – Sina Aghaei, Nathan Justin, Nathanael Jo, Andres Gomez, Phebe Vayanos ODTLearn is an open-source Python package that provides methods for learning optimal decision trees for high-stakes predictive and prescriptive tasks based on the mixed-integer optimization (MIO) framework proposed in (Aghaei et al., 2019) and several of its extensions. The current version of the package provides implementations for learning optimal classification trees, optimal fair classification trees, optimal classification trees robust to distribution shifts, and optimal prescriptive trees from observational data. We have designed the package to be easy to maintain and extend as new optimal decision tree problem classes, reformulation strategies, and solution algorithms are introduced. To this end, the package follows object-oriented design principles and supports both commercial (Gurobi) and open source (COIN-OR branch and cut) solvers. The package documentation

and an extensive user guide can be found at <https://d3m-research-group.github.io/odtlearn/>. Additionally, users can view the package source code and submit feature requests and bug reports by visiting <https://github.com/D3M-Research-Group/odtlearn>.

Finding Regions of Counterfactual Explanations via Robust Optimization – ILKER BIRBIL, UNIVERSITY OF AMSTERDAM – Donato Maragno, Jan-nis Kurtz, Tabea Röber, Rob Goedhart, Dick Den Hertog Counterfactual explanations play an important role in detecting bias and improving the explainability of data-driven classification models. A counterfactual explanation (CE) is a minimal perturbed data point for which the decision of the model changes. Most of the existing methods can only provide one CE, which may not be achievable for the user. In this work we derive an iterative method to calculate robust CEs, i.e., CEs that remain valid even after the features are slightly perturbed. To this end, our method provides a whole region of CEs allowing the user to choose a suitable recourse to obtain a desired outcome. We use algorithmic ideas from robust optimization and prove convergence results for the most common machine learning methods including decision trees, tree ensembles, and neural networks. Our experiments show that our method can efficiently generate globally optimal robust CEs for a variety of common data sets and classification models.

Learning Optimal Prescriptive Trees from Observational Data – NATHANAEL JO, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Sina Aghaei, Andres Gomez, Phebe Vayanos We consider the problem of learning an optimal prescriptive tree (i.e., an interpretable treatment assignment policy in the form of a binary tree) of moderate depth, from observational data. This problem arises in numerous socially important domains such as public health and personalized medicine, where interpretable and data-driven interventions are sought based on data gathered in deployment – through passive collection of data – rather than from randomized trials. We propose a method for learning optimal prescriptive trees using mixed-integer optimization (MIO) technology. We show that under mild conditions our method is asymptotically exact in the sense that it converges to an optimal out-of-sample treatment assignment policy as the number of historical data samples tends to infinity. Contrary to existing literature, our approach: 1) does not require data to be randomized, 2) does not impose stringent assumptions on the learned trees, and 3) has the ability to model domain specific constraints. Through extensive computational experiments, we demonstrate that our asymptotic guarantees translate to significant performance improvements in finite samples, as well as showcase our uniquely flexible modeling power by incorporating budget and fairness constraints.

Counterfactual Explanations for Linear Optimization – JANNIS KURTZ, UNIVERSITY OF AMSTERDAM – Ilker Birbil, Dick Den Hertog In recent years, there has been a rising demand for transparent and explainable machine learning (ML) models. A large stream of works focuses on algorithmic methods to derive so called counterfactual explanations (CE). Although significant progress has been made in generating CEs for ML models, this topic has received minimal attention in the Operations Re-

search (OR) community. However, algorithmic decisions in OR are made by complex algorithms which cannot be considered to be explainable or transparent. In this work we argue that there exist many OR applications where counterfactual explanations are needed. We translate the concept of CEs into the world of linear optimization problems and define three different classes of CEs: strong, weak and relative counterfactual explanations. For all three types we derive problem formulations and analyze the structure of it. We show that the weak and strong CE formulations have some undesirable properties while relative CEs can be derived by solving a convex optimization problem. We test all concepts on a real-world diet problem and we show that relative CEs can be calculated efficiently on NETLIB instances.

2d - ThB200 - 513A - Scaling Sparse Machine Learning Chair: VASSILIS DIGALAKIS 

Projection onto a Capped Rotated Second-Order Cone with Applications to Sparse Regression Relaxations – ISHY ZAGDOUN, BAR ILAN UNIVERSITY – Noam Goldberg, Jonathan Eckstein

Motivated by solving the continuous relaxation of mixed-integer nonlinear programs (MINLPs) with binary indicator variables, we develop a closed-form expression for projecting onto a capped rotated second-order cone. The rapid computation of the projection onto this set enables the development of effective projection-based methods for solving the SOCP perspective relaxation of sparse optimization problems. As a proof of concept for the applicability of using our projection, we apply accelerated projected gradient method (i.e., FISTA) and ADMM to use our projection technique in order to effectively solve the perspective relaxation of a sparse regression problem. We also generalize the basic sparse regression formulation and solution method for group sparsity. Finally, we also develop an ADMM algorithm that solves a nonsmooth unconstrained formulation, which is known to be equivalent to the sparse regression perspective relaxation, and we further generalize this method for the group sparsity setting.

Inexact Stochastic Proximal Methods for Sparse Group Structure Identification – GUANYI WANG, NATIONAL UNIVERSITY OF SINGAPORE – Yutong Dai, Frank E. Curtis, Daniel Robinson
The sparse learning problem is widely used in machine learning and data analysis, enhancing feature interpretability and optimizing model deployability. This paper studies the regularized stochastic sparse learning (RSSL) problem, specifically when the underlying group structures are known. However, existing optimization methods for RSSL require periodically evaluating exact gradients or storing a history of stochastic gradients. Moreover, few existing results ensure consistent support identification across all iterations beyond a sufficiently large index, i.e., the optimal group structures might be identified in some of the iterations but not in the rest, even if the index is sufficiently large. To address these challenges, we propose an inexact stochastic proximal algorithm, denoted as IS-PSstorm, for RSSL without periodic evaluation of exact gradients or extensive storage of history. Our algorithm ensures a near-linear point convergence with a $O(1/k)$ variance reduction property, where k represents the number of iterations. Importantly, we establish that IS-PSstorm consistently

identifies the group sparse structure in every iteration beyond a fixed threshold index with high probability. Numerical experiments on regularized logistic loss problems guarantee that IS-PSstorm outperforms existing methods in various metrics on how efficiently an algorithm optimizes the objective and how robustly an algorithm identifies optimal support.

On the power of greedy local search for sparse statistical problems – ILIAS ZADIK, YALE UNIVERSITY –

In this work, we analyze the greedy local search dynamics for a series of maximum likelihood estimation (MLE) sparse problems, including sparse PCA and sparse regression. The MLE problems can naturally be modeled as integer programs. One interesting conclusion of our work, is the provable suboptimality of these methods as opposed to other time-efficient methods, such as convex relaxations.

Logic Rules and Chordal Graphs for Sparse Learning – ANNA DEZA, UC BERKELEY –

We uncover a new relationship between sparse machine learning and chordal graphs, and develop a fast clique generation method to improve the performance of optimization algorithms for l_0 -regularized and l_0 -constrained machine learning models. We present novel logic rules, implying the simultaneous inclusion, exclusion, or ranking of pairs of features in any optimal solution. We describe the applicable logic rules, prove their validity, and give an algorithm that generates all maximal clique inequalities on the chordal conflict graph in $O(\text{plogp})$ time.

2d - ThC928 - 513A - Machine Learning and Discrete Optimization 2 Chair: RAHUL MAZUMDER 

Deep neural network initialisation: Nonlinear activations impact on the Gaussian process – JARED TANNER, UNIVERSITY OF OXFORD –

Randomly initialised deep neural networks are known to generate a Gaussian process for their pre-activation intermediate layers. We will review this line of research with extensions to deep networks having structured random entries such as block-sparse or low-rank weight matrices. We will then discuss how the choice of nonlinear activations impacts the evolution of the Gaussian process. Specifically we will discuss why sparsifying nonlinear activations such as soft thresholding are unstable, we will show conditions to overcome such issues, and we will show how non-sparsifying activations can be improved to be more stable when acting on a data manifold.

This work is joint with Michael Murray (UCLA), Vinayak Abrol (IIT Delhi), Ilan Price (DeepMind), and from Oxford: Alireza Naderi, Thiziri Nait Saada, Nicholas Daultry Ball, Adam C. Jones, and Samuel C.H. Lam.

Functionally-Identical Pruning of Tree Ensembles – ALEXANDRE FOREL, POLYTECHNIQUE MONTRÉAL – Youssouf Emine, Idriss Malek, Thibaut Vidal

We study how to prune a tree ensemble into a reduced one that is “functionally identical” to the original model. That is, we ensure that its prediction function remains unchanged for any input in a domain of interest. Consequently, this pruning algorithm is also lossless for any aggregated metric. At its

core, our algorithm (i) efficiently solves a mixed integer linear program for ensemble compression, (ii) iteratively detects samples for which prediction differs using a separation oracle, and (iii) dynamically adds constraints to guarantee functional identity. We evaluate our method for different domains of interest: discrete spaces containing the training set, or a subset thereof, and a continuous space characterizing plausible inputs. In an extensive computational campaign, we show that this approach gives a “free lunch”, as it can reduce ensemble size without changing model functionality. Our approach is general and can be extended to other ensemble learners for which an adversarial MILP can be formulated.

Discrete Optimization Methods for Pruning and Compressing Large Neural Networks – RAHUL MAZUMDER, MIT –

Foundation models have achieved remarkable performance across various domains, but their large model sizes lead to high computational costs (storage, inference latency, memory, etc). NN pruning, roughly categorized as unstructured (removing individual weights) and structured (removing entire channels, neurons, or heads in attention layers), aims to reduce these costs by removing less-important parameters while retaining model utility as much as possible. Structured pruning is a practical way to improve inference latency on standard hardware (e.g., TPUs) in contrast to unstructured pruning, requiring specialized hardware and software. We propose several novel algorithmic ideas to improve the efficiency of large vision and language models (on standard hardware) via structured pruning.

2d - FA173 - 513A - **Progress on Machine Learning Approaches for Solving Combinatorial Optimization** Chair: ZAIWEN WEN 

A Monte Carlo Policy Gradient Method with Local Search for Binary Optimization – ZAIWEN WEN, PEKING UNIVERSITY –

Binary integer programming problems are ubiquitous in many practical applications, including the MaxCut and cheeger cut problem, the MIMO detection and MaxSAT, etc. They are NP-hard due to the combinatorial structure. In this talk, we present a policy gradient method using deep Monte Carlo local search to ensure sufficient exploration in discrete spaces. The local search method is proved to improve the quality of integer solutions and the policy gradient descent converges to stationary points in expectation. Numerical results show that this framework provides near-optimal solutions efficiently for quite a few binary optimization problems.

Mixed-Integer Linear Optimization via Learning-Based Two-Layer Large Neighborhood Search – AKANG WANG, SHENZHEN RESEARCH INSTITUTE OF BIG DATA – Wenbo Liu, Wenguo Yang, Qingjiang Shi

Mixed-integer linear programs (MILPs) find widespread application in modeling practical problems such as planning and scheduling. A prominent method for solving those MILPs is large neighborhood search (LNS), which iteratively searches for improved solutions within particular neighborhoods of interest. Recent advancements have explored the integration of machine learning techniques into LNS for judiciously guiding users to

construct neighborhoods. However, a computational bottleneck still exists in the search step, as it relies on off-the-shelf solvers to optimize auxiliary MILPs of relatively large sizes. In this study, we propose a two-layer LNS (TLNS) approach that leverages LNS to effectively address both the original MILP and its auxiliary MILPs, requiring the optimization of only small-sized MILPs via off-the-shelf solvers. Furthermore, we utilize trained graph neural networks to inform the neighborhood design. We conduct extensive computational experiments using public benchmarks. The results indicate that our TLNS approach achieves remarkable performance gains—up to 76% and 97% over LNS and state-of-the-art MILP solvers, respectively.

Understanding and Recomposing the Learning and Searching Techniques in ML4CO: An Deep Dive into the TSP Case – JUNCHI YAN, SHANGHAI JIAO TONG UNIVERSITY –

Despite the rich works on machine learning (ML) for combinatorial optimization (CO), there lacks a principled framework. We develop a unified modular framework for CO incorporating existing major techniques in both learning and search to understand the effects of techniques, which is embodied on the Travelling Salesman Problem (TSP). We show that principles like joint probability estimation, symmetry accommodation, online optimization are desirable for ML4CO, which also guide the development of new method designs within our framework. By extensive evaluation, we also show the advantages of nonautoregressive methods over autoregressive ones, as well as the superiority of supervised paradigms over unsupervised ones. The strategic decoupling and organic recompositions within the framework yield a factory of new solvers, among which some achieve SOTA performance in learningbased solvers of optimality drops of 0.000% for TSP-50 and 100, and 0.109% for TSP-500.

Learning to learn for discrete optimization: some recent developments – JIANYONG SUN, XI’AN JIAOTONG UNIVERSITY –

It has become popular in recent years to research on learning to optimize due to the advancement of deep learning and the rising of the AI for Science research paradigm. In this talk, I will present some of our recent works on learning to optimize for discrete optimization problems. In the first part, a learning to optimize algorithm for integer programming within the branch and bound framework is illustrated. A learning to optimize algorithm for multi-objective optimization problems is then presented. Finally, I will present an application to the MIMO signal detection problem in wireless communication.

2d - FB929 - 513A - **Machine Learning and Discrete Optimization 3** Chair: FEDERICO D’ONOFRIO 

FASTopt: An Optimization Framework for Fast Additive Segmentation in Transparent ML – BRIAN LIU, MIT –

We present FAST, an optimization framework for fast additive segmentation. FAST segments piecewise constant shape functions for each feature in a dataset to produce transparent additive models. The framework leverages a novel optimization procedure to fit these models 2 orders of magnitude faster than existing state-of-the-art methods, such as explainable boosting

machines (Nori et al. 2019). We also develop new feature selection algorithms in the FAST framework to fit parsimonious models that perform well. Through experiments and case studies, we show that FAST improves the computational efficiency and interpretability of additive models.

Reliability evaluation of gaussian processes applied to power systems – PIERRE HOUDOUIN, RTE FRANCE – Manuel Ruiz, Patrick Panciatici

Accurately simulating the electrical grid is crucial in power systems. However, it is also very time-consuming, as the entire electric network must be taken into account. Thus, reducing computation time is crucial. To this end, the use of proxies is garnering growing interest in the community. Indeed, it enables the approximation of some simulation results to reduce the overall computational cost. An increasingly popular proxy technique for learning unknown dynamics is Gaussian processes. These processes are theoretically able to learn any black-box function through a statistical model based on local output correlations. Gaussian processes also enable the incorporation of prior knowledge on the system's physics into the proxy and provide confidence intervals around its prediction. They are particularly suitable for critical systems, where uncertainty quantification is essential. However, like all machine learning algorithms, Gaussian processes can sometimes learn poorly leading to inaccurate predictions and erroneous uncertainty quantification. Before trusting the confidence interval provided with the prediction, it is essential to ensure that the learned model is reliable. Our work aims to present a methodology to evaluate the reliability of Gaussian processes, and avoid the severe consequences implied by trusting a flawed model.

Closed-loop Koopman operator approximation – STEVEN DAHDAH, MCGILL UNIVERSITY – James Richard Forbes

This presentation discusses a previously published method to identify a Koopman model of a feedback-controlled system given a known controller. The Koopman operator allows a nonlinear system to be rewritten as an infinite-dimensional linear system by viewing it in terms of an infinite set of lifting functions. A finite-dimensional approximation of the Koopman operator can be identified from data by choosing a finite subset of lifting functions and solving a regression problem in the lifted space. Existing methods are designed to identify open-loop systems. However, it is impractical or impossible to run experiments on some systems, such as unstable systems, in an open-loop fashion. The proposed method leverages the linearity of the Koopman operator, along with knowledge of the controller and the structure of the closed-loop system, to simultaneously identify the closed-loop and plant systems. The advantages of the proposed closed-loop Koopman operator approximation method are demonstrated experimentally using a rotary inverted pendulum system. An open-source software implementation of the proposed method is publicly available, along with the experimental dataset generated for this paper.

Novel Mixed-Integer Optimization Approaches for Interpretable SVMs – FEDERICO D'ONOFRIO, SAPIENZA UNIVERSITY OF ROME – Immanuel Bomze, Yuri Faenza, Laura Palagi, Bo Peng

Interpretability in machine learning continues to grow in importance, focusing on creating models that offer clear expla-

nations of their decisions for both prediction and debugging purposes. Embedded feature selection methods, which aim at selecting the most relevant features and training the machine learning model simultaneously, are particularly effective in identifying optimal feature subsets for accurate and efficient model performance. In this talk, we will discuss two recent projects on interpretable classification using both linear and nonlinear Support Vector Machines (SVMs). We are interested in embedded feature selection models that can be formulated using Mixed-Integer Nonlinear Optimization (MINLO) techniques. In particular, we analyze the case in which a cardinality constraint is added to the standard primal and dual SVM formulations, which limits the number of features to be used in the training process of the classifier. For linear SVMs, we formulate the problem through a novel Mixed-Integer Quadratic Optimization model with a complementarity constraint. We show how to algorithmically exploit scalable conic relaxations to solve the model for large datasets and present both heuristic and exact procedures. Regarding the nonconvex MINLO problem associated with nonlinear SVM models, which has been minimally explored in the literature, we propose a novel decomposition algorithm based on submodular function maximization which can work with polynomial kernels of any degree. Numerical results demonstrate the effectiveness of our algorithms in solving the addressed optimization models, outperforming off-the-shelf solvers and other solution approaches. We will also hint at possible extensions of our work to additional machine learning problems.

2e - MB153 - 513E - Stochastic Combinatorial Optimization I Chair: NICOLE MEGOW

Stochastic Scheduling of Bernoulli Jobs – MARC UETZ, UNIVERSITY OF TWENTE – Antonios Antoniadis, Ruben Hoeksma, Kevin Schewior

This talk addresses the problem to minimize the total expected completion time of a set of N non-preemptive jobs with stochastic processing times on a set of M identical machines. A special case that has recently gained attention is when the processing times follow Bernoulli distributions. It can be seen as a vanilla version of stochastic scheduling because of its clean combinatorics, while still preserving the core difficulties that make stochastic scheduling problems notoriously hard. Gupta et al. (SODA 23) exhibited an $O(\sqrt{M})$ approximation algorithm, where the performance guarantee suppresses poly-logarithmic factors in the number of jobs. We show that the problem admits a PTAS, given that the number of different job processing times is bounded by a constant. The result rests on a series of transformations of an optimal scheduling policy to a well-formed scheduling policy that makes scheduling decisions at specific points in time only, while losing only a negligible factor in expected cost. An optimal well-formed policy is then computed by dynamic programming. Two technical issues are solved, namely (i) to ensure that, with a negligible delay, the transformed policy has an information advantage over the optimal policy, allowing it to simulate the decisions of that policy, and (ii) to ensure that the delays do not accumulate, thus solving the trade-off between the complexity of the scheduling policy and its expected cost. Our result implies a quasi-polynomial time $O(\log N)$ approximation when there is no bound on the number of different processing times.

Approximation Algorithms for Stochastic Minimum Norm Combinatorial Optimization – SHARAT IBRAHIMPUR, LONDON SCHOOL OF ECONOMICS – Chaitanya Swamy

In this work, we introduce and study stochastic minimum-norm optimization. We have an underlying combinatorial optimization problem where the costs involved are random variables with given distributions; each feasible solution induces a random multidimensional cost vector. We seek a solution that minimizes the expected norm of the induced cost vector, for a given monotone, symmetric norm.

We give a general framework for devising approximation algorithms for stochastic minimum-norm optimization, using which we obtain approximation algorithms for the stochastic minimum-norm versions of the load balancing and spanning tree problems.

Learning Prophet Inequality and Pandora's Box with Limited Feedback – THOMAS KESSELHEIM, UNIVERSITY OF BONN – Khashayar Gatmiry, Sahil Singla, Yifan Wang

The Prophet Inequality and Pandora's Box problems are fundamental stochastic problem with applications in Mechanism Design, Online Algorithms, Stochastic Optimization, Optimal Stopping, and Operations Research. A usual assumption in these works is that the probability distributions of the n underlying random variables are given as input to the algorithm. Since in practice these distributions need to be learned under limited feedback, we initiate the study of such stochastic problems in the Multi-Armed Bandits model.

In the Multi-Armed Bandits model we interact with n unknown distributions over T rounds: in round t we play a policy and only receive its value as feedback. The goal is to minimize the regret, which is the difference over T rounds in the total value of the optimal algorithm that knows the distributions vs. the total value of our algorithm that learns the distributions from the limited feedback. Our main results give near-optimal $O(\text{poly}(n) * \sqrt{T} * \text{polylog}(T))$ total regret algorithms for both Prophet Inequality and Pandora's Box.

Our proofs proceed by maintaining confidence intervals on the unknown indices of the optimal policy. The exploration-exploitation tradeoff prevents us from directly refining these confidence intervals, so the main technique is to design a regret upper bound function that is learnable while playing low-regret Bandit policies.

Set Selection under Explorable Stochastic Uncertainty – NICOLE MEGOW, UNIVERSITY OF BREMEN – Jens Schlöter

Given subsets of uncertain values, we study the problem of identifying the subset of minimum total value by querying as few values as possible. This set selection problem falls into the field of explorable uncertainty and is of intrinsic importance therein as it implies strong adversarial lower bounds for a wide range of classical problems such as knapsack and matchings. We consider a stochastic problem variant and give algorithms that, in expectation, improve upon these adversarial lower bounds. The key to our results is to prove a strong structural connection to a seemingly unrelated covering problem with uncertainty in the constraints via a linear programming formulation. We exploit this connection to derive an algorithmic

framework that can be used to solve both problems under uncertainty, obtaining nearly tight bounds on the competitive ratio. This is the first non-trivial stochastic result concerning the sum of unknown values without further structure known for the set and it may be useful for solving more general problems in the area of explorable uncertainty.

2e - MB239 - 514A - **Network design** Chair: NEIL OLVER

Single-source unsplittable flows in planar graphs – LAURA VARGAS KOCH, UNIVERSITY OF BONN –

The single-source unsplittable flow problem asks to send flow in a digraph with capacities from a common source to different terminals with unrelated demands, each terminal being served through a single path. A seminal result of Dinitz, Garg, and Goemans showed that, whenever a fractional flow exists respecting the capacities, then there is an unsplittable one violating the capacities by at most the maximum demand. Goemans conjectured a very natural cost version of the same result, where the unsplittable flow is required to be no more expensive than the fractional one. So far there are no non-trivial graph classes for which it is known to hold.

In the talk, we will see that a slight weakening of it (with at most twice as large violations) holds for planar graphs. This result is based on a connection to a highly structured discrepancy problem, whose repeated resolution allows us to successively reduce the number of paths used for each terminal, until we obtain an unsplittable flow.

This is joint work with Vera Traub and Rico Zenklusen.

New Structures and Algorithms for Length-Constrained Expander Decompositions – ELLIS HERSHKOWITZ, BROWN UNIVERSITY –

Expander decompositions are one of the most ubiquitous and flexible paradigms for close-to-linear-time graph algorithms. Length-constrained expander decompositions generalize this paradigm to better work for problems with lengths, distances and costs. Roughly, an (h,s) -length ϕ -expander decomposition is a collection of length increases to a graph so that nodes within distance h can route flow over paths of length hs with congestion at most $1/\phi$.

This talk, will cover new close-to-linear-time algorithms for computing length-constrained expander decompositions in graphs with general lengths and capacities. These algorithms allow one to trade off between the size of the decomposition and the length of routing paths: for any $\epsilon > 0$ not too small, it computes in almost-linear-time an (h,s) -length ϕ -decomposition of size H where $s = \exp(\text{poly}(1/\epsilon))$. The key foundations of our algorithm are: (1) a simple yet powerful structural theorem which states that the union of a sequence of sparse length-constrained cuts is itself sparse and (2) new algorithms for computing sparse length-constrained flows.

Based on joint work with Zihan Tan and Bernhard Haeupler.

Ghost Value Augmentation for k-Edge-Connectivity – NATHAN KLEIN, INSTITUTE OF ADVANCED STUDY –

Computing k -edge-connected spanning subgraphs of minimum

cost is a fundamental problem in combinatorial optimization. For $k=1$, this is exactly the minimum spanning tree problem. For general k this is known as the k -edge-connected spanning subgraph (k -ECSS) problem and is NP-Hard. A 2-approximation for this problem can be obtained by a variety of techniques such as Jain's iterative rounding, and obtaining a better-than-2 approximation for any k is a major open question.

In this work, we approach approximating k -ECSS from a resource augmentation perspective, inspired by resource augmentation results on other network design problems like bounded degree spanning trees. I will describe an algorithm that returns a k -ECSS solution of cost no greater than the cheapest $(k+10)$ -ECSS. Equivalently, we can find a $(k-10)$ -ECSS solution of cost no more than the cheapest k -ECSS. Thus, for large k , this presents a promising alternative to a 2-approximation: instead of losing a factor 2 in the cost, one can instead lose $O(1)$ connectivity. As a byproduct of this result, we also obtain a $1+O(1/k)$ for the k -ECSS problem, a variant of k -ECSS in which one can select an edge multiple times at the same cost. This resolves a conjecture of Pritchard from 2011 and improves upon a recent work of Karlin, Klein, Oveis Gharan, and Zhang. We also give a matching lower bound of $1+\Omega(1/k)$, showing that this is the best possible asymptotic ratio unless $P=NP$.

Based on joint work with D Ellis Hershkowitz and Rico Zenklusen.

LP-Based Algorithms for Two-Cost Network Design – RHEA JAIN, UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN –

In several standard network design problems, given an input graph with fixed edge costs and source/sink pairs of vertices, the goal is to find a low-cost subgraph satisfying some connectivity properties. We consider two-cost network design models, in which edges of the input graph have an associated length along with the fixed cost. One fundamental such problem is the multicommodity buy-at-bulk problem: given demand pairs s_i, t_i with demand $d(i)$, the goal is to route $d(i)$ flow between s_i and t_i while minimizing the sum of the fixed cost (total cost of edges used) and routing cost (total length of flow paths). Chekuri et al. gave a polylogarithmic approximation via junction trees and posed the question of establishing an upper bound on the integrality gap of a natural LP relaxation.

In this talk, we will see that we can obtain a polylogarithmic upper bound on the integrality gap using recent results in hop-constrained oblivious routing. We demonstrate a connection between h -hop-constrained network design problems (in which each demand pair must be connected with a path of at most h edges) and obtain bi-criteria approximation algorithms for these problems as well.

This is joint work with Chandra Chekuri.

2e - TA229 - 512A - Algorithms for Pricing and Assortment Optimization Chair: OMAR EL HOUSNI

Submodular Order Functions and Their Applications – RAJAN UDWANI, UC BERKELEY –

We define a new class of set functions that in addition to being monotone and subadditive, also admit a very limited form

of submodularity defined over a permutation of the ground set. We refer to this permutation as a submodular order and give fast and best possible approximation algorithms for offline and online maximization of submodular order functions. We show the utility of this new notion in a variety of applications, including static and dynamic assortment optimization, online allocation of reusable resources, and large-scale optimization in the streaming model.

Optimizing and Learning Assortment Decisions in the Presence of Platform Disengagement – MIKA SUMIDA, UNIVERSITY OF SOUTHERN CALIFORNIA – Angela Zhou

We consider a problem where customers repeatedly interact with a platform. During each interaction with the platform, the customer is shown an assortment of items and selects among these items according to a Multinomial Logit choice model. The probability that a customer interacts with the platform in the next period depends on the customer's past purchase history. The goal of the platform is to maximize the total revenue obtained from each customer over a finite time horizon.

First, we study a non-learning version of the problem where consumer preferences and return probabilities are completely known. We formulate the problem as a dynamic program and prove structural properties of the optimal policy. Next, we provide a formulation in a contextual episodic reinforcement learning setting, where the parameters governing contextual consumer preferences and return probabilities are unknown and learned over multiple episodes. We develop an algorithm based on the principle of optimism under uncertainty for this problem and provide a regret bound.

Previous approaches that address user disengagement often constrain exploration. However, in our model with non-permanent disengagement with assortments, the optimal solution simply offers larger assortments at the beginning of the horizon and exploration is unconstrained during the learning process. We numerically illustrate model insights and demonstrate regimes where our algorithm outperforms naively myopic learning algorithms.

Learning to Price Under Competition for Multinomial Logit Demand – SANJAY MEHROTRA, NORTHWESTERN UNIVERSITY – Vineet Goyal, Shukai Li

We consider a sequential price competition problem a seller faces when multiple sellers selling the same product compete on prices under an unknown multinomial logit (MNL) demand. In particular, at the beginning of each period, each seller first (simultaneously) posts a price and then observes their own realized demand and posted prices of other sellers. However, they do not observe the demand of other sellers. The goal is to find a pricing policy as a function of historical observations to maximize the expected revenue over the selling horizon. This problem is introduced in van de Geer et al. (2019) as a part of the Dynamic Pricing Challenge at the 2017 INFORMS Revenue Management & Pricing Conference. However, to our knowledge, prior to this work there is no known algorithm with provable guarantees. Gallego et al. (2006) study a corresponding single period price competition problem under a known MNL model and show a unique pure-strategy Nash equilibrium. We give a stochastic gradient-based online learning algorithm in the sequential setting and show that it converges to the unique

Nash equilibrium. In particular, we reformulate the learning problem with multiple sellers under a stationary MNL demand to learning under a non-stationary logit demand for a single seller. Furthermore, we show that for any seller i , our algorithm achieves an expected regret at most $O(T^{2/3} \log T \log 2N)$ as compared to the expected revenue of seller i under full information where we fix the sequence of prices of other sellers in all periods and optimize the price for seller i in each period. Here N and T are the numbers of sellers and sequentially arriving customers respectively. This full-information benchmark is stronger than the benchmark of expected revenue for the seller at Nash equilibrium.

Maximum Load Assortment Optimization
 – OMAR EL HOUSNI, CORNELL UNIVERSITY –
 Marouane Ibn Brahim, Danny Segev

Motivated by modern-day applications such as Attended Home Delivery and Preference-based Group Scheduling, where decision makers wish to steer a large number of customers toward choosing the exact same alternative, we introduce a novel class of assortment optimization problems, referred to as Maximum Load Assortment Optimization. In such settings, given a universe of substitutable products, we are facing a stream of customers, each choosing between either selecting a product out of an offered assortment or opting to leave without making a selection. Assuming that these decisions are governed by the Multinomial Logit choice model, we define the random load of any underlying product as the total number of customers who select it. Our objective is to offer an assortment of products to each customer so that the expected maximum load across all products is maximized. We consider both static and dynamic formulations. In the static setting, a single offer set is carried throughout the entire process of customer arrivals, whereas in the dynamic setting, the decision maker offers a personalized assortment to each customer, based on the entire information available at that time. The main contribution of this paper resides in proposing efficient algorithmic approaches for computing near-optimal static and dynamic assortment policies. In particular, we develop a polynomial-time approximation scheme (PTAS) for the static formulation. Additionally, we demonstrate that an elegant policy utilizing weight-ordered assortments yields a $1/2$ -approximation. Concurrently, we prove that such policies are sufficiently strong to provide a $1/4$ -approximation with respect to the dynamic formulation, establishing a constant-factor bound on its adaptivity gap. Finally, we design an adaptive policy whose expected maximum load is within factor $1 - \epsilon$ of optimal, admitting a quasi-polynomial time implementation.

2e - TA930 - 514B - Approximation and Online Algorithms 1 Chair: ALENA OTTO 

Online Matching on 3-Uniform Hypergraphs
 – DANISH KASHAEV, CENTRUM WISKUNDE & INFORMATICA –

The online matching problem was introduced by Karp, Vazirani and Vazirani (STOC 1990) on bipartite graphs with vertex arrivals. It is well-known that the optimal competitive ratio is $1 - 1/e$ for both integral and fractional versions of the problem. Since then, there has been considerable effort to find optimal competitive ratios for other related settings.

In this work, we go beyond the graph case and study the

online matching problem on k -uniform hypergraphs. For $k = 3$, we provide an optimal primal-dual fractional algorithm, which achieves a competitive ratio of $(e - 1)/(e + 1) \approx 0.4621$. As our main technical contribution, we present a carefully constructed adversarial instance, which shows that this ratio is in fact optimal. It combines ideas from known hard instances for bipartite graphs under the edge-arrival and vertex-arrival models.

For $k \geq 3$, we give a simple integral algorithm which performs better than greedy when the online nodes have bounded degree. As a corollary, it achieves the optimal competitive ratio of $1/2$ on 3-uniform hypergraphs when every online node has degree at most 2. This is because the special case where every online node has degree 1 is equivalent to the edge-arrival model on graphs, for which an upper bound of $1/2$ is known.

Nearly optimal node selection in the explorable heap model – SANDER BORST, CWI –
 Daniel Dadush, Sophie Huiberts, Danish Kashaev

An important aspect of the branch-and-bound algorithm is the node selection rule, which decides the order in which the solution space is explored. Ideally, such a rule guarantees not only that the number of unnecessary subproblems that are solved is small, but also that consecutively solved problems are close in the branch-and-bound tree. The design of a good node selection rule is equivalent to the solving the 'explorable heap selection' problem, an online graph exploration problem which was originally proposed by Karp, Saks and Wigderson (FOCS '86).

In this talk we discuss the connection between the two problems and we propose a new randomized algorithm for this problem with significantly better performance guarantees than any previously known algorithm. In the low-space setting this algorithm is nearly optimal in the light of a new lower bound that we provide.

TSP-T3CO: A Classification Scheme for Variants of the Traveling Salesman Problem – ANDREAS KARRENBauer, MAX PLANCK INSTITUTE FOR INFORMATICS –
 Sophia Saller, Jana Koehler

We propose TSP-T3CO as a uniform, easy-to-use and extensible means for the formal and precise classification of TSP variants. It allows to reveal subtle differences within the same named variant and also brings out the differences between variants more clearly. We demonstrate its merits by a comprehensive, concise, and compact display of approximability results for a large number of TSP variants. This makes it easier to understand the approximability landscape and the assumptions under which certain results hold. Open gaps become more evident and results can be compared more easily. [arXiv:2311.00604]

Reliable Adjustable Planning in Case of Dynamically Arriving Information: Example of E-Commerce Warehouses – ALENA OTTO, UNIVERSITY OF PASSAU –
 Catherine Lorenz, Michel Gendreau

Many businesses require intelligent adjustable algorithms, whose behavior is well-understood for a reliable practical deployment. This talk focuses on the application of adjustable planning for picking operations in e-commerce warehouses. Since order picking accounts for a substantial portion (60-70%) of total operating costs in warehouses, it holds a pivotal role.

In e-commerce settings, picking operations are essentially online, with customer orders arriving dynamically in real-time, highlighting the need for adjustable algorithms, or online policies.

Although several online policies for order picking have been suggested in the literature, there is a notable lack of understanding of their performance with respect to optimality. In this talk, we will conduct a comprehensive examination of a prominent online warehousing policy – myopic, immediate reoptimization (Reopt) – and establish analytical performance bounds using probabilistic- and worst-case analysis. We demonstrate that, under broad stochastic assumptions, Reopt is almost surely asymptotically optimal and, notably, we validate its near-optimal performance empirically.

2e - TC358 - 514B - **Approximation in Algorithmic Game Theory** Chair: CHAITANYA SWAMY



Fair Allocation of Indivisibles Beyond Additive Valuations – POOJA KULKARNI, UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN –

Fair Allocation of resources is a fundamental problem that has been studied by the Economics and Computer Science communities. This problem consists of dividing a set of resources among a set of agents in a fair way. Fairness resists a single definition and various notions of fairness have been defined and investigated in literature. In this talk we consider two fairness notions – AnyPrice Share (APS) and Nash Social Welfare (NSW). We study the algorithmic aspects of these fairness notions when valuation functions of the agents are submodular or fractionally subadditive (XOS).

APS is a share-based notion where each agent determines a value that she thinks is fair and we have to output an allocation where the agents get this value. We will discuss a simple greedy algorithm that ensures $1/3$ fraction of APS to every agent when the valuation functions of agents are submodular. This is the current best-known approximation of APS for submodular valuations. Nash social welfare is a single objective where the aim is to maximize the geometric mean of values received by the agents. I will touch upon a sublinear approximation algorithm for maximizing NSW with XOS valuation functions. This algorithm uses Capped Welfare maximization to maximize NSW. This was the first algorithm to use social welfare for maximization of Nash social welfare.

Budget-Feasible Mechanism Design: Simpler, Better Mechanisms and General Payment Constraints – RIAN NEOGI, UNIVERSITY OF WATERLOO –

In the setting of budget feasible mechanism design, a buyer wants to purchase items from a set of sellers. Each seller i supplies one item, upon which they incur a cost of c_i . The buyer wants to buy a set of items from the sellers so as to maximize a certain valuation function, subject to a knapsack constraint stating that the total cost of the items bought is at most a certain budget. The true cost c_i is private information that is only known to seller i and not to the buyer. Thus, the buyer must work with the seller's reported costs. The goal is to design a mechanism that is truthful, in the sense that the sellers do not have an incentive to deviate from reporting their true

costs, and budget feasible, in the sense that the total payments made to the sellers is within a given budget B .

Bei et al. considered the setting when the valuation function is XOS and gave the first polynomial time $O(1)$ approximation for the problem. Their mechanism has essentially remained the only mechanism for XOS valuations, and requires both a demand oracle and an XOS oracle. They also showed the existence of a $O(1)$ approximation for when the valuation function is subadditive. However, this result is existential in nature, in the sense that we do not know what the mechanism is that obtains this approximation, we just know that one such mechanism exists.

In this talk, I will speak about our recent results in budget feasible mechanism design. When the valuation function is XOS, we design budget feasible mechanisms that are simpler and better, both in terms of the approximation factors they obtain and in terms of requiring weaker oracle access to the valuation function. When the valuation function is subadditive, we give the first polynomial time explicitly constructed constant factor approximation to the problem, whereas previously only an existential result was known. Moreover, we introduce a novel class of mechanism design problems that we dub generalized budget feasible mechanisms, wherein we allow more intricate payment constraints, as opposed to a single constraint on the total payment.

This talk is based on joint work with Chaitanya Swamy and Konstantin Pashkovich.

Multidimensional Budget-Feasible Mechanism Design – CHAITANYA SWAMY, UNIVERSITY OF WATERLOO –

In budget-feasible mechanism design, a buyer wishes to procure a set of items of maximum value from self-interested rational players. We have a ground set U of items, and a nonnegative valuation function v , where $v(S)$ specifies the value obtained from set S of items. The entirety of current work on budget-feasible mechanisms has focussed on the single-dimensional setting, wherein each player holds a single item e and incurs a private cost c_e for supplying item e . We introduce multidimensional budget feasible mechanism design: the universe U is now partitioned into item-sets $\{G_i\}$ held by the different players, and each player i incurs a private cost $c_i(S_i)$ for supplying set S_i of items, which is a subset of G_i . A budget-feasible mechanism is a mechanism (i.e., an algorithm and a payment scheme) that is truthful, i.e., where players are incentivized to report their true costs, and where the total payment made to the players is at most some given budget B . The goal (as in the single-dimensional setting) is to devise a budget-feasible mechanism that procures a set of items of large value.

2e - WA931 - 514B - **Approximation and Online Algorithms 2** Chair: JAROSLAV FOWKES



Scheduling with Obligatory Tests – THOMAS ERLEBACH, DURHAM UNIVERSITY – Konstantinos Dogeas, Ya Chun Liang

Motivated by settings such as medical treatments or aircraft maintenance, we consider a scheduling problem with jobs that consist of two operations, a test and a processing part. The time required to execute the test is known in advance while the time required to execute the processing part becomes known

only upon completion of the test. We use competitive analysis to study algorithms for minimizing the sum of completion times for n given jobs on a single machine. As our main result, we prove using a novel analysis technique that the natural 1-SORT algorithm has competitive ratio at most 1.861. For the special case of uniform test times, we show that a simple threshold-based algorithm has competitive ratio at most 1.585. We also prove a lower bound that shows that no deterministic algorithm can be better than $\sqrt{2}$ -competitive even in the case of uniform test times.

An Exact Method for the Job Sequencing and Tool Switching Problem – ALBERTO LOCATELLI, UNIVERSITY OF MODENA AND REGGIO EMILIA – Jean François Côté, Leandro C. Coelho

In the Job Sequencing and Tool Switching Problem (SSP), a single flexible machine with a capacitated tool magazine has to process a set of jobs, each requiring a specific set of tools. The SSP aims to simultaneously determine the processing sequence and the subsets of tools present in the magazine for each job in order to minimize the total number of tool switches. We propose a new Quadratic Travelling Salesman Problem-based formulation for the SSP. We then develop an exact method, which decomposes the problem into a Master Problem (MP) for determining the processing sequence, and a Sub Problem (SP) for determining the subsets of tools present in the magazine for each job. Optimality cuts, generated by the SP, are dynamically added to the MP to gradually obtain a tighter approximation of the objective function at the optimal solution. Computational tests on instances from the literature were performed to evaluate the effectiveness of the proposed approach.

Recent progress for correlation clustering – ALANTHA NEWMAN, CNRS UNIVERSITÉ GRENOBLE ALPES – Nairen Cao, Vincent Cohen Addad, Euiwoong Lee, Shi Li, Lukas Vogl

In the correlation clustering problem, we are given a complete graph with each edge labeled as + (similar) or - (dissimilar). The goal is to find a clustering (a partition) of the vertices that minimizes the number of dissimilar intracluster edges and the number of similar intercluster edges. Until recently, the best known approximation guarantee was 2.06, based on a modification of the natural LP-based pivot algorithm. We now know how to go below 1.5 using various new tools and insights. In this talk, we will give an overview of these recent developments for the correlation clustering problem.

Approximating large-scale Hessian matrices using secant equations – JAROSLAV FOWKES, SCIENCE AND TECHNOLOGY FACILITIES COUNCIL – Nick Gould, Jennifer Scott

Large-scale optimization algorithms frequently require sparse Hessian matrices that are not readily available. Existing methods for approximating large sparse Hessian matrices either do not impose sparsity or are computationally prohibitive. To try and overcome these limitations, we propose a novel approach that seeks to satisfy as many componentwise secant equations as necessary to define each row of the Hessian matrix. A naive application of this approach is prohibitively expensive on Hessian matrices that have some relatively dense rows, but by carefully taking into account the symmetry and connectivity of the Hessian matrix we are able to devise an approximation al-

gorithm that is fast and efficient with scope for parallelism. Example sparse Hessian matrices from the CUTEst test problem collection for optimization illustrate the effectiveness and robustness of our proposed method.

2e - WB146 - 514B - Stochastic Combinatorial Optimization II Chair: VISWANATH NAGARAJAN [↔](#)

Fast Online Node Labeling for Very Large Graphs – YIFAN SUN, STONY BROOK UNIVERSITY

– We consider the online node labeling problem over very large graphs, where the prediction requires computing a large matrix inverse, with sparsity dictated by the graph structure. In particular, we consider the regime where the number of nodes is so large that common sparsifiers and repeated multiplications are still computationally intractable. Here, we examine local methods, such as the approximate Page-Rank method, whose computational complexity is altogether independent of graph size. Using such methods, we recompute the regret bounds for online learning under this more realistic computational framework, and propose a framework for its acceleration.

Semi-Bandit Learning for Monotone Stochastic Optimization – ROHAN GHUGE, GEORGIA TECH – Arpit Agarwal, Viswanath Nagarajan

Stochastic optimization is a widely used approach for optimization under uncertainty, where uncertain input parameters are modeled by random variables. Exact or approximation algorithms have been obtained for several fundamental problems in this area. However, a significant limitation of this approach is that it requires full knowledge of the underlying probability distributions. Can we still get good (approximation) algorithms if these distributions are unknown, and the algorithm needs to learn them through repeated interactions? We resolve this question for a large class of “monotone” stochastic problems, by providing a generic online learning algorithm with $\sqrt{T \log T}$ regret relative to the best approximation algorithm (under known distributions). Importantly, our online algorithm works in a semi-bandit setting, where in each period, the algorithm only observes samples from the random variables that were actually probed. Our framework applies to several fundamental problems in stochastic optimization such as prophet inequality, Pandora’s box, stochastic knapsack, stochastic matchings and stochastic submodular optimization.

Power of Adaptivity for Stochastic Submodular Cover – VISWANATH NAGARAJAN, UNIVERSITY OF MICHIGAN – Rohan Ghuge, Anupam Gupta

We study the stochastic submodular cover problem, where one needs to select a subset of stochastic items to cover a submodular function at minimum expected cost. Solutions in this setting correspond to sequential decision processes that select items one-by-one adaptively, depending on prior observations. While such adaptive solutions achieve the best objective, they are inherently sequential, which is undesirable in many applications. We show how to obtain solutions that approximate fully-adaptive solutions using only a few “rounds” of adaptivity. We study both independent and correlated settings, proving smooth tradeoffs between the number of adaptive rounds and the solution quality. We also present experimental results demonstrating that a few rounds of adaptivity suffice to obtain

high-quality solutions in practice.

Recent Advances in Stochastic Function Evaluation: Score Classification and Voting – KEVIN SCHEWIOR, UNIVERSITY OF SOUTHERN DENMARK –

In stochastic function evaluation, one is given a (succinctly represented) Boolean function and n (typically independent) random variables with known probability distributions. Each of the variables can be queried at known cost, which reveals the value of the respective variable. The goal is to efficiently compute a strategy for querying variables that determines the value of the function applied to these variables at (approximately) minimum expected cost. In the first part of this talk, we consider score-classification functions (Gkenosis, Grammel, Hellerstein, and Kletenik; 2018). Here, variables are Boolean and the function value only depends on which given subinterval of $[0, n]$ the total number of successes is in. We give simple non-adaptive constant-factor approximation algorithms. This part is joint work with Benedikt Plank. In the second part of this talk, we consider a setting where the random variables correspond to votes, their values correspond to candidates, and the value of the function is the outcome of an election. We give adaptive constant-factor approximations for the absolute-majority and relative-majority versions of this problem. This part is joint work with Lisa Hellerstein and Naifeng Liu.

2e - ThA293 - 514B - Approximation Algorithms for the Traveling Salesman Problem

Chair: DAVID WILLIAMSON

A Lower Bound for the Max Entropy Algorithm for the Traveling Salesman Problem – BILLY JIN, CORNELL UNIVERSITY – Nathan Klein, David Williamson

One of the most famous conjectures in combinatorial optimization is the four-thirds conjecture, which states that the integrality gap of the subtour LP relaxation of the TSP is equal to $4/3$. For 40 years, the best known upper bound was 1.5. Recently, Karlin, Klein, and Oveis Gharan showed that the max entropy algorithm for the TSP gives an improved bound of $1.5 - 10^{-36}$. In this paper, we show that the approximation ratio of the max entropy algorithm is at least 1.375, even for graphic TSP. Thus the max entropy algorithm does not appear to be the algorithm that will ultimately resolve the four-thirds conjecture in the affirmative, should that be possible.

Approximation Algorithms for Traveling Salesman Problems – JENS VYGEN, UNIVERSITY OF BONN – Vera Traub

Our new book (with the same title as this talk) describes and advances the state of the art on approximation algorithms for the main variants of the traveling salesman problem. Here we present some new results from this book (including a better approximation ratio for the asymmetric TSP and its path version), outline the state of the art, and mention some important open problems.

Small additive error for unsplittable multicommodity flow in outerplanar graphs – RICHARD SHAPLEY, CORNELL – David Shmoys

We consider an unsplittable version of the minimum max-load multicommodity flow problem, where each demand must be

routed along a single path. The objective is to minimize the maximum load on any edge in the network. In a seminal work, Schrijver, Seymour, and Winkler showed how to efficiently solve this problem on a cycle to within an additive term of $\frac{3}{2}W$ of the optimal value, where W is the largest demand between any two nodes. We extend their result to outerplanar graphs and provide an efficient algorithm for this problem that exceeds the optimal value by an additive term of no more than $O(W \log k)$, where k is the number of faces in the graph. This implies an $O(\log k)$ approximation ratio. We also extend this result to planar graphs with bounded treewidth and demands on the outer face, for which we also achieve an additive $O(W \log k)$ error term.

A $4/3$ -Approximation Algorithm for Half-Integral Cycle Cut Instances of the TSP – DAVID WILLIAMSON, CORNELL UNIVERSITY – Billy Jin, Nathan Klein

A long-standing conjecture for the traveling salesman problem (TSP) states that the integrality gap of the standard linear programming relaxation of the TSP (sometimes called the Subtour LP or the Held-Karp bound) is at most $4/3$ for symmetric instances of the TSP obeying the triangle inequality. We consider the half-integral case, in which a feasible solution to the LP has solution values in $0, 1/2, 1$. Karlin, Klein, and Oveis Gharan, in a breakthrough result, were able to show that in the half-integral case, the integrality gap is at most 1.49993; Gupta et al. show a slight improvement of this result to 1.4983. Both of these papers consider a hierarchy of critical tight sets in the support graph of the LP solution, in which some of the sets correspond to ‘cycle cuts’ and the others to ‘degree cuts’. Here we show that if all the sets in the hierarchy correspond to cycle cuts, then we can find a distribution of tours whose expected cost is at most $4/3$ times the value of the half-integral LP solution; sampling from the distribution gives us a randomized $4/3$ -approximation algorithm. We note that known bad cases for the integrality gap have a gap of $4/3$ and have a half-integral LP solution in which all the critical tight sets in the hierarchy are cycle cuts; thus our result is tight. This is joint work with Nathan Klein and Billy Jin.

2e - ThB244 - 514B - Advanced algorithms for chip design Chair: STEPHAN HELD

Depth Optimization in Binary Addition – SUSANNE ARMBRUSTER, UNIVERSITY OF BONN – Ulrich Brenner

We consider the fundamental problem of constructing fast circuits for the carry bit computation in binary addition. This problem basically reduces to computing And-Or paths, i.e., functions $(t_1 \cap (t_2 \cup (t_3 \cap \dots t_m) \dots))$ for some Boolean variables t_i . We present an algorithm based on Grinchuk that computes an And-Or path circuit with the smallest known depth, which is a depth of $\log_2(m) + \log_2(\log_2(m))$ for all $m \geq 4$. This approaches the best known asymptotic lower bound of $\log_2(m) + \log_2(\log_2(m)) - 5$ due to Commentz-Walter for the minimum depth of any adder circuit. Our technique is based on recursive splitting of the standard And-Or path circuit into smaller circuits for extended And-Or paths.

Tighter Approximation for the Uniform Cost-Distance Steiner Tree Problem – STEPHAN HELD, UNIVERSITY OF BONN – Josefine Foos, Yan-

nik K.D. Spitzley

Uniform cost-distance Steiner trees minimize the sum of the total length and weighted path lengths from a dedicated root to the other terminals. They are applied when the tree is intended for signal transmission, e.g. in chip design or telecommunication networks. They are a special case of general cost-distance Steiner trees, where different distance functions are used for total length and path lengths. We improve the best published approximation factor for the uniform cost-distance Steiner tree problem from 2.39 [15] to 2.05. If we can approximate the minimum-length Steiner tree problem arbitrarily well, our algorithm achieves an approximation factor arbitrarily close to $1 + 1/\sqrt{2}$. This bound is tight in the following sense. We also prove the gap $1 + 1/\sqrt{2}$ between optimum solutions and the lower bound which we and all previous approximation algorithms for this problem use. Similarly to previous approaches, we start with an approximate minimum-length Steiner tree and split it into subtrees that are later re-connected. To improve the approximation factor, we split it into components more carefully, taking the cost structure into account, and we significantly enhance the analysis.

Resource Sharing revisited: Local weak duality and optimal convergence – DANIEL BLANKENBURG, UNIVERSITY OF BONN –

We revisit the (block-angular) min-max resource sharing problem, which is a well-known generalization of fractional packing and the maximum concurrent flow problem. It consists of finding a resource allocation for every customer, such that the maximum resource usage is minimized. Resource sharing has several applications in VLSI design. Among others, it is used to model the global routing problem. We improve on the currently fastest known FPTAS in various ways. A major novelty of our analysis is the concept of local weak duality, which illustrates that the algorithm optimizes (close to) independent parts of the instance separately and that the second-highest entry of the computed solution is approximately minimal. Further, we provide a lower bound on the number of required oracle calls for a natural class of algorithms. Our FPTAS is optimal within this class; its running time matches the lower bound precisely, and thus improves on the previously best-known running time for the primal as well as the dual problem.

Decision-guided SAT for hard routing problems – JAN MALTE SCHÜRKS, UNIVERSITY OF BONN – Stefan Hougardy

We present a new SAT based algorithm for packing Steiner trees with additional constraints. Such problems arise as routing problems in chip design with complex design rules. Our approach uses a modification of the standard CDCL SAT solving algorithm and incorporates strategies similar to ripup-and-reroute. The resulting algorithm is significantly faster than ILP routers, produces good routings in practice, and guarantees to find a solution if it exists.

2e - FA150 - 514B - Submodularity and Beyond in Approximation Algorithms Chair:

KARTHEKEYAN CHANDRASEKARAN



Improved Approximation Algorithms Beyond Uncrossable Functions – ISHAN BANSAL, CORNELL UNIVERSITY – Joseph Cheriyan, Sharat Ibrahimpur, Logan Grout

We address long-standing open questions raised by Williamson, Goemans, Vazirani and Mihail pertaining to the design of approximation algorithms for problems in network design via the primal-dual method (Combinatorica 15(3):435-454, 1995). Williamson et al. prove an approximation guarantee of two for connectivity augmentation problems where the connectivity requirements can be specified by so-called uncrossable functions. They state: “Extending our algorithm to handle non-uncrossable functions remains a challenging open problem. The key feature of uncrossable functions is that there exists an optimal dual solution which is laminar. This property characterizes uncrossable functions... A larger open issue is to explore further the power of the primal-dual approach for obtaining approximation algorithms for other combinatorial optimization problems.” Our main result proves an $O(1)$ -approximation guarantee via the primal-dual method for a class of functions that generalizes the notion of an uncrossable function. We mention that the support of every optimal dual solution could be non-laminar for instances that can be handled by our methods. We present two applications of our main result: 1. An $O(1)$ -approximation algorithm for augmenting the family of near-minimum cuts of a graph. 2. An $O(1)$ -approximation algorithm for the model of $(p, 2)$ -Flexible Graph Connectivity. This is joint work with Joseph Cheriyan, Logan Grout and Sharat Ibrahimpur and has appeared at ICALP’23.

Polyhedral Aspects of Feedback Vertex Set and Pseudoforest Deletion Set – SHUBHANG KULKARNI, UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN – Karthekeyan Chandrasekaran, Chandra Chekuri, Samuel Fiorini, Stefan Weltge

In the feedback vertex set (FVS) problem, the objective is to delete a minimum cost subset of vertices to make the graph acyclic. FVS is a fundamental NP-hard problem. Approximation algorithms for FVS have relied on the local ratio technique, which have subsequently led to primal-dual algorithms via LPs. While these algorithms achieve the best approximation factor possible, the lack of LP-based primal rounding algorithms for FVS with optimal approximation guarantees has been a barrier for resolving the approximability of the more general Subset-FVS problem.

In this talk, I will present some recent progress towards this goal of designing primal rounding algorithms for FVS. In particular, I will focus on the closely related pseudoforest deletion set (PFDS) problem, where the objective is to delete a minimum cost subset of vertices so that every connected component has at most one cycle. PFDS is also NP-hard and shares the same approximability and algorithmic status as FVS. I will present primal rounding algorithms for PFDS that achieve nearly the best approximation factor possible. The rounding algorithms are based on the iterative rounding framework, and rely on extreme point properties of the associated polyhedra. I will prove one of these extreme point properties by showing that the constraints conditionally exhibit a supermodularity property. This leads to conditional uncrossing, a new technique which may be of independent interest.

Quotient sparsification for submodular functions – KENT QUANRUD, PURDUE UNIVERSITY – Graph sparsification has been an important topic with many structural and algorithmic consequences. Recently hypergraph sparsification has come to the fore and has seen exciting

progress. In this paper we take a fresh perspective and show that they can be both be derived as corollaries of a general theorem on sparsifying matroids and monotone submodular functions.

Quotients of matroids and monotone submodular functions generalize k -cuts in graphs and hypergraphs. We show that a weighted ground set of a monotone submodular function f can be sparsified while approximately preserving the weight of every quotient of f with high probability in randomized polynomial time.

This theorem conceptually unifies cut sparsifiers for undirected graphs [BK15] with other interesting applications. One basic application is to reduce the number of elements in a matroid while preserving the weight of every quotient of the matroid. For hypergraphs, the theorem gives an alternative approach to the hypergraph cut sparsifiers obtained recently in [CKN20], that also preserves all k -cuts. Another application is to reduce the number of points in a set system while preserving the weight of the union of every collection of sets. We also present algorithms that sparsify hypergraphs and set systems in nearly linear time, and sparsify matroids in nearly linear time and queries in the rank oracle model.

Maximizing a Monotone Submodular Function Under an Unknown Knapsack Capacity – SABINE MÜNCH, TRIER UNIVERSITY – Sven De Vries, Stephen Raach

Consider the problem of maximizing a monotone increasing, submodular function defined on a set of weighted items under an unknown knapsack capacity. We require that the knapsack capacity is larger or equal than the heaviest item. Assume that the items are packed one by one into the knapsack and that the knapsack capacity can only be accessed through an oracle that answers whether a currently packed knapsack is feasible or not. If an item is attempted to be added to the present solution and fits into the knapsack, it is packed irrevocably. If the addition of an item more than exhausts the residual capacity of the knapsack, the item is removed from further consideration. We analyze the performance of universal policies, thus nonadaptive packing according to a predetermined sequence. It is well known that Wolsey's greedy algorithm (1982) solves the same problem with known knapsack capacity with an approximation factor of 0.357, where the approximation factor is the worst case ratio between the solution obtained by the algorithm and the optimal solution. We present an algorithm to compute universal policies that perform, for any unknown but fixed knapsack capacity, at least as good as Wolsey's algorithm.

2e - FA245 - 512A - **TSP and vehicle routing**
Chair: STEFAN HOUGARDY

Improved guarantees for the a priori TSP – LUISE PUHLMANN, RESEARCH INSTITUTE FOR DISCRETE MATHEMATICS, UNIVERSITY OF BONN – Jannis Blauth, Meike Neuwohner, Jens Vygen

In the a priori TSP, we are given a metric space V of customers and an activation probability for each customer. We ask for a TSP tour T for V that minimizes the expected length after cutting T short by skipping the inactive customers. All known approximation algorithms select a nonempty subset S of the customers and construct a master route solution, consisting of

a TSP tour for S and two edges connecting every customer not in S to a nearest customer in S . We address the following questions: If we randomly sample the subset S , what should be the sampling probabilities? How much worse than the optimum can the best master route solution be? This is joint work with Jannis Blauth, Meike Neuwohner, and Jens Vygen.

The k-Opt algorithm for the Traveling Salesman Problem has exponential running time for $k \geq 5$ – STEFAN HOUGARDY, UNIVERSITY OF BONN – Sophia Heimann, Hung P. Hoang Hoang

The k-Opt algorithm is a local search algorithm for the Traveling Salesman Problem. Starting with an initial tour, it iteratively replaces at most k edges in the tour with the same number of edges to obtain a better tour. We prove that for $k \geq 5$ the k-Opt algorithm can have exponential running time for all pivot rules.

Anytime Optimization Approaches for Online dial-a-ride problem – ELAHE AMIRI, POLYTECHNIQUE MONTRÉAL –

This study presents the development of a real-time dispatching system tailored for a ride-sharing service capable of accommodating a large volume of passengers and trips. The primary objective is to address the dynamic Dial-and-Ride Problem that ensures short travel time and service for each customer while minimizing waiting times. Departing from conventional methods that rely on fixed time intervals for batching requests, this research introduces a novel rolling-horizon strategy with adjustable time epochs. This approach combines the advantages of batch optimization with smaller batch sizes, ensuring rapid and effective dispatching. Two primal-based algorithms within a column generation framework are proposed and compared against a traditional column generation approach. A thorough evaluation is conducted using datasets derived from historical taxi trips in New York City, encompassing scenarios with up to 30,000 requests per hour.

Vehicle Routing with Time-Dependent Travel Times: Theory, Practice, and Benchmarks – DIRK MÜLLER, UNIVERSITY OF BONN – Jannis Blauth, Stephan Held, Niklas Schlömer, Vera Traub, Thorben Tröbst, Jens Vygen

We develop theoretical foundations and practical algorithms for vehicle routing with time-dependent travel times. We devise a faster algorithm to compute the pointwise minimum of a set of piecewise linear functions and a monotonicity-preserving variant of the Imai-Iri algorithm to approximate an arrival time function with fewer breakpoints. Next, we show how to evaluate insertion and deletion operations in tours efficiently and update the underlying data structure faster than previously known when a tour changes. Evaluating a tour also requires a scheduling step which is non-trivial in the presence of time windows and time-dependent travel times. We show how to perform this in linear time.

Based on these results, we develop a local search heuristic to solve real-world vehicle routing problems with various constraints efficiently and report experimental results on classical benchmarks, as well as on new benchmarks with time-dependent travel times that we created based on real-world data. Our results demonstrate the importance of considering time-dependent travel times in instances with tight time windows.

Improved Approximation Algorithms for the Joint Replenishment Problem with Outliers, and with Fairness Constraints – VARUN SURIYA-NARAYANA, CORNELL UNIVERSITY – Sivashankar Varun, Gollapudi Siddharth, David Shmoys

The joint replenishment problem (JRP) is a classical inventory management problem. We consider a natural generalization with outliers, where we are allowed to reject (that is, not service) a subset of demand points. In this paper, we are motivated by issues of fairness - if we do not serve all of the demands, we wish to “spread out the pain” in a balanced way among customers, communities, or any specified market segmentation. One approach is to constrain the rejections allowed, and to have separate bounds for each given customer. In our most general setting, we consider a set of C features, where each demand point has an associated rejection cost for each feature, and we have a given bound on the allowed rejection cost incurred in total for each feature. This generalizes a model of fairness introduced in earlier work on the Colorful k-Center problem in which (analogously) each demand point has a given color, and we bound the number of rejections of each color class. We give the first constant approximation algorithms for the fairness-constrained JRP with a constant number of features; specifically, we give a 2.86-approximation algorithm in this case. Even for the special case in which we bound the total (weighted) number of outliers, this performance guarantee improves upon bounds previously known for this case. Our approach is an LP-based algorithm that splits the instance into two subinstances. One is solved by a novel iterative rounding approach and the other by pipage-based rounding. The standard LP relaxation has an unbounded integrality gap, and hence another key element of our algorithm is to strengthen the relaxation by correctly guessing key attributes of the optimal solution, which are sufficiently concise, so that we can enumerate over all possible guesses in polynomial time - albeit exponential in C , the number of features.

Assortment Optimization with Visibility Constraints – MAROUANE IBN BRAHIM, CORNELL TECH – Théo Barré, Omar El Housni, Andrea Lodi, Danny Segev

Motivated by applications in e-retail and online advertising, we study the problem of assortment optimization under visibility constraints, that we refer to as APV. We are given a universe of substitutable products and a stream of T customers. The objective is to determine the optimal assortment of products to offer to each customer in order to maximize the total expected revenue, subject to the constraint that each product is required to be shown to a minimum number of customers. The minimum display requirement for each product is given exogenously and we refer to these constraints as visibility constraints. We assume that customer choices follow a Multinomial Logit model (MNL). We provide a characterization of the structure of the optimal assortments and present an efficient polynomial time algorithm for solving APV. To accomplish this, we introduce a novel function called the “expanded revenue” of an assortment and establish its supermodularity. Our algorithm takes advantage of this structural property. Additionally, we

demonstrate that APV can be formulated as a compact linear program. Next, we consider APV with cardinality constraints, which we prove to be strongly NP-hard and not admitting a Fully Polynomial Time Approximation Scheme (FPTAS), even in the special case where all the products have identical prices. Subsequently, we devise a Polynomial Time Approximation Scheme (PTAS) for APV under cardinality constraints with identical prices. Our algorithm starts by linearizing the objective function through a carefully crafted guessing procedure, then solves the linearized program, and finally randomly rounds the obtained solution to derive a near optimal solution for APV with cardinality constraints. We also examine the revenue loss resulting from the enforcement of visibility constraints, comparing it to the unconstrained version of the problem. To offset this loss, we propose a novel strategy to distribute the loss among the products subject to visibility constraints. Each vendor is charged an amount proportional to their product’s contribution to the revenue loss. Finally, we present the results of our numerical experiments providing illustration of the obtained outcomes.

Fairness, Randomization, and Approximation Algorithms – ARAVIND SRINIVASAN, UNIVERSITY OF MARYLAND –

We discuss approaches to fair allocations/service in combinatorial optimization, wherein approximation and randomization will both be crucial. Clustering problems will be among those studied through the lens of fairness; specifically, we will show how randomized-rounding algorithms for linear-programming relaxations can lead to useful (per-user) probabilistic guarantees for a range of problems.

Algorithmic Tools for Redistricting: Fairness via Analytics – DAVID SHMOYS, CORNELL UNIVERSITY –

The American winner-take-all congressional district system empowers politicians to engineer electoral outcomes by manipulating district boundaries. To date, computational solutions mostly focus on drawing unbiased maps by ignoring political and demographic input, and instead simply optimize for compactness and other related metrics. However, we maintain that this is a flawed approach because compactness and fairness are orthogonal qualities; to achieve a meaningful notion of fairness, one needs to model political and demographic considerations, using historical data.

We will discuss a series of papers that explore and develop this perspective. In the first (joint with Wes Gurnee), we present a scalable approach to explicitly optimize for arbitrary piecewise-linear definitions of fairness; this employs a stochastic hierarchical decomposition approach to produce an exponential number of distinct district plans that can be optimized via a standard set partitioning integer programming formulation. This enables a large-scale ensemble study of congressional districts, providing insights into the range of possible expected outcomes and the implications of this range on potential definitions of fairness. Further work extending this (joint with Julia Allen & Wes Gurnee), shows that many additional real-world constraints can be easily adapted in this framework (such as minimal county splits as was recently required in Alabama legislation in response to the US Supreme Court decision *Milligan v. Alabama*). In another paper (joint with Nikhil Garg, Wes Gurnee, and David Rothschild), we study the design of multi-

member districts (MMDs) in which each district elects multiple representatives, potentially through a non-winner-takes-all voting rule (as was proposed in H.R. 4000). We carry out large-scale analyses for the U.S. House of Representatives under MMDs with different social choice functions, under algorithmically generated maps optimized for either partisan benefit or proportionality. We find that with three-member districts using Single Transferable Vote, fairness-minded independent commissions can achieve proportional outcomes in every state (up to rounding), and this would significantly curtail the power of advantage-seeking partisans to gerrymander.

2e - FB295 - 512A - **Algorithms with Predictions** Chair: CLIFFORD STEIN 

Fair Secretaries with Unfair Predictions – ERIC BALKANSKI, COLUMBIA UNIVERSITY – Will Ma, Andreas Maggiori

Algorithms with predictions, also called learning-augmented algorithms, is a recent framework for decision-making that leverages the power of machine-learned predictions without making any assumption about their quality. The goal is to improve the algorithm's performance when predictions are accurate while maintaining acceptable guarantees when predictions are erroneous. Due to challenges with fairness in machine learning, a serious concern is that learning-augmented algorithms could leverage biased predictions and, as a result, make unfair decisions. We study the classical secretary problem in the learning-augmented setting, and prove that such a concern is indeed legitimate. Current learning-augmented algorithms for that problem are initially given predictions about the candidates' values and select a candidate that is guaranteed to have value that is a $\max(\Theta(1), 1 - O(\epsilon))$ approximation to the maximum value, where ϵ is the approximation error. However, these algorithms may pick the best candidate with vanishing probability. Our main result is a novel algorithm that achieves a $\max(\Theta(1), 1 - O(\epsilon))$ approximation while also guaranteeing that the best candidate is selected with constant probability. We also extend our result to the k -secretary problem and complement our theoretical analysis with experiments.

Incremental Topological Ordering and Cycle Detection with Predictions – BENJAMIN MOSELEY, CARNEGIE MELLON UNIVERSITY – Shikha Singh, Aidin Niaparasat, Mccauley Samuel

This paper leverages the framework of algorithms-with-predictions to design data structures for two fundamental dynamic graph problems: incremental topological ordering and cycle detection. In these problems, the input is a directed graph on n nodes, and the m edges arrive one by one. The data structure must maintain a topological ordering of the vertices at all times and detect if the newly inserted edge creates a cycle. The theoretically best worst-case algorithms for these problems have high update cost (polynomial in n and m). In practice, greedy heuristics (that recompute the solution from scratch each time) perform well but can have high update cost in the worst case.

In this paper, we bridge this gap by leveraging predictions to design a learned new data structure for the problems. Our data structure guarantees consistency, robustness, and smooth-

ness with respect to predictions – that is, it has the best possible running time under perfect predictions, never performs worse than the best-known worst-case methods, and its running time degrades smoothly with the prediction error. Moreover, we demonstrate empirically that predictions, learned from a very small training dataset, are sufficient to provide significant speed-ups on real datasets

Algorithms with Prediction Portfolios – THOMAS LAVASTIDA, THE UNIVERSITY OF TEXAS AT DALLAS – Michael Dinitz, Sungjin Im, Benjamin Moseley, Sergei Vassilvitskii

The research area of algorithms with predictions has seen recent success showing how to incorporate machine learning into algorithm design to improve performance when the predictions are correct, while retaining worst-case guarantees when they are not. Most previous work has assumed that the algorithm has access to a single predictor. However, in practice, there are many machine learning methods available, often with incomparable generalization guarantees, making it hard to pick a best method a priori. In this work we consider scenarios where multiple predictors are available to the algorithm and the question is how to best utilize them.

Ideally, we would like the algorithm's performance to depend on the quality of the best predictor. However, utilizing more predictions comes with a cost, since we now have to identify which prediction is the best. We study the use of multiple predictors for a number of fundamental problems, including matching, load balancing, and non-clairvoyant scheduling, which have been well-studied in the single predictor setting. For each of these problems we introduce new algorithms that take advantage of multiple predictors, and prove bounds on the resulting performance.

Energy-Efficient Scheduling with Predictions – CLIFFORD STEIN, COLUMBIA UNIVERSITY – Eric Balkanski, Noemie Perivier, Hao Ting Wei

An important goal of modern scheduling systems is to efficiently manage power usage. In energy-efficient scheduling, the operating system controls the speed at which a machine is processing jobs with the dual objective of minimizing energy consumption and optimizing the quality of service cost of the resulting schedule. Since machine-learned predictions about future requests can often be learned from historical data, a recent line of work on learning-augmented algorithms aims to achieve improved performance guarantees by leveraging predictions. In particular, for energy-efficient scheduling, Bamas et. al. and Antoniadis et. al. designed algorithms with predictions for the energy minimization with deadlines problem and achieved an improved competitive ratio when the prediction error is small while also maintaining worst-case bounds even when the prediction error is arbitrarily large. In this talk, we consider a general setting for energy-efficient scheduling and provide a flexible learning-augmented algorithmic framework that takes as input an offline and an online algorithm for the desired energy-efficient scheduling problem. We show that, when the prediction error is small, this framework gives improved competitive ratios for many different energy-efficient scheduling problems, including energy minimization with deadlines, while also maintaining a bounded competitive ratio regardless of the prediction error. Finally, we empirically demonstrate that this framework achieves an improved performance on real

and synthetic datasets.

2f - MB344 - 513C - **Modeling Software Chair:**
DAN STEFFY 

FICO®Xpress: New and Augmented Interfaces for Modeling Optimization Problems

– GREGOR HENDEL, FAIR ISAAC DEUTSCHLAND GMBH –

A fast, scalable, and easy-to-use programming interface (API) is an essential component when working with commercial high-end optimization solvers. As such, FICO®Xpress has been offering interfaces for an ever-growing set of programming languages (C, C++, Java, C#, Python, R).

This talk gives an overview of these interfaces, with a particular focus on the most recent enhancements. We will focus on the APIs for Java, C#, and Python. The new, augmented FICO®Xpress APIs for Java and C# leverage modern language constructs. They have been designed as an object-oriented layer ensuring a memory-efficient and reliable experience for the user. With the augmented API, the use of solver features such as callbacks becomes easier, and it gives access to the full set of problem types available with FICO®Xpress.

Practical guidelines for model improvement and reformulation

– DAN STEFFY, GUROBI OPTIMIZATION –

I will share insights and lessons learned from helping Gurobi customers from a wide range of industries adjust their optimization models to improve solver performance and numerical behavior. We will look at the challenges that we see most often in LP, MIP and MINLP models, and discuss our approach and typical recommendations to help address them. We will also consider some well-known modeling “rules of thumb” and discuss how applicable they are in 2024.

What’s new in JuMP

– MILES LUBIN, HUDSON RIVER TRADING – Oscar Dowson, Joaquim Dias Garcia, Joey Huchette, Benoît Legat, Juan Pablo Vielma

JuMP is an algebraic modeling language for mathematical optimization embedded in the Julia programming language. JuMP combines the convenience of a high-level, solver-independent modeling syntax with great performance, making it suitable for a range of applications spanning from teaching to continent-scale energy system modeling. In this talk, we present major developments since the 1.0 release two years ago. These include a new nonlinear programming interface and support for problem classes like multi-objective optimization, constraint programming, nonlinear mixed-complementarity problems, rational linear programming, and others.

OptiChat: An AI Assistant for Explaining Optimization Models Powered by LLM

– CAN LI, PURDUE UNIVERSITY –

Mathematical optimization has wide applications in real-world decision-making problems such as aircraft crew scheduling, smart grid operation, and health care. One of the primary barriers to deploying optimization models in practice is the challenge of helping practitioners understand and interpret such models. The research to help non-experts understand optimization models has a long history. In the 1980s, Harvey Greenberg developed the ANALYZE software, an expert sys-

tem that relies on a specific syntax tailored for optimization problems. The downside of expert systems is that they still require a fair amount of domain knowledge and optimization background. In the pioneering paper, Greenberg insightfully foresaw that “Ideally, one would like to move toward a natural language query system where the descriptions are resident with the model in a way that permits more automation to obtain answers.” However, to the best of our knowledge, such a natural language-based system envisioned by Greenberg has not been developed due to the lack of a general-purpose language model. Recently, large language models (LLMs) such as GPT-4 and LLaMA have achieved remarkable success in diverse applications, opening new ways of developing software for analyzing optimization models. In this presentation, we develop a natural language-based system, OptiChat, for analyzing (mixed-integer) linear programs. It is an agent-based autonomous system that coordinates the user, the optimization model, the optimization solver, and the LLM. The non-expert user is interested in understanding the optimization model written in an algebraic modeling language that has symbolic expressions of the decision variables, the input parameters, the constraints, and the objective. OptiChat can provide high-level model descriptions, diagnose and resolve infeasibility for infeasible problems, explain why a solution is optimal, and answer any interactive questions of the user such as “what if” questions.

2f - MB932 - 513B - **Computational Discrete and Integer Optimization 1 Chair:** FRED GARDI 

Optimizing Airport Selection for Electric Aircraft Operations Across Europe

– FREDERIC KRONER, TU BRAUNSCHWEIG – Imke Joormann, Lisa Marie Manke, Andreas Tillmann

Due to the ambitious emission targets the international community as well as individual states have set, the development and deployment of electric aircraft will be further pursued in the upcoming years. However, it is not possible to replace conventional with electric aircraft without equipping airports with the necessary infrastructure. Since extensions of airports are extremely costly, we present a model that determines the optimal selection of airports in Europe for the development of infrastructure to support the operation of electric aircraft, while accommodating the routing of the passenger demand and staying within emission targets. To identify the most beneficial flight routes based on the electrification status of each airport the passenger demand is defined as a multi-commodity flow and we apply a path-based formulation within a column generation framework. We have gathered real-world data, which are utilized to analyze the model with regards to solvability in relation to the instance size and the established emission limits.

A speed-up for Helsgaun’s TSP heuristic by relaxing the positive gain criterion

– SABRINA AMMANN, TU BRAUNSCHWEIG – Birte Ostermann, Sebastian Stiller, Timo De Wolff

The Traveling Salesman Problem (TSP) is one of the most extensively researched and widely applied combinatorial optimization problems. It is NP-hard even in the symmetric and metric case, mostly considered here. The Lin-Kernighan-Helsgaun (LKH) heuristic has led to many best known solu-

tions in particular for very large instances and often finds an optimal solution for small instances. We propose variations of LKH that perform significantly faster on large instances. LKH repeatedly improves an initially random tour by exchanging edges along alternating circles. Thereby, it respects several criteria designed to quickly find alternating circles that give a feasible improvement of the tour. We relax one of those criteria - the positive gain criterion - carefully leading to improvement steps hitherto undiscovered by LKH. We confirm this improvement experimentally via extensive simulations on various benchmark libraries for TSP. Our computational study shows that for large instances our method is on average 13% faster than the latest version of LKH while the quality of solutions stays virtually the same.

A New Labeling Approach for a Special Case of the Multicommodity Flow Problem – LISA MARIE MANKE, TU BRAUNSCHWEIG, INSTITUTE OF AUTOMOTIVE MANAGEMENT AND INDUSTRIAL PRODUCTION – Imke Joormann

The standard multicommodity flow problem is defined on a directed graph and possesses origin-destination pairs as commodities. In this presentation, we consider an additional flow with no predefined origin or destination. This is applicable to an air transportation network, where it should be possible to refuel an aircraft for more than one flight regardless of the origin-destination passenger flow, for example when introducing hydrogen-powered aircraft with a lower flight range. Interesting instances of this problem are relatively large and cannot be solved to optimality via generic approaches. Hence, we introduce a new labeling procedure that is based on Dijkstra's shortest path algorithm. The most relevant changes are shifting the labels to the edges and expanding them so that each commodity has its own label. The additional flow, which does not inherit a predefined origin or destination, is then incorporated as precomputed data represented by an additional capacity from each vertex to all other vertices. Computational results showed a speed-up compared to the generic approach.

Hexaly, a new kind of global optimization solver – FRED GARDI, HEXALY –

Hexaly is a new kind of global optimization solver. Its modeling interface is nonlinear and set-oriented. It also supports user-coded functions, thus enabling black-box optimization and simulation optimization. Thus, Hexaly APIs unify modeling concepts from mixed-linear, nonlinear, and constraint programming. Under the hood, Hexaly combines various exact and heuristic optimization methods: spatial branch-and-bound, simplex methods, interior point methods, augmented Lagrangian methods, automatic Dantzig-Wolfe reformulation, column and row generation, propagation methods, local search, direct search, population-based methods, and surrogate modeling techniques.

Regarding performance benchmarks, Hexaly distinguishes itself against the leading solvers in the market, like Gurobi, IBM Cplex, and Google OR Tools, by delivering fast and scalable solutions to problems in the spaces of Supply Chain and Workforce Management like Routing, Scheduling, Packing, Clustering, Matching, Assignment, and Location problems. For example, on notoriously hard problems like the Pickup and Delivery Problem with Time Windows or Flexible Job Shop Scheduling with Setup Times, Hexaly delivers

solutions with a gap to the best solutions known in the literature smaller than 1% in a few minutes of running times on a basic computer.

2f - MC339 - 513E - Extensions of Integer Programming Chair: MARIANNA DE SANTIS

On the accurate detection of the Pareto frontier for bi-objective mixed integer linear problems – MARIANNA DE SANTIS, UNIVERSITÀ DI FIRENZE – Lavinia Amorosi

We propose a criterion space search algorithm for bi-objective mixed integer linear programming problems. The Pareto frontier of these problems can have a complex structure, as it can include isolated points, open, half-open and closed line segments. Therefore, its exact detection is a hard computational task. Our algorithm handles this issue by alternating the resolution of single objective mixed integer linear problems with bi-objective linear ones. The performance of the algorithm is improved using suitably defined cuts and related strategies. Under specific assumptions, we can prove that the exact Pareto frontier can be detected in a finite number of iterations. Experimental results on a test-bed of instances and a comparison with available solvers is presented, showing the notably good performance of our approach in terms of accuracy of the Pareto frontier detected.

PySCIPOpt-ML: Embedding Trained Machine Learning Models into Mixed-Integer Programs – MARK TURNER, ZUSE INSTITUTE BERLIN – Antonia Chmiela, Thorsten Koch, Michael Winkler

A standard tool for modelling real-world optimisation problems is mixed-integer programming (MIP). However, for many of these problems there is either incomplete information describing variable relations, or the relations between variables are highly complex. To overcome both these hurdles, machine learning (ML) models are often used and embedded in the MIP as surrogate models to represent these relations. Due to the large amount of available ML frameworks, formulating ML models into MIPs is highly non-trivial. We present PySCIPOpt-ML, a tool for the automatic MIP formulation of trained ML models, allowing easy integration of ML constraints into MIPs. In this presentation we will discuss the MIP formulations of the embedded ML models, and present computational results on the scale of ML models that can easily be embedded. In addition, we introduce a library of MIP instances with embedded ML constraints.

QUBO.jl: A tale of implementation and benchmarking of a Quantum Optimization Ecosystem in Julia – PEDRO MACIEL XAVIER, FEDERAL UNIVERSITY OF RIO DE JANEIRO – Pedro Ripper, Joaquim Dias Garcia, Nelson Maculan, David Bernal Neira

We present QUBO.jl, an end-to-end Julia package for working with QUBO (Quadratic Unconstrained Binary Optimization) instances. This tool aims to convert a broad range of JuMP problems for straightforward application in many physics and physics-inspired solution methods whose standard optimization form is equivalent to the QUBO. These methods include quantum annealing, quantum gate-circuit optimization

algorithms (Quantum Optimization Alternating Ansatz, Variational Quantum Eigensolver), other hardware-accelerated platforms, such as Coherent Ising Machines and Simulated Bifurcation Machines, and more traditional methods such as simulated annealing. Besides working with reformulations, QUBO.jl allows its users to interface with the aforementioned hardware, sending QUBO models in various file formats and retrieving results for subsequent analysis. QUBO.jl was written as a JuMP / MathOptInterface (MOI) layer that automatically maps between the input and output frames, thus providing a smooth modeling experience. As a result of the infrastructure developed to work with QUBO formulations and solvers, we present QUBOLib, a software library and instance collection designed to enable the continuous benchmarking of QUBO solvers as they evolve.

2f - TA269 - 513E - **Solver Software I** Chair: TIMO BERTHOLD 

What's New in FICO Xpress Solver? – TIMO BERTHOLD, FAIR ISAAC DEUTSCHLAND GMBH –

In this presentation, we will give an overview of the latest enhancements, the newest features, and the most recent performance improvements in the FICO Xpress Solver for mixed-integer linear and nonlinear optimization problems. These include new heuristics, cutting and branching techniques, an augmented API, and updates to our global MINLP solver.

What's new in Gurobi 11 – MICHAEL WINKLER, GUROBI OPTIMIZATION –

We will give an overview about recent developments and performance improvements within Gurobi 11. Gurobi provides an optimization solver software package, free for academics, able to tackle numerous types of as mathematically formulated optimization problems. Its capabilities include solving linear programming (LP), quadratic programming (QP), quadratically constrained programming (QCP), mixed-integer linear and non-linear programming (MIP incl. MILP, MIQP, MICQP, MIQCQP and MINLP) to proven optimality. Additionally, multi-objectives, scenarios, and callback support, an irreducible infeasible subsystem (IIS) finder, a model instance tuner, a lazy constraint feature and a variety of general purpose heuristics are integrated into Gurobi.

Enhancing the HiGHS MIP solver – JULIAN HALL, UNIVERSITY OF EDINBURGH –

The practical performance of the HiGHS MIP solver has led the open-source world since 2022. This talk will discuss the recent performance enhancements and new features that have extended the MIP solver's capabilities. An overview of use cases and other recent developments in HiGHS will also be given.

Recent Progress in the Cardinal Optimizer – NILS CHRISTIAN KEMPKE, CARDIAL OPERATIONS/COPT GMBH –

In this talk, we present the recent advances in the Cardinal Optimizer (COPT). We highlight some key computational and theoretical improvements and evaluate their impact on the performance of our MIP solver. We provide an overview of the performance numbers of the latest COPT release for all supported problem classes.

2f - TB340 - 513E - **Solver Software II** Chair: PHILIPP CHRISTOPHEL 

A modeling language-based approach to automatically recommend first-order optimization methods. – SOFIANE TANJI, UCLOUVAIN – François Glineur

Since the advent of modern computational mathematics, the literature on optimization algorithms has been ever-growing, and a wide range of methods are now available. Usually, published algorithms are shown to be applicable to given templates of optimization problems, and are often accompanied by proofs characterizing their convergence rates.

There exist many variations of such templates. Indeed, a template may describe the objective as a single function or a sum of functions with different characteristics (say, smoothness, convexity, existence of a computable proximal operator, etc.), accompanied by constraints that can also be described in various ways (e.g. linear constraints, feasible set with a computable projection operator, functional constraints, etc.).

Hence, for a given instance of an optimization problem, it is not always easy to identify which templates it can match, especially if one wants to consider equivalent reformulations of the problem. This makes the task of choosing an optimization method tedious for the practitioner.

In this talk, we present a modeling language for mathematical programming that aims at describing oracle-based optimization problem formulations. Using this language, we propose a framework that automatically checks whether a user-provided optimization problem fits a known template. The framework, implemented as a Python package, relies on an extensive library of optimization methods, described using the above-mentioned language, with their associated known convergence rate for one or several templates, with the goal of automatically ranking applicable optimization methods according to their worst-case theoretical guarantees. The framework also handles some reformulation techniques, allowing more corresponding templates to be potentially identified.

Symmetry detection in Mixed-Integer Conic Programming – SVEN WIESE, MOSEK APS –

We describe how to detect symmetry in Mixed-Integer Conic Programming. We present our framework as an extension to the Mixed-Integer Linear Programming case, resulting in a relatively small formal overhead. To do so, we introduce the concept of symmetry labelings for cones and study their properties, with a focus on what labelings are useful in practice. We also report on computational experiments performing symmetry detection in the optimization software package MOSEK on a series of Mixed-Integer Conic Programming problems.

Hybridizing combinatorial heuristics and continuous optimization methods for Mixed-Integer Programming – JULIEN DARLAY, HEXALY –

Hexaly is a new kind of global optimization solver that combines exact and heuristic methods. Heuristics like local search are known for delivering quality primal solutions to large-scale combinatorial optimization problems in short running times. However, these are difficult to apply when continuous decisions are involved, as in inventory routing, nuclear power plant outage scheduling, portfolio optimization with limited assets,

and sparse regression. To address these problems, Hexaly automatically decomposes them into a pure combinatorial problem and a continuous subproblem parametrized by the combinatorial decisions. The combinatorial space is explored by applying local search techniques. Once a new feasible solution is found for integer decisions, an exact continuous algorithm is launched to determine continuous decisions optimally: simplex methods when the continuous subproblem is linear and interior point methods when it is nonlinear. Sensitivity analysis is performed to guide the heuristic search in the combinatorial space. The talk will present the overall approach and illustrate it on the problems previously mentioned.

Recent Advances in the SAS Linear and Mixed-integer Optimization Solvers – PHILIPP CHRISTOPHEL, SAS INSTITUTE GMBH –

In this talk we present the most recent improvements to the SAS linear and mixed-integer optimization solvers and discuss future developments. Topics include simplex implementational aspects, the latest ideas for crossover, new presolve techniques and improvement for existing techniques, and the latest developments in cut generation, selection and management.

2f - WA281 - 513E - New Algorithms for Large-Scale Optimization Chair: DANIEL ALOISE

Dynamic constraint aggregation for solving the minimum sum-of-squares clustering problem – ANTONIO M. SUDOSO, SAPIENZA UNIVERSITY OF ROME – Daniel Aloise

The minimum sum-of-squares clustering problem (MSSC), also known as k-means clustering, refers to the problem of partitioning n data points into k clusters, with the objective of minimizing the total sum of squared Euclidean distances between each point and the center of its assigned cluster. We propose an efficient algorithm for solving large-scale MSSC instances in two-dimensional Euclidean space, which combines column generation (CG) with dynamic constraint aggregation (DCA) to effectively reduce the number of constraints considered in the CG master problem. DCA reduces degeneracy by utilizing an aggregated restricted master problem obtained from an aggregating partition of the set of data points into disjoint clusters. Our computational results demonstrate that the proposed algorithm significantly outperforms existing CG-based approaches available in the literature.

Generalized integral simplex – ALPHA SALIOU BARRY, POLYTECHNIQUE MONTRÉAL –

The set partitioning problem can be generalized. One generalization is to consider that certain tasks must be covered several times. So it's no longer a partition that's sought, but a set of parts that verify the right number of coverages.

Unlike the set partitioning problem, the graph of integer solutions is no longer connected when the possible arcs are simplex pivots. However, in order to restore the connectedness of the graph, it is possible to add artificial variables locally, which allow new pivots to be made while preserving the integrity of the solution.

We give priority to pivots that keep the solution integer. When this is no longer possible, we solve a linear program to obtain a non-degenerate sequence of simplex pivots to perform

in order to improve the solution. When this improved solution is still not integer, in some cases the pivot sequence allows to add an artificial column associated with an integer artificial point. The linear program is solved again from this point to find the remaining pivots to be performed in order to obtain a better integer solution. If integrity is still not reached, a branching is performed.

We present results for large pairing problems (around 2,000 constraints and 300,000 variables) and show that the method is around 2 to 5 times faster than traditional methods (CPLEX) and that in many cases, the addition of the artificial variable effectively allows moves on integer points not directly accessible by pivot.

Integral Simplex for the Set Partitioning Problem – FRANCOIS SOUMIS, POLYTECHNIQUE MONTREAL UNIVERSITY –

For the set partitioning problem, the existence of a sequence of integer solutions with non-increasing costs produced by simplex pivots, permitting to reach optimality was proved by Balas, Padberg 1972. Finding the following term in the sequence can be very time-consuming due to degeneracy. When there is no entering variable that leads to a better integer solution, we use a linear programming sub-problem to find a group of variables to enter the basis to obtain an improved solution which is integer most of the time. When this solution is not integer, we solve a small integer programming problem to obtain a better integer solution. We first, present results for large-scale problems (with up to 500 000 variables) for which optimal integer solutions are obtained, most of the time, without any branching. We also, present the integration of this new approach with column generation and parallelism to solve large-scale bus driver and airline pilot problems. Problems with many thousand tasks are solved 5 to 10 faster than with branch and price based on CPLEX.

New complementary problem formulation for the improved primal simplex – YOUSSEF EMINE, GERAD –

The primal simplex algorithm is still one of the most used algorithms by the operations research community. It moves from basis to adjacent one until optimality. The number of bases can be very huge, even exponential, due to degeneracy or when we have to go through all of the extreme points very close to each other. The improved primal simplex algorithm (IPS) is efficient against degeneracy but when there is no degeneracy, it behaves exactly as a primal simplex and consequently, it may suffer from the same limitations.

We present a new formulation of the complementary problem, i.e., the auxiliary subproblem used by the improved primal simplex to find descent directions, that guarantees a significant improvement of the objective value at each iteration until we reach an ϵ -approximation of the optimal value. We prove that the number of needed directions is polynomial.

2f - WB341 - 513E - Computational Mixed Integer Programming Chair: PIERRE BONAMI

Generalized Resolution Conflict Analysis in MIP Solvers – GIONI MEXI, ZUSE INSTITUTE BERLIN –

Conflict analysis has been successfully generalized from Boolean satisfiability (SAT) solving to mixed integer programming (MIP) solvers, but although MIP solvers operate with general linear inequalities, the conflict analysis in MIP has been limited to reasoning with the more restricted class of clausal constraints. This is in contrast to how conflict analysis is performed in so-called pseudo-Boolean solving, where solvers can reason directly with 0-1 integer linear inequalities rather than with clausal constraints extracted from such inequalities. In this work, we investigate how pseudo-Boolean conflict analysis can be generalized for MIP and integrated in MIP solving. Phrased in MIP terminology, conflict analysis can be understood as a sequence of linear combinations and cuts. We leverage this perspective to design a new conflict analysis algorithm based on mixed integer rounding (MIR) cuts, which theoretically dominates the state-of-the-art Chvátal-Gomory-based method in pseudo-Boolean solving. We implement our new conflict analysis algorithm in the open-source MIP solver SCIP and evaluate it on a large and diverse set of MIP instances from MIPLIB 2017.

Recent additions to the collection of cutting planes in Gurobi – PIERRE BONAMI, GUROBI – Cutting planes are one of the fundamental pieces for efficiently solving mixed integer linear optimization problems. Gurobi Optimizer 11 implements 21 different types of cutting plane separation algorithms.

We will discuss some of the most recent techniques that were added to our algorithm. We will briefly sketch the ideas behind them and present some computational results on a large set of models.

Leveraging Discounted Pseudocosts in MILP for Enhanced Branching Strategies – KRUNAL KISHOR PATEL, CERC, POLYTECHNIQUE MONTRÉAL –

In this talk, we introduce the concept of discounted pseudocosts, inspired by the notion of discounted total reward in reinforcement learning, and explore their application in mixed-integer linear programming (MILP). Pseudocosts traditionally estimate the change in the objective function value resulting from changes in variable bounds during the branch-and-bound process. By integrating ideas from reinforcement learning, we propose a novel approach incorporating a forward-looking perspective into pseudocost estimation.

We present the motivation behind discounted pseudocosts and discuss how they represent the anticipated reward for branching after one level of exploration in the MILP problem space. Leveraging this concept, we aim to guide the branch-and-bound algorithm towards more promising regions of the solution space, potentially improving its efficiency and convergence rate.

Furthermore, we present the results of our initial experiments conducted to evaluate the effectiveness of discounted pseudocosts in MILP solvers. Through a series of computational experiments on MIPLIB 2017 benchmark instances, we demonstrate the impact of incorporating discounted pseudocosts on the performance of the branch-and-bound algorithm. Our findings suggest that this approach yields promising results, showcasing its potential to enhance branching strategies and accelerate the solution process for challenging MILP problems.

Variable Fixing for Integer Linear Programming via Objective Rotation – LUKAS SCHÜRMANN, UNIVERSITY OF BONN –

Domain propagation of variables throughout the solution process of a branch-and-bound algorithm is a powerful tool that is typically implemented in sophisticated mixed integer programming solvers. The variety of techniques spreads over combinatorial propagation, constraint propagation, and reduced cost strengthening. We present an approach that extends the idea of the latter by rotating the objective to retrieve more information about the underlying polytope. This leads to an increased number of fixed variables compared to simply applying reduced cost strengthening, especially since it is applicable to both, basic and non-basic variables. A careful analysis of the parameters used by our algorithm allows us to detect dead-ends early on and therefore restricts the running time. This approach can be seen as a heuristic look-ahead branching and is related to the estimation of the strong branching scores. Our computational study on the MIPLIB instances shows that the integration of this technique in a branch-and-bound framework can lead to an improvement of the running time. Additionally, we observe a large benefit for the Vehicle Routing Problem with Heterogeneous Time Windows.

2f - WC266 - 513E - **Certified Optimization**
Chair: JONAD PULAJ 

An SMT-based checker for VIPR certificates of MILP results – JONAD PULAJ, DAVIDSON COLLEGE – Hammurabi Mendes, Hien Tran, Haoze Wu, Daniel Zhou

Verification of results returned from mixed-integer linear programming (MILP) solvers is particularly desirable in the context of highly sensitive applications such as hardware verification, compiler optimization, or experimental mathematics. We present an approach for validating proof certificates in the VIPR format (generated for MILP problems) via Satisfiability Modulo Theories (SMT) solvers. The resulting checker is based on the equivalence between the correctness of a VIPR certificate and the satisfiability of a formula in the theory of linear/integer real arithmetic encoded as an SMT instance. We demonstrate the effectiveness of this approach and its variations by evaluating them on existing benchmark instances.

Certificates of Correctness for Presolving Binary Programs – ALEXANDER HOEN, ZUSE INSTITUTE BERLIN – Andy Oertel, Ambros Gleixner, Jakob Nordström

It is well known that reformulating the original problem can be crucial for the performance of mixed-integer programming (MIP) solvers. To ensure correctness, all transformations must preserve the feasibility status and optimal value of the problem, but there is currently no established methodology to express and verify the equivalence of two mixed-integer programs. In this work, we take a first step in this direction by showing how the correctness of MIP presolve reductions on 0-1 integer linear programs can be certified by using (and suitably extending) the VeriPB tool for pseudo-Boolean proof logging. Our experimental evaluation on both decision and optimization instances demonstrates the computational viability of the approach and leads to suggestions for future revisions of the proof

format that will help to reduce the verbosity of the certificates and to accelerate the certification and verification process further.

The Four-Color Ramsey Multiplicity of Triangles – ALDO KIEM, ZUSE INSTITUTE BERLIN, TECHNISCHE UNIVERSITÄT BERLIN –

Goodman established in 1959 that asymptotically $1/4$ of all triangles in a two-coloring of the edges of a complete graph are monochromatic. By using the semidefinite programming method of Flag Algebras, we generalize this result to four-colored complete graphs resulting in the tight bound of $1/256$. This is the largest Flag Algebra based computation performed to date. We will introduce the basic theory of Flag Algebra semidefinite programming and highlight our new methods such as symmetry decompositions and exact rational LP solving that we used to make this computation possible.

2f - ThA343 - 513E - **Decomposition Methods**
Chair: MARCO LÜBBECKE 

A scalable Master Problem in Dual Decomposition of Stochastic MIP – KIBAEK KIM, ARGONNE NATIONAL LABORATORY –

The dual decomposition of stochastic mixed-integer programming (SMIP) has been instrumental for providing a scalable solution of SMIP with integer variables and large scenarios by taking the Lagrangian relaxation of the original problem with respect to the nonanticipativity constraints. The Lagrangian relaxation results in the master problem of solving the Lagrangian dual problem and the subproblems that are defined for each scenario and can be solved in parallel. While the bundle method and its variants have been successfully used to solve the master problem, the master problem steps can significantly degrade the parallel efficiency as is typically solved in serial. We present a novel algorithm for solving the master problem in parallel by adopting the alternating direction method of multiplier algorithm. We implement the algorithm and present the numerical results from various test instances.

A Row-and-Column Generation Approach to Min-Sum Job-Shop Scheduling – MARCO LÜBBECKE, RWTH AACHEN UNIVERSITY – Sarah Roth

We report on an exact approach to min-sum job-shop scheduling which builds on and combines ingredients that have been around in the literature for some time. The first is the idea of using an extended formulation for the problem, in this case essentially a network flow model on a time-indexed graph. Secondly, such formulations can be decomposed (leading to Lagrangian or Dantzig-Wolfe based approaches), either with so-called job-level or machine-level subproblems, or both, leading to strong dual bounds. We use machine-level subproblems. Thirdly, working with time-indexed models directly is often computationally limited because of their sheer size, and the so-called row-and-column generation paradigm suits this well in order to dynamically generate the network. And this is it (well, with some branching, cutting, and primal heuristics), and it worked for us. For several instances (minimum weighted completion time objective) from the literature we can prove optimality for the first time, or improve the best known dual or the primal bounds, respectively. These computational results are almost ten years old, but we never reported on them. In

the end we look at a practical application from scheduling in a health care environment, which we came across only recently.

Decomposition method for a capacitated multi-vehicle covering tour problem with intermediate facilities – ANTOINE LEGRAIN, POLYTECHNIQUE MONTREAL – Vera Fischer, David Schindl

This talk considers a waste collection problem with intermediate disposal facilities to accommodate the usage of smaller, but more sustainable vehicles, with less capacity than the traditional waste collection trucks. The optimization problem consists in determining the locations to place the collection points and the routes of a capacitated collection vehicle that visits these locations. We first present a mixed-integer linear programming formulation that exploits the sparsity of the road network. To efficiently solve practical instances, we propose a solution method that decomposes the problem into a set covering problem to select the locations and a capacitated vehicle routing problem with intermediate facilities to determine the routes and price the set covering solution. We apply column generation to solve this routing problem and present a novel approach for the set covering problem that exploits the properties of the problem to efficiently find a set cover by using a graph theoretical approach. We compare our method with the set construction method and the heuristic Benders decomposition approach presented in a previous work. Computational experiments on instances derived from real-life data confirm the difficulty of our problem and show the superior performance of the developed decomposition method with respect to the number of best solutions and the average gaps to the best solution.

Integer programming column generation: Accelerating branch-and-price for set covering, packing, and partitioning problems – ELINA RÖNNBERG, LINKÖPING UNIVERSITY – Stephen J. Maher

Large-neighbourhood search (LNS) heuristics are important mathematical programming techniques that search for primal feasible solutions by solving an auxiliary problem with a restricted feasible region. Extending such powerful generic LNS heuristics to the branch-and-price context is inherently challenging. The most prominent challenges arise from the fact that in branch-and-price algorithms, columns are generated with the sole aim to solve linear programming relaxations. Hence, the ability to form integer feasible solutions is not considered during the generation of columns. Without any changes to the standard pricing schemes, the potential of deploying generic LNS heuristics within a branch-and-price procedure is severely limited.

We propose a matheuristic, based on an LNS heuristic framework, where the novelty is a customised pricing scheme for generating columns to solve an auxiliary problem. The theoretical foundation for this pricing scheme is a set of optimality conditions for integer programs. From this foundation, a column generation strategy is developed for finding columns that are likely to be of use in high-quality primal feasible solutions for the original problem. The proposed matheuristic is implemented in the generic branch-price-and-cut solver GCG. On a broad test set comprising classical block diagonal structured instances and general instances from the MIPLIB 2017 Collection, the computational results show a significant improvement

to the solving performance of GCG.

2f - ThB342 - 513E - **Combinatorial Optimization and Applications** Chair: JAAP PEDERSEN ↩

Branch and Price for the Length-Constrained Cycle Partition Problem – MOHAMMED GHANAM, ZUSE INSTITUTE BERLIN – Ambros Gleixner, Gioni Mexi, Edward Lam

The length-constrained cycle partition problem (LCCP) is a graph optimization problem in which a set of nodes must be partitioned into a minimum number of cycles. Every node is associated with a critical time and the length of every cycle must not exceed the critical time of any node in the cycle. We formulate LCCP as a set partitioning model and solve it using an exact branch-and-price approach. We use a dynamic programming-based pricing algorithm to generate improving cycles, exploiting the particular structure of the pricing problem for efficient bidirectional search and symmetry breaking. Computational results show that the LP relaxation of the set partitioning model produces strong dual bounds and our branch-and-price method improves significantly over the state of the art. It is able to solve closed instances in a fraction of the previously needed time and closes 13 previously unsolved instances, one of which has 76 nodes, a notable improvement over the previous limit of 52 nodes.

Cutting planes for the maximum k-vertex cover problem – SHENG JIE CHEN, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES – Wei Kun Chen, Yu Hong Dai

The Maximum k-Vertex Cover (Max k-VC) problem involves finding a subset of at most k vertices in an undirected graph G to maximize the number of covered edges. The need for the Max k-VC problems arises in various fields, such as system risk management, network robustness measurement, and airplane seating assignment. In this research, we focus on developing cutting plane strategies for solving this problem exactly. We begin by studying the polytope associated with the integer programming formulation of the Max k-VC problem. In particular, we prove that certain facet-defining inequalities for the polytope can be derived from those for the projected and low-dimensional polytopes obtained by restricting G to its subgraphs. Based on this insight, we propose three families of strong valid inequalities for the Max k-VC problem, utilizing different structured subgraphs including odd cycles, cliques, and generalized stars. We also provide necessary and sufficient conditions for these inequalities to be facet-defining, and discuss the corresponding separation problems. Finally, we integrate the proposed valid inequalities as cutting planes into the branch-and-bound algorithm of CPLEX, a widely used state-of-the-art integer programming solver. Extensive computational results demonstrate the effectiveness of incorporating these cutting planes for solving the Max k-VC problems, in terms of strengthening the linear programming relaxation and improving the solution efficiency.

Quota Steiner Tree Problem and its application on Wind Farm Planning – JAAP PEDERSEN, ZUSE INSTITUTE BERLIN –

In this talk, we present an approach for the integrated layout

and cable routing problem of onshore wind farm planning using the Quota Steiner tree problem (QSTP). We present a transformation of the QSTP that significantly outperforms standard out-of-the-box MIP solvers in the context of regional wind farm planning. Secondly, we show that solving integrated layout and cable routing problem can lead to up to 20% cost reduction compared to a sequential approach. Finally, we explore the possibilities of extending the QSTP formulation by relevant technical constraints such that our approach becomes applicable to the design of a single wind farm.

A New Class of Compact Formulations for Vehicle Routing Problems – JULIAN YARKONY, OPTYM – Udayan Mandal, Louis Martin Rousseau, Amelia Regan

This paper introduces a novel compact mixed integer linear programming (MILP) formulation and a discretization discovery-based solution approach for the Vehicle Routing Problem with Time Windows (VRPTW). We aim to solve the optimization problem efficiently by constraining the linear programming (LP) solutions to use only flows corresponding to time and capacity-feasible routes that are locally elementary (prohibiting cycles of customers localized in space).

We employ a discretization discovery algorithm to refine the LP relaxation iteratively. This iterative process alternates between two steps: (1) increasing time/capacity/elementarity enforcement to increase the LP objective, albeit at the expense of increased complexity (more variables and constraints), and (2) decreasing enforcement without decreasing the LP objective to reduce complexity. This iterative approach ensures we produce an LP relaxation that closely approximates the optimal MILP objective with minimal complexity, facilitating an efficient solution via an off-the-shelf MILP solver.

The effectiveness of our method is demonstrated through empirical evaluations on classical VRPTW instances. We showcase the efficiency of solving the final MILP and multiple iterations of LP relaxations, highlighting the decreased integrality gap of the final LP relaxation. We believe that our approach holds promise for addressing a wide range of routing problems within and beyond the VRPTW domain.

2f - ThC270 - 513E - **Augmentation Methods and Dynamic Programming** Chair: DOMINIC YANG ↩

Augmentation search for integer programmes defined on polyhedra – TOLGA BEKTAS, UNIVERSITY OF LIVERPOOL –

This talk will describe a primal search algorithm to optimise an integer programme defined over a polyhedron, wherein the search is conducted on the lattice described by the linear constraints of the model. Search directions are derived in the spirit of Graver bases and are extracted dynamically using a feasibility-seeking residual problem. The algorithm also allows for the use of Graver-base specific valid inequalities based on the specific properties of the coefficient matrix of the constraints of the integer programme. Computational results will be presented to show the potential of the algorithm, particularly on 0-1 programming formulations with nonlinear objective functions.

Algebra-Inspired Randomized Augmentation

Methods in Combinatorial Optimization – DOMINIC YANG, ARGONNE NATIONAL LABORATORY – In recent years, significant work has been focused on methods for exploring and sampling from and optimizing over the set of feasible points of an integer program, known as a fiber. These methods focus on producing a basis of moves which enable transitions between feasible points and facilitate Markov Chain Monte Carlo methods over the fiber. In this talk, we will address integer programs with very large bases for which computation using the whole set is intractable and analyze the performance of augmentation methods for optimization when using a small sample of the basis. An intelligent sampling approach could both potentially enable the solution of problems of a much higher dimension and also enable the use of larger bases with more attractive properties for smaller problems (e.g., Graver bases for which a greedy augmentation method produces an optimal solution). In particular, we will study a specific total variation-regularized optimization problem and present theoretical and experimental results for this problem.

A lifted-space dynamic programming algorithm for the Cubic Knapsack Problem (CKP) – IBRAHIM DAN DIJE, ESG UQAM – Franklin Djeumou Fomeni, Leandro Coelho

The Cubic Knapsack Problem (CKP) is respectively a generalization of Quadratic Knapsack Problem (QKP) and Knapsack Problem (KP). CKP is known as combinatorial optimization problem in which, a cubic function of binary variables is maximizing subject to one dimension knapsack constraint. It has many applications in biology, project selection, capital budgeting problem, etc. The CKP is NP-hard problem in the strong sense, because it's a generalization of QKP that is NP-hard, then there is no polynomial exact algorithm that can solve it. In this work, we introduced a dynamic programming heuristic in the space of lifted variable to solve CKP by finding a near optimal solution. The solution of dynamic programming heuristic is compared with exact solution for and with other heuristic find in the literature.

2f - FA224 - 513C - **Tightening MILPs by Reformulation** Chair: DOMINIK KAMP

Augmented subproblem grinding for the lot-type design problem – JÖRG RAMBAU, UNIVERSITY OF BAYREUTH – Miriam Kießling, Sascha Kurz
We consider a fashion discounter supplying its many branches with integral multiples from a set of available lot-types. The lot-type design problem (LDP) [Gaul et al. 2010] asks: which (integral) multiples of which (integral) lot-types (assortments of sizes) should be supplied to a set of branches in order to meet a (fractional) expected demand as closely as possible? There is a compact LDP-model; however, its integral gap is so large that it cannot be solved for most practical instances. On the other hand, the tightest LDP-model known so far [Gaul et al. 2010] can have billions of variables. (For 12 different sizes, reasonable for lingerie or children's clothing, there are 1,159,533,584 different lot-types, if we assume at most 5 items of each size and a total number of items in a lot-type between 12 and 30.) Thus, not for all instances the tight model can be fed to a computer statically. We show how the tight model, which can be interpreted as a Dantzig-Wolfe decomposition

(a standard-reformulation to tighten ILP-models) of the compact model, can be solved by Augmented Subproblem Grinding, which is a new Branch-and-Cut-and-Price variant, well-suited for tight models. It enforces a binary branch-and-price tree. One branch consists of a single promising subproblem that is solved immediately ("ground off"). The other branch is tightened by a new constraint ("augmented") that excludes all solutions consisting of only known columns (a non-standard reformulation). The theoretical core component is the identification of a characteristic lifting of dual variables in the reduced master problem for all the constraints that would emerge with new variables. Computational results on real-world instances as well as on randomized stress-tests show that for the tight LDP-model Augmented Subproblem Grinding speeds up the solution by a factor of more than 100 on average compared to the solution of the static model and solves instances (up to 9,867,114,720,320 variables) that cannot be solved statically at all.

An acceleration scheme for column generation in production-maintenance scheduling problems – JOÃO DIONÍSIO, CMUP/FCUP – Ambros Gleixner, João Pedro Pedroso, Ksenia Bestuzheva

Given a group of machines, the objective is to decide each machine's production schedule, and when maintenance actions should be performed. Our focus is on variants where the (independent) machines are complex, with many interacting components, and their joint production must satisfy a predetermined demand. Furthermore, the production load of a machine impacts its need for maintenance.

A direct formulation can be very slow, especially when there are many identical machines, which increases the model's symmetry. The machines being complex also leads to some very dense blocks in the otherwise sparse constraint matrix.

A reasonable initial approach is a Dantzig-Wolfe reformulation, given the model's block-diagonal structure. While it can be faster than the compact formulation when there are many identical machines, it remains prohibitively slow when that is not the case, as its pricing problems remain difficult mixed-integer (nonlinear) programs.

This problem can be thought of as two connected subproblems. In one, we need an assignment to the continuous production variables (a production pattern), and in the other an assignment to the integer maintenance variables (a maintenance pattern).

Thinking of these two subproblems, decomposing the original problem in a way that handles them separately might be beneficial. We would have an easy subproblem that produces production patterns and a hard subproblem that produces maintenance patterns. This comes with additional difficulties, as it requires simultaneous column-and-row generation, and the introduction of a third, more difficult, subproblem (which is essentially the same as the subproblem of the original reformulation). The hope is that by only calling a subproblem whenever its easier peers have exhausted all negative reduced cost columns, the overall algorithm is faster. As in our problem one maintenance pattern can be feasibly completed by many production patterns, we expect that the easy subproblems will be called much more often.

This idea of generating partial columns in different sub-

problems can also be applied in some other Dantzig-Wolfe decompositions.

A new approach for inter-sample avoidance using an integer polytope formulation – JOHANNES SCHMIDT, BRANDENBURG UNIVERSITY OF TECHNOLOGY COTTBUS SENFTENBERG – Armin Fügenschuh

To model the physics of the used vehicles in trajectory planning problems, Newton's equations of motion must be incorporated, discretized by numerical solution schemes. Thus, the position of the vehicle is in the model known only at a discrete set of points. Due to this lack of information, corner cutting can occur, i.e., the vehicle can abbreviate through an obstacle if it remains outside at every evaluation point again. Increasing the number of discrete time steps cannot resolve this problem completely and leads to higher computation times. To obtain always realistic solutions, corner cutting must be restricted, yielding so called inter-sample avoidance methods. We present a new approach to ensure inter-sample avoidance for polyhedral obstacles. The hyperplanes describing the obstacle divide the mission space into a set of subareas which can be encoded by binary vectors. These are used to ensure feasible movements without requiring more evaluation points or additional variables. We give a procedure to generate an integer polytope containing the feasible movements and derive an upper bound on the number of necessary constraints to describe it. The new approach is qualitatively tested against other inter-sample avoidance methods from the literature and its computational efficiency is studied in numerical experiments, using a mixed-integer linear program for trajectory planning of unmanned aerial vehicles and the state-of-the-art solver Gurobi.

Dynamic Demand Propagation in Guaranteed Service Models – DOMINIK KAMP, UNIVERSITY OF BAYREUTH, CHAIR OF ECONOMATHEMATICS – Jörg Rambau

In supply chain management a challenging task is to find an efficient allocation of safety stocks which are sufficient to maintain a reliable supply process. In the presence of multiple supply options at each stage this becomes even more interesting because the selection of suppliers influences the internally propagated demands which the upstream safety stock requirements rely on. The Stochastic Guaranteed Service Model with Demand Propagation (SGSM-DP) models the safety stock requirements in a multi-echelon inventory network by bilinear constraints on replenishment times and propagated demands bounding the base stock levels at each node in multiple static demand scenarios. Applying a standard big-M-linearization leads to a mixed integer linear program which is compact but intractable in practice due to a large integrality gap. Regarding the flow conservation structure of the demand propagation, an equivalent formulation (F-SGSM-DP) was developed acting on a time-expanded flow network. Although larger in size, solvers show a consistently better performance caused by the almost tight reformulation. Using this approach, realistic instances could be solved to optimality for the first time and evaluated in a dynamic simulation environment. With these insights, further improvement potential on dynamic demand characteristics was revealed. A novel generalization based on periodic time cycles (P-SGSM-DP) corresponding to periodic demand-sequence scenarios makes it finally possible to reach

cost-efficient results on a real-world spare-parts distribution system with highly dynamic demand (rare over time but substantial in amount) compared to the approximate approaches usually applied in practice.

2f - FA933 - 513E - Computational Discrete and Integer Optimization 2 Chair: QUENTIN LOU-VEAUX

Column Elimination – ANTHONY KARAHALIOS, CARNEGIE MELLON UNIVERSITY – Willem Jan Van Hoeve

Column elimination is an exact algorithm to solve discrete optimization problems via a 'column' formulation in which variables represent a combinatorial structure such as a machine schedule or truck route. Column elimination works with a relaxed set of columns, from which infeasible ones are iteratively eliminated. As the total number of columns can typically be exponentially large, we use relaxed decision diagrams to compactly represent and manipulate the set of columns. In this talk, we provide an introduction to the column elimination method and present recent algorithmic improvements resulting in competitive performance on large-scale vehicle routing problems. Specifically, our approach closes a large vehicle routing benchmark instance with 1,000 locations for the first time.

Sequential Mixed-Integer Programming for Matrix Sparsification – ANDREAS TILLMANN, TU BRAUNSCHWEIG –

Sparsity of matrices can be exploited in various applications, e.g., to speed up the solution of linear systems or in second-order optimization algorithms. This motivates considering the matrix sparsification problem (MS), in which we seek an equivalent representation of a given matrix with as few nonzero entries as possible. While MS and the equivalent sparse nullspace basis problem have been of interest for decades, the only known solution approaches appear to be heuristics or methods relying on restrictive assumptions that may be unlikely to hold in practice. We propose a sequential mixed-integer linear programming algorithm for matrix sparsification that leverages relations to matroid theory, discuss connections to a machine learning task known as dictionary learning, and show some numerical results.

An improved weighting local search for large-scale binary integer programs – SHUNJI UMETANI, RECRUIT CO., LTD. –

We have developed a 4-flip neighborhood local search (LS) algorithm for binary integer programs (BIP) that incorporates an efficient incremental evaluation of solutions and an adaptive control of penalty weights. The proposed algorithm attained good performance for large-scale set covering and partitioning problems consisting of homogenous constraints. However, we also realized two defectives of the proposed algorithm that (1) it does not well perform for some combinatorial optimization problems such as bin packing problems consisting of heterogeneous constraints, and (2) the 4-flip neighborhood search performs less effectively than expected while taking much computation time. To overcome them, we first introduce an improved extraction method of variable associations that identifies promising pairs of flipping variables in the neighborhood

search. We focus on that many of BIPs in real applications consist of typical types of constraints such as knapsack, set packing, covering and partitioning, bin packing, variable bounding and so on. We then evaluate promising pairs of flipping variables by computing inner product of partial column vectors associated to each typical constraint type. We also introduce iterated local search (ILS) algorithm instead of the 4-flip neighborhood search for diversification of the search. The ILS is a simple and effective framework that repeats LS from different initial solutions generated by perturbing (a.k.a. kick operation) the best solution obtained so far. It is known that a naive ILS often fails to escape local minimum solutions due to rewinding to the initial solutions. However, it is difficult to design an effective kick operation deriving few distinctive features of BIP. We accordingly develop a generalized diversification procedure of ILS that prevents rewinding to the kick operation by freezing the flipped variables. Based on the proposed algorithm, we develop an open heuristic solver for BIP and report our computational results for a wide variety of test instances.

Sensitivity analysis for linear change of the constraint matrix of an LP – QUENTIN LOUVEAUX, UNIVERSITY OF LIÈGE – Guillaume Derval, Damien Ernst, Bardhyl Miftari

We focus on the parametric linear problem $f(l) = \min cx$ such that $(A + lD)x \leq b$, where D is a matrix modifying the coefficients of the constraint matrix A with the parameter l , which can be used to analyze the impact of multiple linear changes in the constraint matrix. As a new approach to the problem, we derive five bounds which allow us to avoid recomputing the problem for numerous l and provide guarantees on the objective function's behavior. We discuss an iterative algorithm that use these bounds to compute an approximation to the original function within a given error threshold. We analyze the performance of the bounds on toy and real-world problems and demonstrate the effectiveness of the approach.

2f - FB268 - 513E - **Integer Programming Applications in Mobility and Traffic** Chair: NIELS LINDNER

On the Split Closure of the Periodic Timetabling Polytope – NIELS LINDNER, FREIE UNIVERSITÄT BERLIN –

The Periodic Event Scheduling Problem (PESP) is the central mathematical tool for periodic timetable optimization in public transport. PESP can be formulated in several ways as a MILP with typically general integer variables. However, even for rather small instances, good dual bounds are notoriously hard to obtain. We therefore investigate the split closure of the periodic timetabling polytope and show that split inequalities are identical with so-called flip inequalities. While split inequalities are a general mixed-integer programming technique, flip inequalities are defined in purely combinatorial terms. We prove that pseudo-polynomial-time separation of split/flip cuts is best possible unless $P = NP$, but also observe that the complexity becomes linear-time if the cycle defining the flip inequality is fixed. Moreover, introducing mixed-integer-compatible maps, we compare the split closures of different formulations, and show that reformulation or binarization by subdivision do not lead to stronger split closures. Finally, we estimate computationally how much of the opti-

mality gap of the instances of the benchmark library PESPLib can be closed exclusively by split cuts, and that a combination of a heuristic and an exact cut generation approach is able to provide better dual bounds for five instances.

Integrated Periodic Event Scheduling and Infrastructure Assignment: approaches via Cyclic Orders and Graph Matchings – ENRICO BORTOLETTO, ZUSE INSTITUTE BERLIN –

We present an approach that extends the well-known Periodic Event Scheduling Problem (PESP) by integrating safety constraints and platform assignment options, offering a comprehensive and efficient solution for real-world transportation networks. While the traditional PESP framework focuses solely on scheduling events with periodicity constraints, it often overlooks the crucial aspect of infrastructure utilization. To address this limitation, we showcase two mixed-integer programming ideas. The first makes use of constraints based on cyclic orders of conflicting activities, and we show how this formulation is able to deal with simple instances and requirements. Secondly, we present a formulation based on a matching problem, capable of modelling much more complex situations, even allowing more efficient infrastructure capacity utilization. To demonstrate the practical efficacy of our approach, we evaluate it on real-world instances, including the S-Bahn Berlin network.

A Flexible Model for Integrated Line Planning and Periodic Timetabling with Track Choice in Practice – BERENIKE MASING, ZUSE INSTITUTE BERLIN – Niels Lindner, Christian Liebchen

The line plan, timetable, and vehicle schedule are key components of public transportation planning, and high-quality service can only be provided if they are well coordinated. We present the Integrated Line Planning and Periodic Timetabling with Track Choice Model, which addresses each of these aspects simultaneously and can cover different modeling requirements, depending on the detail of planning intended. We introduce the extended infrastructure-aware event-activity network (EAN) as the core object: Based on an implicitly given line plan, it encodes all potential vehicle paths that a line could be realised as, taking into account the underlying infrastructure.

We obtain a mixed-integer linear optimization formulation by generalising the periodic event scheduling problem (PESP) to accommodate the extended infrastructure-aware EAN. Vehicle routing decisions are made by an imposed binary flow on the EAN, which in turn activates traditional PESP constraints. The line planning aspect is included by frequency coverage constraints.

We demonstrate applicability of our model by using it for re-planning purposes in the context of construction sites: Infrastructure becomes unavailable and the regular service cannot be upheld as capacity issues on tracks arise. The goal is then to recover as much of the regular service as possible while satisfying safety requirements. We show that our model is well suited to this purpose on multiple real-life construction site scenarios in the network of S-Bahn Berlin.

It becomes evident that the solving process benefits from good starting solutions. As has been observed for other PESP-related problems in the past, we demonstrate a SAT (Boolean Satisfiability) approach to be effective as a primal heuristic.

The Approximation Error Caused by Linearizing Charging Behavior for Electric Vehicle Scheduling Problems – FABIAN LÖBEL, ZUSE INSTITUTE BERLIN – Ralf Borndörfer, Steffen Weider

Public transport operators are increasingly adopting electric battery-powered bus fleets with a significant impact on associated vehicle scheduling problems. Available driving ranges are still limited compared to combustion-powered counterparts, such that buses may have to recharge in-service. This needs to be taken into account within vehicle schedules via additional downtime and possibly detours. The relationship between charge states and charge duration is non-linear due to the complex chemical processes within batteries, however, electric vehicle scheduling models usually rely on a linearized description in order to access the powerful tools offered by mixed-integer linear programming.

In this talk, we examine the approximation error this introduces and its impact on solution feasibility. We highlight worst-case behavior, study piecewise linear charge curves considered state-of-the-art in the literature, and propose an alternative.

2g - WA243 - 512F - Constraints and Machine Learning Chair: CHENBO SHI

Optimal counterfactual explanations for k-Nearest Neighbors using Mathematical Optimization and Constraint Programming – RICARDO FUKASAWA, UNIVERSITY OF WATERLOO – Claudio Contardo, Louis Martin Rousseau, Thibaut Vidal

Within the topic of explainable AI, counterfactual explanations to classifiers have received significant recent attention. We study counterfactual explanations that try to explain why a data point received an undesirable classification by providing the closest data point that would have received a desirable one. Within the context of one the simplest and most popular classification models – k-nearest neighbors (k-NN) – the solution to such optimal counterfactual explanation is still very challenging computationally. In this work, we present techniques that significantly improve the computational times to find such optimal counterfactual explanations for k-NN.

MORBDD: Multiobjective Restricted Binary Decision Diagrams by Learning to Sparsify – RAHUL PATEL, UNIVERSITY OF TORONTO – Elias Khalil, David Bergman

In multicriteria decision-making, a user seeks a set of non-dominated solutions to a (constrained) multiobjective optimization problem, the so-called Pareto frontier. In this work, we seek to bring a state-of-the-art method for exact multiobjective integer linear programming into the heuristic realm. We focus on binary decision diagrams (BDDs) which first construct a graph that represents all feasible solutions to the problem and then traverse the graph to extract the Pareto frontier. Because the Pareto frontier may be exponentially large, enumerating it over the BDD can be time-consuming. We explore how restricted BDDs, which have already been shown to be effective as heuristics for single-objective problems, can be adapted to multiobjective optimization through the use of machine learning (ML). MORBDD, our ML-based BDD sparsifier, first trains a binary classifier to eliminate BDD nodes

that are unlikely to contribute to Pareto solutions, then post-processes the sparse BDD to ensure its connectivity via optimization. Experimental results on multiobjective knapsack problems show that MORBDD is highly effective at producing very small restricted BDDs with excellent approximation quality, outperforming width-limited restricted BDDs and the well-known evolutionary algorithm NSGA-II.

Learning Lagrangian Multipliers for the Travelling Salesman Problem – LOUIS MARTIN ROUSSEAU, POLYTECHNIQUE MONTRÉAL – Augustin Parjadis, Aaron Ferber, Quentin Cappart, Bistra Dilkina

Lagrangian relaxation is a versatile mathematical technique employed to relax constraints in an optimization problem, enabling the generation of dual bounds to prove the optimality of feasible solutions and the design of efficient propagators in constraint programming (such as the weighted circuit constraint). However, the conventional process of deriving Lagrangian multipliers (e.g., using subgradient methods) is often computationally intensive, limiting its practicality for large-scale or time-sensitive problems. To address this challenge, we propose an innovative unsupervised learning approach that harnesses the capabilities of graph neural networks to exploit the problem structure, aiming to generate accurate Lagrangian multipliers efficiently. We apply this technique to the well-known Held-Karp Lagrangian relaxation for the traveling salesman problem. The core idea is to predict accurate Lagrangian multipliers and to employ them as a warm start for generating Held-Karp relaxation bounds. These bounds are subsequently utilized to enhance the filtering process carried out by branch-and-bound algorithms. In contrast to much of the existing literature, which primarily focuses on finding feasible solutions, our approach operates on the dual side, demonstrating that learning can also accelerate the proof of optimality. We conduct experiments across various distributions of the metric traveling salesman problem, considering instances with up to 200 cities. The results illustrate that our approach can improve the filtering level of the weighted circuit global constraint, reduce the optimality gap by a factor of two for unsolved instances up to a timeout, and reduce the execution time for solved instances by 10%.

Constraint Learning to Define Trust Regions in Optimization Over Pre-Trained Predictive Models – CHENBO SHI, UNIVERSITY OF CONNECTICUT – Mohsen Emadikhia, Leonardo Lozano, David Bergman

There is a recent proliferation of research on the integration of machine learning and optimization. One expansive area within this research stream is optimization over pre-trained predictive models, which proposes the use of pre-trained predictive models as surrogates for uncertain or highly complex objective functions. In this setting, features of the predictive models become decision variables in the optimization problem. Despite a recent surge in publications in this area, only a few papers note the importance of incorporating trust-region considerations in this decision-making pipeline, i.e., enforcing solutions to be similar to the data used to train the predictive models. Without such constraints, the evaluation of the predictive model at solutions obtained from optimization cannot be trusted and the practicality of the solutions may be unreasonable. In this paper, we provide an overview of the approaches appearing in the literature to construct a trust region, and propose three alter-

native approaches. Our numerical evaluation highlights that trust-region constraints learned through our newly proposed approaches, compare favorably with previously suggested approaches, both in terms of solution quality and computational time.

2g - WC241 - 514B - CPAIOR at ISMP: Mathematical Programming, Constraint Programming, and Hybrid Methods Chair: LAURENT MICHEL 

A reinterpretation of cutting planes – JOHN HOOKER, CARNEGIE MELLON UNIVERSITY – Danial Davarnia, Atefeh Rajabalizadeh

The constraint programming community has intensively studied the concept of consistency as a means to reduce backtracking. We show how cutting planes can achieve a form of consistency by excluding inconsistent partial assignments – that is, partial assignments that cannot be extended to a full feasible assignment. This leads to a reinterpretation of cutting planes as a consistency maintenance mechanism rather than a means of strengthening bounds. We find that cuts selected for this task can, at least in some contexts, reduce branching search more effectively than traditional separating cuts. In fact, traditional branch-and-cut algorithms can benefit from this feature of cutting planes even while using them to tighten bounds. Consistency concepts may therefore offer a new perspective on integer programming that can lead to a better understanding of what makes branching methods work.

A Constraint-Programming-Based Branch-and-Check Approach for Stochastic Consistent Home Health Care Routing and Scheduling Problem – SEYEDE SAEDE HOSSEINI, POLYTECHNIQUE MONTREAL – Yossiri Adulyasak, Louis Martin Rousseau

We study a multi-period home health care routing and scheduling problem (HHCRSP), which requires consistent assignment of care workers and visit times for each patient. This problem is similar to the consistent vehicle routing problem (ConVRP), known as NP-hard. The uncertainty in travel and service times is a real-world assumption, which makes finding consistent routes and schedules even more challenging. This paper proposes a novel approach that integrates stochastic travel and service times into the deterministic consistent HHCRSP through a chance-constraint approach. We model these chance-constraints in two ways: a scenario-based and an extreme value theory-based (ETV-based) approximation. To solve practical-sized instances of this problem, we decompose it into a master problem (MP) and a set of subproblems (SP). We use branch-and-check (B&Ch) to solve the MP through mixed integer linear programming and the SPs through constraint programming. Our computational experiments demonstrate that our stochastic models, especially the ETV-based approximation model, outperform the deterministic model in terms of service levels at a low cost. We study the efficiency of this approach under different types of uncertainty.

C++DDOpt : A Modeling and Solving Interface for Decision Diagram-based Optimization. – LAURENT MICHEL, U. OF CONNECTICUT – Willem Jan Van Hoeve

This talk introduces C++DDOpt, a C++ library offering both

a modeling and a solving interface for solving combinatorial optimization problems using decision diagram-based optimization solvers. Currently this optimization technology is accessible via lower-level programming languages which limits easy modeling experimentation and wider adoption. C++DDOpt allows to represent state-based models in an intuitive manner. These models are specifications which are used to automatically compile relaxed and restricted decision diagrams, as well as the search process. We demonstrate the functionality of C++DDOpt on a variety of classic combinatorial optimization problems and demonstrate its effectiveness.

2g - FB242 - 512G - Learning and Explaining Constraints Chair: NADJIB LAZAAR 

Domain-Independent Dynamic Programming – J. CHRISTOPHER BECK, UNIVERSITY OF TORONTO – Ryo Kuroiwa

Domain-independent dynamic programming (DIDP) is a general-purpose system for modeling and solving combinatorial optimization problems based on dynamic programming (DP). The aim of the DIDP system is to declaratively model problems in a DP-based formulation and then solve the formulation with a general-purpose solver. To date, we have implemented a number of solvers based on heuristic search as studied extensively in AI. Our numerical results have shown that DIDP outperforms state-of-the-art MIP and CP approaches on a variety of combinatorial problem benchmarks including traveling salesperson with time windows, orienteering, single machine scheduling, and assembly line balancing. In this presentation, we provide an overview of the DIDP modeling and solving approaches and recent numerical results. We then provide a detailed examination of the models and solver behavior for selected problem classes to elucidate problem characteristics that lend themselves to strong DIDP performance.

Lazy MIP Solving with Methods from SAT – ALEXANDER TESCH, BOOL AI –

In the past, there have been many approaches to integrate ideas from SAT into MIP solving and vice versa. In particular, conflict analysis, originally developed for SAT, has been established as a standard in modern MIP solvers to compute stronger cutting planes from conflicts that were derived during the search. An important difference between conflict analysis in MIP and SAT is, however, that MIP uses conflict analysis mostly as an add-on to improve the LP-relaxation during a branch-and-bound procedure while for SAT it constitutes the actual solving method, also called Conflict-Driven Clause learning (CDCL). This raises the question if we could similarly develop a MIP solver that is built exclusively on lazy methods rather than branch-and-bound, i.e. on successively learning constraints from conflicts that occur during the search. In this talk, we will present a first version of such a lazy MIP solver which we developed as an extension to our constraint programming (CP) engine LazyCP which previously used pseudo-Boolean methods to solve finite-domain CPs. We will present some first computational results and discuss some observed strengths, weaknesses and potential improvements to the “lazy approach” of solving MIPs.

Automated Streamlining for Constraint Satisfaction Problems. – IAN MIGUEL, UNIVERSITY OF ST ANDREWS – Nguyen Dang, Ozgur Akgun, Patrick

Spracklen

Before a combinatorial search problem can be solved with constraint solving or a related technique it must first be modelled: a set of variables chosen to represent the decisions we must make in solving the problem, and a set of constraints specifying allowed combinations of variable assignments. Since the search spaces involved can be vast, an effective model of a problem of interest has a very significant impact on the solving process. Steps such as breaking symmetry and/or adding implied constraints to aid the solver in making inferences and therefore detecting dead ends earlier in search can be used to improve an initial model. If performance remains unsatisfactory, a more aggressive step that can be taken is the addition of streamliner constraints, which are not guaranteed to be sound but are designed to focus effort on a highly restricted but promising portion of the search space. When the concept was originally introduced, generating effective streamliner constraints was a manual, difficult and time-consuming task. This talk will present an automated approach to the generation, search and selection of streamliner portfolios to produce a substantial reduction in search across a diversity of problems.

Query-Based Constraint Acquisition – NADJIB LAZAAR, LIRMM, UNIVERSITY OF MONTPELLIER, CNRS –

Constraint Acquisition (CA) is a fundamental symbolic learning approach which learns concepts under the form of constraint networks. By classifying positive and negative instances, CA progressively eliminates constraints from a bias containing all possible constraints between variables. When a CA process raises queries to an external oracle, the learning approach becomes active. In recent years, query-based CA has become a popular method to acquire knowledge under the form of constraint networks, to handle partial answers over positive and negative instances, to deal with uncertainty in data instances, and acquire qualitative constraints. This talk surveys existing query-based CA methods and presents its main application areas which range from robotics to software engineering. Deploying query-based CA for explaining the results of classification models developed with opaque machine learning models is an important perspective of this symbolic learning approach.

3a - MB107 - 515B - **Multistage Stochastic Programs** Chair: ZHIPING CHEN

A Bayesian approach to data-driven multi-stage stochastic optimization – ZHIPING CHEN, XI'AN JIAOTONG UNIVERSITY – Wentao Ma

Aimed at sufficiently utilizing available data and prior distribution information, we introduce a data-driven Bayesian-type approach to solve multi-stage convex stochastic optimization, which can easily cope with the uncertainty about data process's distributions and their inter-stage dependence. To unravel the properties of the proposed multi-stage Bayesian expectation optimization (BEO) problem, we establish the consistency of optimal value functions and solutions. Two kinds of algorithms are designed for the numerical solution of single-stage and multi-stage BEO problems, respectively. A queuing system and a multi-stage inventory problem are adopted to numerically demonstrate the advantages and practicality of the

new framework and corresponding solution methods, compared with the usual formulations and solution methods for stochastic optimization problems.

Multistage stochastic optimization of an elementary hydrogen infrastructure – RAIAN NOUFEL LEFGOUM, ÉCOLE DES PONTS PARISTECH CERMICS – Jean Philippe Chancelier, Michel De Lara, Pierre Carpentier, Sezin Afsar

Hydrogen displays promising features for decarbonization industry, transportation and building sectors. The desired transition towards a hydrogen economy requires hydrogen costs to come down, through optimal choices of infrastructure design and operation. Most of the literature models hydrogen problems in a deterministic manner and solves them using linear programming. In this talk, we present an approach based on multistage stochastic optimization, mixing design choices with operational decisions taken on an hourly basis. We consider an elementary hydrogen infrastructure which consists of an electrolyzer, a compressor and a storage to serve a transportation demand. This infrastructure is powered by three sources of energy (on site photovoltaics, renewable electricity through power purchase agreement, power grid) with their own characteristics. The modelling of the electrolyzer covers its different functioning modes and the nonlinear relation between the production of hydrogen and the electricity consumption. The optimization problem is to minimize the expected operational costs over a week by making hourly decisions in an uncertain context. This involves managing the electrolyzer's load and mode, quantifying the electricity supplied by each source, and determining the amount of hydrogen extracted from the storage to satisfy the demand. Renewable energy sources are emphasized in the hydrogen production process to ensure eligibility for a subsidy, which is awarded if the proportion of nonrenewable electricity usage stays under a predetermined threshold. We consider uncertainties affecting on site photovoltaics production and hydrogen demand. All decisions are taken prior to the occurrence of uncertainties, except for the decision regarding the grid electricity supply, which is treated as a recourse following the realization of these uncertainties. In this work, we formulate a multistage stochastic optimization problem, assuming stagewise independence of the noise. Building on this assumption, we develop decomposition algorithms based on dynamic programming. We present numerical results for a given infrastructure design. Then, we consider various combinations of infrastructure designs and their subsequent optimal operation. With this, we discuss the optimal sizing of equipment, especially the sensitivity of electrolyzer and storage designs to the uncertainties.

The Kelly Strategy - Modifications to Address Market Changes and Investor Preferences – LEONARD MACLEAN, DALHOUSIE UNIVERSITY – Yonggan Zhao

The log-optimal or Kelly strategy has many desirable properties such as maximizing the asymptotic long run growth of capital and minimizing the expected time to financial goals. The strategy is aggressive and can have considerable short run risk, where risk is the chance of falling short of a downside threshold. The risk with the Kelly strategy is exacerbated if the distribution of future returns is incorrectly specified. In this paper we consider the risk associated with the error in the estimation of returns, and the risk associated with constraints on the log-

optimal investment model. It is assumed that the volatility of asset returns is affected by global market forces and individual asset trading characteristics. An improvement in the estimation of returns is achieved with a hierarchical model, where the uncertain financial market is segmented into regimes with the dynamics of the stochastic regime process being Markovian. The driving force for between regime changes is indicators which characterize economic states. Within a regime the asset prices are lognormal, and distribution parameters depend on the uncertain regimes. To accommodate investor risk aversion, the log-optimal model obtains maximum growth while staying above a specified downside wealth threshold with high probability, and shortfalls below the threshold are penalized in the objective. The enhanced asset pricing model and the controlled growth decision model reduce the risk of investments in uncertain markets. The solution from the decision model has the familiar Kelly formula, with weights to aggregate regime distribution parameters and an investment fraction to protect against shortfalls. The Kelly format reduces the optimization to determining the regime weights and investment fraction in the Kelly portfolio. The methods are illustrated with investments in ETF's over time.

Multi-stage distributionally robust optimization with Bayesian learning moments uncertainty – JIA LIU, XI'AN JIAOTONG UNIVERSITY – Zhiping Chen

In this talk, we introduce multi-stage distributionally robust portfolio selection (MSDRPS) problems with reinforcement learning moments uncertainty. With uncertainty sets predetermined by fixed reference moments, we derive the closed-form solution of the MSDRPS problem. With multiple prior reference moments, we construct multiple prior uncertainty sets and solve the MSDRPS problem in a dynamic programming approach. When the reference moments information is learnable in a Bayesian rule, we construct two Bayesian reinforcement learning MSDRPS schemes with real investment feedback or reference investment feedback. We propose a two-level decomposition framework to solve the Bayesian reinforcement learning MSDRPS problem using reference investment feedback. Finally, numerical results on newly listed and sub-new stocks examine the practicality and superior performance of the Bayesian reinforcement learning strategy in MSDRPS problems. (this is a joint work with Zhiping Chen, Roujia Li and Zongben Xu)

3a - MC934 - 515A - **Continuous Stochastic Programming 1** Chair: JUNYU ZHANG

A Sampling-based Progressive Hedging Algorithm for Stochastic Programming – DI ZHANG, UNIVERSITY OF SOUTHERN CALIFORNIA – Yihang Zhang, Suvrajeet Sen

The progressive hedging algorithm (PHA) is a cornerstone in tackling large-scale stochastic programming challenges. However, its traditional implementation is hindered by several limitations, including the requirement to solve all scenario subproblems in each iteration, reliance on an explicit probability distribution, and a convergence process that is highly sensitive to the choice of penalty parameters. This paper introduces a sampling-based PHA that aims to overcome these limitations. Our approach employs a dynamic selection process for the number of scenario subproblems solved per iteration.

It incorporates adaptive sequential sampling for determining sample sizes, stochastic conjugate subgradient methods for direction finding, and a line-search technique to update the dual variables. Experimental results demonstrate that this novel algorithm not only addresses the bottlenecks of the conventional PHA but also potentially surpasses its scalability, representing a substantial improvement in the field of stochastic programming.

Generalized adaptive partition-based method for two-stage stochastic linear programs: geometric oracle and analysis – VINCENT LECLERE, ENPC –

Adaptive Partition-based Methods (APM) are numerical methods that solve, in particular, two-stage stochastic linear problems (2SLP). We say that a partition of the uncertainty space is *adapted* to the current first stage control $\tilde{x}$ if we can aggregate scenarios while conserving the true value of the expected recourse cost at $\tilde{x}$. The core idea of APM is to iteratively constructs an *adapted* partition to all past tentative first stage controls.

Relying on the normal fan of the dual admissible set, we give a necessary and sufficient condition for a partition to be adapted even for non-finite distribution, and provide a geometric method to obtain an adapted partition. Further, by showing the connection between APM and the L-shaped algorithm, we prove convergence and complexity bounds of the APM methods. We discussed the fixed recourse case, with elements to forgo this assumption.

Stochastic Bregman Proximal Gradient Method Revisited: Kernel Conditioning and Painless Variance Reduction – JUNYU ZHANG, NATIONAL UNIVERSITY OF SINGAPORE –

We investigate Bregman proximal gradient (BPG) methods for solving nonconvex composite stochastic optimization problems. Instead of the standard gradient Lipschitz continuity (GLC) assumption, the objective function only satisfies a smooth-adaptability assumption w.r.t. some kernel function. An in-depth analysis of the stationarity measure is made in this paper, where we reveal an interesting fact that the widely adopted Bregman proximal gradient mapping in the existing works may not correctly depict the near stationarity of the solutions. To resolve this issue, a new Bregman proximal gradient mapping is proposed and analyzed in this paper. Second, a thorough analysis is made on the sample complexities of the stochastic Bregman proximal gradient methods under both the old and the newly proposed gradient mappings are analyzed. Note that the absence of GLC disables the standard analysis of the stochastic variance reduction techniques, existing stochastic BPG methods only obtain an $O(\epsilon^{-2})$ sample complexity under the old gradient mapping, we show that such a limitation in the existing analyses mainly comes from the insufficient exploitation of the kernel's properties. By proposing a new kernel-conditioning regularity assumption on the kernel, we show that a simple epoch bound mechanism is enough to enable all the existing variance reduction techniques for stochastic BPG methods. Combined with a novel probabilistic argument, we show that there is a high probability event conditioning on which $O(\sqrt{n}\epsilon^{-1})$ sample complexity for the finite-sum stochastic setting can be derived for the old gradient mapping. Moreover, with a novel adaptive step size

control mechanism, we also show an $O(\sqrt{n}L_h(X)\epsilon^{-1})$ complexity under the new gradient mapping. Worst case instances are designed to show the tightness of the analysis.

3a - TA115 - 515A - Advances in First-Order Methods for Continuous and Stochastic Optimization Chair: TIANJIAO LI ←

Asymptotic Normality and Optimality in Nonsmooth Stochastic Optimization – LI-WEI JIANG, CORNELL UNIVERSITY – Damek Davis, Dmitriy Drusvyatskiy

Polyak and Juditsky famously showed that the stochastic gradient method for minimizing smooth and strongly convex functions enjoys a central limit theorem: the error between the running average of the iterates and the minimizer, normalized by the square root of the iteration counter, converges to a normal random vector. Moreover, the covariance matrix of the limiting distribution is in a precise sense “optimal” among any estimation procedure. A long-standing open question is whether similar guarantees – asymptotic normality and optimality – exist for nonsmooth optimization and, more generally, for equilibrium problems.

In this talk, we obtain such guarantees under mild conditions that hold both in concrete circumstances (e.g. nonlinear programming) and under generic linear perturbations. The key insight is that, for various constrained or nonsmooth problems, there exists an active manifold (aka. active set) in algorithmically useful ways. Based on this, we can prove that the shadow sequence of the (stochastic) subgradient method along the active manifold is precisely an inexact Riemannian gradient method with an implicit retraction. As a consequence, we can leverage the “partial smoothness” of nonsmooth problems and extend results for the smooth case.

This talk is based on joint work with Damek Davis and Dmitriy Drusvyatskiy.

A simple uniformly optimal method without line search for convex optimization – TIANJIAO LI, GEORGIA INSTITUTE OF TECHNOLOGY – Guanghui Lan

Line search (or backtracking) procedures have been widely employed into first-order methods for solving convex optimization problems, especially those with unknown problem parameters (e.g., Lipschitz constant). In this paper, we show that line search is superfluous in attaining the optimal rate of convergence for solving a convex optimization problem whose parameters are not given a priori. In particular, we present a novel accelerated gradient descent type algorithm called auto-conditioned fast gradient method (AC-FGM) that can achieve an optimal rate of convergence for smooth convex optimization without requiring the estimate of a global Lipschitz constant or the employment of line search procedures. We then extend AC-FGM to solve convex optimization problems with Hölder continuous gradients and show that it automatically achieves the optimal rates of convergence uniformly for all problem classes with the desired accuracy of the solution as the only input. Finally, we report some encouraging numerical results that demonstrate the advantages of AC-FGM over the previously developed parameter-free methods for convex optimization.

Universal First-Order Methods for Convex Optimization Problems – JIAMING LIANG, UNIVERSITY OF ROCHESTER –

We present a generic algorithmic framework for convex optimization without knowing the levels of convexity and smoothness of the problem. In particular, the framework includes an adaptive subgradient method and an adaptive proximal bundle method. One advantage of the framework is that it does not perform any line search on the strong convexity parameter. To the best of our knowledge, this is the first universal method that does not use a backtracking procedure on the strong convexity parameter. We present complexity results in terms of both the optimality gap and the stationarity. Their complexity bounds are as good as those obtained with known convexity and smoothness parameters.

Optimal and parameter-free gradient minimization methods for convex and nonconvex optimization – JIMMY ZHANG, IE, PURDUE UNIVERSITY –

We propose novel optimal and parameter-free algorithms for computing an approximate solution with small (projected) gradient norm. Specifically, for computing an approximate solution such that the norm of its (projected) gradient does not exceed ϵ , our complexity results match the lower complexity bounds of the convex and strongly convex cases up to some universal constant, and achieve the best known complexity bound for the nonconvex case for the first time in the literature. Moreover, for all the convex, strongly convex, and nonconvex cases, we propose parameter-free algorithms that do not require the input of any problem parameters. To the best of our knowledge, there do not exist such parameter-free methods before especially for the strongly convex and nonconvex cases. Since most regularity conditions (e.g., strong convexity and lower curvature) are imposed over a global scope, the corresponding problem parameters are notoriously difficult to estimate. However, gradient norm minimization equips us with a convenient tool to monitor the progress of algorithms and thus the ability to estimate such parameters in-situ.

3a - TB111 - 515A - PDE-constrained optimization under uncertainty Chair: DANIEL WALTER ←

Optimality Conditions with Probabilistic State Constraints – CAROLINE GEIERSBACH, WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS – René Henrion, Pedro Pérez Aros

In this talk, we discuss optimization problems subject to random state constraints, where we distinguish between the chance-constrained case and the almost sure formulation. We highlight some of the difficulties in the infinite-dimensional setting, which is of interest in physics-based models where a control belonging to a Banach space acts on a system described by a partial differential equation (PDE) with random inputs or parameters. We study the setting in which the obtained state should be bounded uniformly over the physical domain with high probability, or even probability one. We apply our results to a model with a random elliptic PDE, where the randomness is induced by the right-hand side. For the chance-constrained setting, this structure allows us to obtain an explicit representation for the Clarke subdifferential of the probability function

using the spherical radial decomposition of Gaussian random vectors. This representation is used for the numerical solution in a discretize-then-optimize approach. For the almost sure setting, we use a Moreau-Yosida regularization and solve a sequence of regularized problems in an optimize-then-discretize approach. The solutions are compared, providing insights for the development of further algorithms.

Convergence rates for ensemble-based solutions to optimal control of uncertain dynamical systems – JOHANNES MILZ, GEORGIA INSTITUTE OF TECHNOLOGY – Olena Melnikov

We consider optimal control problems with nonlinear ordinary differential equations with uncertain inputs. We approximate the optimal control problem using the sample average approximation, yielding optimal control problems with ensembles of deterministic dynamical systems. Utilizing techniques developed for deriving metric entropy bounds for suprema of sub-Gaussian processes, we derive nonasymptotic Monte Carlo-type convergence rates for the ensemble-based solutions. Moreover, we explore the challenges associated with extending our approach to establishing convergence rates for optimal control under uncertainty with time-dependent partial differential equations. We illustrate our theoretical framework on a catalyst mixing problem under uncertainty.

Adaptive Surrogate Modeling for Trajectory Optimization with Model Inexactness – MATTHIAS HEINKSCHLOSS MATTHIAS, RICE UNIVERSITY –

In many applications one must compute optimal trajectories from imperfect knowledge of the dynamics. For example, solving trajectory optimization problems for hypersonic vehicles requires the computation of lift and drag coefficients at many flight configurations. Determining these coefficients requires expensive high-fidelity computations using detailed representations of the hypersonic vehicle. This talk proposes the use of computationally inexpensive adaptive Gaussian process models constructed from high-fidelity samples to approximate the components of the dynamics that are expensive to evaluate. To reduce the effect of model errors on the optimal trajectory, the current Gaussian process model is updated as needed at the cost of evaluating the components of the dynamics at a small number of additional sample points. First, the optimal control problem is solved using the mean of the current Gaussian process model to represent the dynamics. Next, sensitivity analysis is combined with properties of the Gaussian process model, such as its variance, to determine whether the Gaussian process model needs to be updated, and if so, at which samples the dynamics should be evaluated to update the Gaussian process model. This talk outlines our current model refinement procedure and demonstrates its performance on a trajectory optimization problem for a hypersonic vehicle with lift and drag models that are known, but expensive to evaluate.

Empirical estimators for risk-neutral composite optimal control with applications to bang-bang control – DANIEL WALTER, HUMBOLDT UNIVERSITÄT ZU BERLIN –

We consider risk-neutral composite optimal control problems where the objective functional consists of an expectation functional involving the solution operator of a parametrized PDE and a nonsmooth but convex regularizer. Our particular inter-

est lies in cases in which the latter lacks strong convexity, e.g. “bang-bang-off” regularizers given by the sum of a sparsifying term and additional box constraints. For the practical realization, the expectation term is treated by a Monte Carlo sample-based approach. We study the asymptotic consistency of this ansatz and derive nonasymptotic sample size estimates. Our analyses leverage problem structures commonly encountered in PDE-constrained optimization problems, including compact embeddings. Our theoretical results are confirmed by extensive numerical results for both, linear and bilinear, control problems.

3a - WA102 - 515A - Multistage Optimization Programs under Uncertainty Chair: FRANCESCA MAGGIONI DARINKA DENTCHEVA

Multistage distributional robustness, time consistency and optimal policy – RUI GAO, UNIVERSITY OF TEXAS AT AUSTIN –

In this talk, I will discuss distributional robustness in a dynamic setting, where the deviation is measured by the causal transport distance between stochastic processes. Such a choice accounts for information evolution, making it hedge against a plausible family of data processes. I will present a recursive reformulation to evaluate the worst-case risk of any given random sequence and explore the intricacies of time consistency in dynamic risk measures. Furthermore, I will present dynamic programming reformulations for finding the optimal robust policy.

Distributionally robust chance-constrained Markov decision processes – ABDEL LISSER, UNIVERSITY PARIS SACLAY, CENTRALESUPELEC – Hoang Nam Nguyen, Vikas Vikram Singh

Markov decision process (MDP) is a decision making framework where a decision maker is interested in maximizing the expected discounted value of a stream of rewards received at future stages at various states which are visited according to a controlled Markov chain. Many algorithms including linear programming methods are available in the literature to compute an optimal policy when the rewards and transition probabilities are deterministic. In many practical situations, it is often the case that only a partial information about the underlying distribution is available based on the historical data. In that case, a distributionally robust approach, is used to model the uncertainties, which assumes that the true distribution belongs to an uncertainty set based on its partially available information. In this paper, we consider an MDP problem where the reward vector is known and the transition probability vector is a random vector which follows a discrete distribution whose information is not completely known. We formulate the MDP problem using distributionally robust chance-constrained optimization framework under various types of moments based uncertainty sets, and statistical-distance based uncertainty sets defined using ϕ -divergence and Wasserstein distance metric. For each uncertainty set, we propose an equivalent mix-integer bilinear programming problem or a mix-integer semidefinite programming problem with bilinear constraints. As an application, we study a machine replacement problem and perform numerical experiments on randomly generated instances.

Risk Evaluation and Control for Distributed Multi-Agent Systems – DARINKA DENTCHEVA,

STEVENS INSTITUTE OF TECHNOLOGY – Aray Almen

We discuss risk evaluation and risk-averse optimization of complex distributed systems with general risk functionals. We postulate a set of axioms for the functionals evaluating the total risk of the system and show the dual representation for the systemic risk measures. Furthermore, we propose two new families of measures constructed by using either collections of linear scalarizations or non-linear risk aggregation. The new framework facilitates risk-averse sequential decision making by distributed methods. The proposed approach is compared theoretically and numerically to some of the systemic risk measurements in the existing literature. We also formulated a multi-stage decision problem arising in communication networks when a team of robots explores an area and each robot reports relevant information. The goal is to determine a few reporting points so that the communication is conducted most efficiently while managing the risk of losing information. A distributed numerical method is developed for solving the problem. In this context, we compare the new risk measures to other methods of systemic risk evaluation. We show that the proposed framework is less conservative and results in a substantially better solution of the problem at hand as compared to a linear aggregation of the risk of individual agents as well as other methods.

3a - WB232 - 515A - **Decision-Focused Learning and Bilevel Models for Stochastic Optimization** Chair: TITO HOMEM DE MELLO

Assessing solution quality in risk-averse stochastic programs – RUBEN VAN BEESTEN, ERASMUS UNIVERSITY ROTTERDAM – Nick Koning, David Morton

In stochastic programming algorithms, assessing whether a candidate solution is near-optimal is of vital importance. For risk-neutral stochastic programs, this is typically done by estimating the optimality gap using a Monte Carlo estimator. For risk-averse stochastic programs, however, these estimators can underestimate the true optimality gap due to a bias in the Monte Carlo estimator of a risk measure. In this presentation, we propose a Monte Carlo method that yields valid estimates for the optimality gap in risk-averse stochastic programs with a coherent risk measure. Our method avoids the aforementioned bias by using two independent samples, each estimating a different element constituting the optimality gap. We address both tractability of our approach and the quality of the resulting optimality gap estimates.

An implementation of the Adaptive Partition Method to Stochastic Bilevel Problems – FRANCISCO JARA MORONI, UNIVERSIDAD DIEGO PORTALES – Carlos Muñoz

The Adaptive Partition Method (APM) has proven to be an efficient way to solve two-stage stochastic problems. The underlying idea is that many scenarios are similar in some sense. The method aims to find a partition of the sample space and solve an aggregated master problem that leads to the same optimal solution as the original problem. This work focuses on extending the APM to stochastic bilevel problems. For the first implementation, we investigate the more convenient structures

of such problems and then discuss how to deal with problems where this structure is absent. We complement our work with numerical experiments, which show that APM is a valid approach for these problems.

Decision-focused predictions via pessimistic bilevel optimization: a computational study

– VÍCTOR BUCAREY, UNIVERSIDAD DE O'HIGGINS – Calderón Sophia, Gonzalo Muñoz, Frederic Semet

Dealing with uncertainty in optimization parameters is an important and longstanding challenge. Typically, uncertain parameters are predicted accurately, and then a deterministic optimization problem is solved. However, the decisions produced by this so-called predict-then-optimize procedure can be highly sensitive to uncertain parameters. In this work, we contribute to recent efforts in producing decision-focused predictions, i.e., to build predictive models that are constructed with the goal of minimizing a regret measure on the decisions taken with them.

We formulate the exact expected regret minimization as a pessimistic bilevel optimization model. Then, using duality arguments, we reformulate it as a non-convex quadratic optimization problem. Finally, we show various computational techniques to achieve tractability. We report extensive computational results on shortest-path instances with uncertain cost vectors. Our results indicate that our approach can improve training performance over the approach of Elmachtoub and Grigas (2022), a state-of-the-art method for decision-focused learning.

Application-Driven Optimal Pointwise Forecasts for Stochastic Optimization – TITO HOMEM DE MELLO, UNIVERSIDAD ADOLFO IBÁÑEZ

– Juan Valencia, Felipe Lagos, Guido Lagos

We consider the class of two-stage stochastic programs with uncertainty only on the right hand side. Such a class encompasses practical many problems, especially in inventory models. We show that, under certain conditions, there exist an optimal scenario, in the sense that solving the problem with that scenario yields the same optimal solution as the original problem. In a data-driven context, the result means that pointwise forecasts of the uncertain variables can be optimal, and we show that such an optimal forecast can in principle be obtained by a decision-focused learning approach.

3a - WC258 - 515A - **Theory and Algorithms for Chance Constrained Optimization** Chair:

ANIRUDH SUBRAMANYAM

Joint chance-constrained optimization with continuous distributions – DANIELA BERNHARD,

FRIEDRICH ALEXANDER UNIVERSITÄT ERLANGEN NÜRNBERG – Frauke Liers, Michael Stingl

Input parameters of optimization problems are often affected by uncertainties. Chance constraints are a common way to model stochastic uncertainties in the constraints. We consider joint chance-constrained optimization problems with continuous distributions and non-convex constraint functions. For this difficult class of problems, no efficient general solution approaches exist. We propose a gradient-based approach to approximately solve the joint chance-constrained problems. Therefore, we approximate the original problem by smoothing indicator functions and penalize a violation of the smoothed

constraint in the objective function. The approximation problem is solved with the Continuous Stochastic Gradient method, a variant of the stochastic gradient descent with improved convergence properties that has recently been introduced in the literature. Furthermore, we provide a convergence theory for the smoothing and penalty approximations with sequences of increasing smoothing and penalty parameters. Under very mild assumptions, our approach is applicable to a wide range of joint chance-constrained optimization problems. We illustrate the efficiency of our approach for difficult practical problems with an application to gas networks. The numerical experiments demonstrate that the approach quickly finds nearly feasible solutions for difficult joint chance-constrained problems with non-convex constraint functions and continuous distributions, even for realistically-sized instances.

Self-Structured Importance Sampling for Chance-Constrained Optimization – SAI RAKSHITH, PSG COLLEGE OF TECHNOLOGY – Anand Deo, Shanyin Tong, Karthyek Murthy, Anirudh Subramanyam

Chance constraints naturally emerge in systems where reducing risk is crucial, particularly those exposed to high-impact but rare events such as natural disasters, infrastructure failures, and financial crashes. However, the rarity of such events renders the solution of chance-constrained models impractical using naive Monte Carlo methods. Importance sampling can be a viable solution method in such cases, as it employs an alternate probability distribution to generate samples more frequently from rare-event regions. Nonetheless, the challenge of selecting an appropriate alternate distribution that is effective for all decision variables across various problems remains unresolved. Most existing strategies are problem-specific, necessitating an in-depth understanding of the system’s dynamics. Our work adapts the self-structured importance sampling (SSIS) method, recently proposed by Deo & Murthy, to develop a new algorithm for solving chance-constrained problems. This method transforms samples generated through naive Monte Carlo to samples that more accurately represent rare region behavior and can be universally applied across a wide range of constraint functions and probability distributions. Our results demonstrate that the proposed algorithm can substantially outperform Monte Carlo sampling in terms of both feasibility and optimality, even with small sample sizes.

Chance-Constrained Optimization for Uncertain Contact-Rich Systems – YUKI SHIRAI, MIT-SUBISHI ELECTRIC RESEARCH LABORATORIES – Devesh Jha, Arvind Raghunathan

Designing controllers for contact-rich systems is known to be challenging due to the non-smooth dynamics. In general, the parameters involved in contact-rich systems such as friction coefficients are very difficult to estimate perfectly. This can lead to poor performance during execution. However, considering uncertainties for contact-rich systems is even more challenging and very little understood in open literature. In this talk, we talk about novel insight regarding stochastic contact-rich systems and present a novel contact-aware covariance controller for the same. The proposed formulation highlights a very fundamental treatment of stochastic contact-rich systems using sample average approximation. Furthermore, a novel non-linear programming method using bilevel optimization is proposed, resulting in much better computation compared to

the mixed-integer optimization-based stochastic optimization, which has been used in the covariance steering community. This paper makes an important contribution toward safe manipulation and locomotion in the real world using stochastic contact-rich models.

3a - ThA101 - 515A - **Decision-making under Preference Uncertainty** Chair: WENJIE HUANG



Risk Budgeting Allocation for Dynamic Risk Measures – SILVANA PESENTI, UNIVERSITY OF TORONTO –

We define and develop an approach for risk budgeting allocation - a risk diversification portfolio strategy - where risk is measured using a dynamic time-consistent risk measure. Specifically, a dynamic risk budgeting strategy is a portfolio allocation where at every trading point each asset contributes a predefined percentage to the future risk of the portfolio. For this, we introduce a notion of dynamic risk contributions that generalise the classical Euler contributions and which allow us to obtain dynamic risk contributions in a recursive manner. We prove that, for the class of dynamic coherent distortion risk measures, the risk allocation problem may be recast as a sequence of strictly convex optimisation problems. Moreover, we show that any self-financing dynamic risk budgeting strategy with initial wealth of 1 is a scaled version of the unique solution of the sequence of convex optimisation problems. Furthermore, we develop an actor-critic approach, leveraging the notion of conditional elicibility of dynamic risk measures, to solve for risk budgeting strategies using deep learning. We illustrate the methodology on a simulated market model, with a dynamic risk measures constructed through conditional Expected Shortfalls, discuss that the dynamic risk budgeting strategy is non-Markovian and how the agent invests.

Nonparametric Inverse Optimization: Theory and Applications – REZA VALIMORADI, UNIVERSITY OF OTTAWA – Jonathan Li

Inverse optimization is a powerful tool for understanding decision-making processes by inferring the objective functions behind observed decisions. Traditional approaches rely heavily on parametric models, which assume the objective function’s structure can be precisely known. This assumption can be impractical and prone to model mis-specification in complex real-world scenarios. Addressing this challenge, we introduce a novel nonparametric framework for inverse optimization. Our approach stands out by its flexibility to model a large class of objective functions without predefined parametric structures and that it can be implemented in data-efficient fashion. We prove the computational tractability of the nonparametric models and demonstrate their advantages via numerical experiments.

Preference Ambiguity and Robustness in Multistage Decision Making – HUIFU XU, THE CHINESE UNIVERSITY OF HONG KONG –

In this work, we consider a multistage expected utility maximization problem where the decision maker’s utility function at each stage depends on historical data and the information on the true utility function is incomplete. To mitigate adverse impact arising from ambiguity of the true utility, we propose a maximin robust model where the optimal policy is based on the

worst-case sequence of utility functions from an ambiguity set constructed with partially available information about the decision maker's preferences. We then show that the multistage maximin problem is time consistent when the utility functions are historical-path-dependent and demonstrate with a counter example that the time consistency may not be retained when the utility functions are historical-path-independent. With the time consistency, we show the maximin problem can be solved by a recursive formula whereby a one-stage maximin problem is solved at each stage beginning from the last stage. Moreover, we propose two approaches to construct the ambiguity set: a pairwise comparison approach and a zeta-ball approach where a ball of utility functions centered at a nominal utility function under zeta-metric is considered. To overcome the difficulty arising from solving the infinite dimensional optimization problem in computation of the worst-case expected utility value, we propose piecewise linear approximation of the utility functions and derive error bound for the approximation under moderate conditions. Finally, we use the stochastic dual dynamic programming (SDDP) method and the nested Benders' decomposition method to solve the multistage historical-path-dependent preference robust problem and the scenario tree method to solve the historical-path-independent problem, and carry out comparative analysis on the efficiency of the computational schemes as well as out-of-sample performances of the historical-path-dependent and historical-path-independent models. The preliminary results show that the historical-path-dependent preference robust model solved by SDDP algorithm displays overall superiority. This is a joint work with Jia Liu and Zhiping Chen in Xi'an Jiaotong University, China.

Multi-attribute Preference Robust Optimization with Quasi-Concave Choice Functions – WENJIE HUANG, THE UNIVERSITY OF HONG KONG – Jian Wu, William B. Haskell, Huifu Xu

In behavioural economics, a decision maker's preferences are expressed by choice functions. Preference robust optimization (PRO) is concerned with problems where the true choice function which represents the decision maker's preferences is ambiguous, and the optimal decision is based on the worst-case choice function from a set of plausible choice functions constructed with elicited preference information. In this paper, we propose a PRO model to support choice functions that are: (i) monotonic (prefer more to less), (ii) quasi-concave (prefer diversification), and (iii) multi-attribute (have multiple objectives/criteria). As a main result, we show that the robust choice function can be constructed efficiently solved by solving a sequence of linear programming problems. In decision making with worst-case choice function, we demonstrate how the maximin problem can be efficiently solved by a sequence of convex optimization problems. To examine the behavior and scalability of the proposed model and computational schemes, we apply them to a portfolio optimization problem and a capital allocation problem and report the numerical results.

3a - ThB935 - 515A - **Continuous Stochastic Programming 2** Chair: DIONYSIS KALOGERIAS ↗

Duality in Optimal Stopping – NICOLAS GAGNÉ, UNIVERSITÉ DE MONTRÉAL – Fabian Bastin Optimal stopping problems, commonly encountered in financial mathematics, involve determining the most opportune time

to execute an action (e.g., exercising an option) to maximize expected returns. We introduce the dual problem associated to an optimal stopping problem by way of a perturbation function. We give an intuitive interpretation of the dual problem as well as the role it plays in practice.

Adaptive Partitioning for Chance-Constrained Problems with Finite Support – MARIUS ROLAND, POLYTECHNIQUE MONTRÉAL – Alexandre Forel, Thibaut Vidal

This presentation considers chance-constrained stochastic problems (CCSPs) with finite support. We present an iterative algorithm that solves a reduced size chance-constrained model. This reduced model is obtained by partitioning the scenario set and, when solved to optimality, yields a bound on the optimal objective value of the original CCSP. Moreover, the algorithm ensures finite termination at an optimal solution of the original CCSP. At the heart of the algorithm lie two fundamental operations: refinement and merging. Refinement involves splitting a subset of the partition, while merging combines two subsets. These operations are executed to enhance the bound obtained in each step of the algorithm while minimally increasing the size of the reduced model to be solved. Under mild conditions, we prove that, for specific refinement and merge operations, the bound obtained after solving each reduced model strictly improves throughout the iterative process. Furthermore, the algorithm structure allows for the seamless integration of various computational enhancements, significantly reducing the computational time required to solve the reduced chance-constrained problems. Based on theoretical arguments we also provide strategies to initialize the partition considered in the algorithm. The efficiency of the algorithm is assessed through numerical experiments. We study the impact of every component of the algorithm and compare its performance against state-of-the-art methods on chance constrained multidimensional knapsack problems.

Consistent Assortment Optimization under the MNL with Exogenous Data – CARLOS CARDONHA, UNIVERSITY OF CONNECTICUT, SCHOOL OF BUSINESS – Aritanan Gruber

This work investigates the Assortment Optimization Problem under the Multinomial Logit choice model in scenarios where the firm does not know the attractiveness factors associated with each product. In our model, the firm relies on exogenous data consisting of the assortments offered by other firms and random variables defining a set of conversion rates, which are multiplicative factors indicating how the products' attractiveness adopted in the external assortments differ from the respective values for the firm. The notion of optimality is nuanced in this model, so we focus on identifying consistent assortments in the sense that the assortment decisions made by the firm should not conflict with exogenous data. We investigate the structural properties and different solution techniques for the problem based on assumptions about the conversion rates. We show that the expectation-based deterministic formulation of the problem can be solved efficiently. We also study some probabilistic-constrained formulations where such problems can be recast as second-order cone programs assuming the conversion rates are normally distributed. Our results show that the model does not require unrealistically large amounts of exogenous data to deliver accurate results, thus providing evidence that our framework can be adopted in practice.

Strong Duality in Risk-Constrained Non-convex Functional Programming – DIONYSIS KALOGERIAS, YALE UNIVERSITY – Pougkakiotis Spyridon

We show that risk-constrained functional optimization problems with general integrable nonconvex instantaneous reward/constraint functions exhibit strong duality, regardless of nonconvexity. We consider risk constraints featuring convex and positively homogeneous risk measures admitting dual representations with bounded risk envelopes, generalizing expectations. Popular risk measures supported within our setting include the conditional value-at-risk (CVaR), the mean-absolute deviation (MAD, including the non-monotone case), certain distributionally robust representations and more generally all real-valued coherent risk measures on the space L^1 . We highlight the usefulness of our results by further discussing various generalizations of our base model, extensions for risk measures supported on L^p for $p \geq 1$, implications in the context of mean-risk tradeoff models, as well as more specific applications in wireless systems resource allocation, and supervised constrained learning. Our core proof technique appears to be new and relies on risk conjugate duality in tandem with J. J. Uhl's weak extension of A. A. Lyapunov's convexity theorem for vector measures taking values in general infinite-dimensional Banach spaces.

For more information, please have a look at our full and detailed preprint available on arXiv through the link: <https://arxiv.org/pdf/2206.11948.pdf>

3a - ThC252 - 515A - Advances in Risk Optimization and Design Chair: QUANYAN ZHU ←

Federated Learning for Discrete Optimal Transport with Large Population under Incomplete Information – YINGDONG LU, IBM – Juntao Chen

Optimal transport (OT) formulation and algorithms allow for optimal allocation of limited resources in a network consisting of sources and targets. The standard OT paradigm does not extend well over a large population of different types. In this paper, we formulate innovate OT problem and design algorithms to cope with resource allocation with a large and heterogeneous population of target nodes. The heterogeneity of targets is described by a type distribution function. We consider two instances in which the distribution is known and unknown to the sources, i.e., transport designer. For the former case, we propose a fully distributed algorithm to obtain the solution. For the latter case in which the targets' type distribution is not available to the sources, we develop a collaborative learning algorithm to compute the OT scheme efficiently. We evaluate the performance of the proposed learning algorithm using a case study.

Data driven decision making under uncertainty with entropic risk measure – UTSAV SADANA, UNIVERSITY OF MONTREAL – Erick Delage, Angelos Georgioud

Entropic risk measure is widely used to account for tail risks associated with an uncertain parameter. We study distributionally robust optimization problems with entropic risk measures

and use cross-validation to tune the radius (hyperparameter) of the ambiguity set. However, with limited data, the empirical average of the entropic risk associated with scenarios in the validation set underestimates the true entropic risk. We provide two procedures that first learn Gaussian mixture models and then use bootstrapping to identify scaling parameters to correct the bias in the entropic risk estimated using the validation data. We show that one can achieve lower risk using our debiasing procedure in project selection and portfolio optimization problems.

Stackelberg risk preference design – SHUTIAN LIU, NEW YORK UNIVERSITY – Quanyan Zhu

Decision-makers are often assumed to possess fixed risk preferences. However, evidences have shown that individual risk preferences are unstable and are subject to manipulation. This work proposes a Stackelberg risk preference design (STRIPE) problem to capture a designer's incentive to influence DMs' risk preferences. STRIPE consists of two levels. In the lower level, individual DMs in a population, known as the followers, respond to uncertainties according to their risk preference types. In the upper level, the leader influences the distribution of the types to induce targeted decisions and steers the follower's preferences to it. Our analysis centers around the solution concept of approximate Stackelberg equilibrium that yields suboptimal behaviors of the players. We show the existence of the equilibrium, derive a single-level reformulation of the design problem, and suggest two applications for the proposed framework.

3a - FA247 - 515A - Latest Developments in Stochastic Optimization Chair: JONATHAN LI ←

Reinforcement Learning for Portfolio Optimization in Model with Regime-Switching – DAVID SAUNDERS, UNIVERSITY OF WATERLOO –

We study a reinforcement learning approach to portfolio selection when the underlying asset price process follows a regime-switching diffusion model. An analytical solution to the HJB equation for the exploratory mean-variance model under regime switching is derived and implementation in practice employing a deep neural network will be discussed. Empirical results investigating robustness with respect to model mis-specification will also be presented.

Optimal trade execution with learning – ANDREW LIM, NATIONAL UNIVERSITY OF SINGAPORE – Galvin Ng

A trader would like to maximize expected revenue from selling an inventory of stocks by a specified terminal time when expected returns are linear in the trading rate but with unknown intercept (drift) and slope (price impact). Since the trading rate affects the data, this is intrinsically a problem of exploration and exploitation. To sidestep the challenges of solving the high-dimensional dynamic program, we consider a high-noise asymptotic regime and use Taylor Series to extract the exploration component of the optimal policy. Exploration depends, in a complex but intuitive way, on the time horizon, inventory level, and parameter uncertainty.

On Generalization and Regularization via Wasserstein Distributionally Robust Optimization – QINYU WU, UNIVERSITY OF WATERLOO – Jonathan Li, Tiantian Mao

Wasserstein distributionally robust optimization (DRO) has found success in operations research and machine learning applications as a powerful means to obtain solutions with favorable out-of-sample performances. Two compelling explanations for the success are the generalization bounds derived from Wasserstein DRO and the equivalency between Wasserstein DRO and the regularization scheme commonly applied in machine learning. Existing results on generalization bounds and the equivalency to regularization are largely limited to the setting where the Wasserstein ball is of a certain type and the decision criterion takes certain forms of an expected function. In this paper, we show that by focusing on Wasserstein DRO problems with affine decision rules, it is possible to obtain generalization bounds and the equivalency to regularization in a significantly broader setting where the Wasserstein ball can be of a general type and the decision criterion can be a general measure of risk, i.e., nonlinear in distributions. This not only accommodates important applications in classification, regression, and risk minimization, but also expands to general decision-theoretical frameworks previously unaddressed using Wasserstein DRO. Our results are strong in that the generalization bounds do not suffer from the curse of dimensionality and the equivalency to regularization is exact. As a byproduct, our regularization results broaden considerably the class of Wasserstein DRO models that can be solved efficiently via regularization formulations. These general assurances should serve as a strong basis for expanding the application of Wasserstein DRO across diverse domains of data-driven decision problems.

Model Aggregation for Risk Evaluation and Robust Optimization – Tiantian Mao, University of Science and Technology of China

We introduce a new approach for prudent risk evaluation based on stochastic dominance, which will be called the model aggregation (MA) approach. In contrast to the classic worst-case risk (WR) approach, the MA approach produces not only a robust value of risk evaluation but also a robust distributional model, independent of any specific risk measure. The MA risk evaluation can be computed through explicit formulas in the lattice theory of stochastic dominance, and under some standard assumptions, the MA robust optimization admits a convex-program reformulation. The MA approach for Wasserstein and mean-variance uncertainty sets admits explicit formulas for the obtained robust models. Via an equivalence property between the MA and the WR approaches, new axiomatic characterizations are obtained for the Value-at-Risk (VaR) and the Expected Shortfall (ES, also known as CVaR). The new approach is illustrated with various risk measures and examples from portfolio optimization.

3a - FB936 - 515A - Continuous Stochastic Programming 3 Chair: SHIBSHANKAR DEY 

Stable Matching with Contingent Priorities – Federico Bobbio Bobbio, University of Montreal – Ignacio Rios, Alfredo Torrico, Margarida Carvalho

We study stable matching problems under contingent priorities, whereby the clearinghouse prioritizes some agents based on the allocation of others. Using school choice as a motivating

example, we first introduce a stylized model of a many-to-one matching market where the clearinghouse aims to prioritize applicants with siblings assigned to the same school and match them together. We provide a series of guidelines to implement these contingent priorities and introduce two novel concepts of stability that account for them. We study some properties of the corresponding mechanisms, including the existence of a stable assignment under contingent priorities, its incentive properties, and the complexity of finding one if it exists. Moreover, we provide mathematical programming formulations to find such stable assignments whenever they exist. Finally, using data from the Chilean school choice system, we show that our framework can significantly increase the number of siblings assigned together while having no large effect on students without siblings.

On the Chance-Constrained Optimization with Gaussian Mixture Models – Shibshankar De, Industrial Engineering and Management Sciences, Northwestern University – Sanjay Mehrotra, Anirudh Subramanyam, Tianyang Yi, D. Adrian Maldonado

This study develops mixed-integer quadratically constrained program (MIQCP) and second-order cone program (SOCP) formulations for chance-constrained optimization problems with Gaussian mixture models. By leveraging the convex-concave characteristics of the standard normal cumulative distribution function (CDF), we provide both piecewise-linear inner and outer approximations of the original problem. Optimal objectives from these approximations are shown to lie within an epsilon difference of the true value with an adequate number of linear pieces. Additionally, for the conservative SOCP formulation, an explicit upper bound of the objective value is provided. To expedite computation, we propose a heuristic that selects a significantly lower number of linear pieces than the worst-case complexity, without sacrificing desired approximation accuracy. Through extensive numerical experiments analyzing the trade-off between solution time and approximation quality, we find that our formulations offer a significant computational advantage over classical Sample Average Approximation (SAA) methods across various scenarios, including different problem classes and sizes, mixture components, chance constraint reliability levels, and reference distributions.

3b - MB287 - 512H - Recent modeling language extensions: Stochastics and beyond
Chair: MICHAEL FERRIS 

GAMSPy: Algebraic Modeling and Python – Steven Dirkse, GAMS Development Corporation – Muhammet Soytürk, Adam Christensen
Optimization applications combine technology and expertise from many different areas, including model-building, algorithms, and data-handling. Often, the gathering, pre/post-processing, and visualization of the data is done by a diverse organization-spanning group that shares a common bond: their skill in and appreciation for Python and the vast array of available packages it provides. For this reason, GAMS offers a new comfortable way to integrate with Python on the data-handling and modeling side. In this talk, we will explore the benefits of our Python library GAMSPy.

DAG-based modeling framework for extended mathematical programming – OLIVIER HUBER, UNIVERSITY OF WISCONSIN MADISON –

We are investigating the modeling and design of algorithms for models with multiple optimization problems. This structure is present in many problem classes like bilevel/MPEC/multilevel programming or Nash equilibrium problems or whenever value functions present in the mappings of an optimization problem, like in multistage stochastic programming with coherent risk measures. Here, we tackle the challenge of modeling setups with any combination of the above structures. For instance, decentralized energy markets with part of the problem data being subjected to uncertainty can feature a Nash equilibrium problem where market participants have a nonsmooth objective function due to the presence of a coherent risk measure.

For such complex but structured setups, we propose a modeling framework based on a directed acyclic graph (DAG) structure. The nodes are of two types: the first one subsumes parametric optimization problems and variational inequalities, while the second one indicates a Nash behavior between its children nodes. Edges between two optimization problems specifies their interaction: either a hierarchical one (bilevel/MPEC) or the value function of the child appears in the parent problem.

This DAG structure is leveraged by model transformations to transform part or all of the problem in a form amenable to computation by existing solvers. When tackling some discrete stochastic programs, the reformulated problems do then belong to a classical model type and high-performance solvers for discrete optimization problem can be used to compute a solution. The DAG is also useful in the design of decomposition algorithms for solving complex instances. An implementation of these ideas is present in ReSHOP, a reformulation solver for hierarchical optimization problems.

Advances in Automated Conversion of Optimization Problems – ROBERT FOURER, AMPL OPTIMIZATION INC. – Gleb Belov

We take it for granted that an optimization package accepts both minimization and maximization problems, recognizes them as equivalent, and converts all minimizations to maximizations (or vice-versa) before solving. This is only the very simplest example of the many conversions carried out routinely by large-scale optimization software. The range of expressions recognized by modeling languages and solvers has been steadily extended in ways that make optimization models easier to describe, validate, and maintain — but that make conversion possibilities ever more numerous and complex.

Continuing this trend, automatic conversions have been a central feature in the design of a new solver interface framework for the AMPL modeling language. This presentation describes a range of challenges that have been faced in detecting formulations that solvers can handle, and in implementing conversions to forms that solvers require. Examples combining a variety of discrete and nonlinear expressions lead to some general recommendations for design and implementation of automated conversions.

Stochastic Equilibrium Problems – MICHAEL FERRIS, UNIVERSITY OF WISCONSIN MADISON – Olivier Huber

We present a mechanism for describing and solving collections

of optimization problems that are linked by equilibrium conditions. The general framework of MOPEC (multiple optimization problems with equilibrium constraints) captures many example applications that involve independent decisions coupled by shared resources. Included in this class are classical models such as complementarity, general equilibria, and agent based formulations arising from Nash Games.

Stochastic equilibria can be used to model many data driven applications, including many dynamic models of competition. In this model, players solve risk averse optimization problems that are coupled by a shared scenario tree. Players are linked by equilibrium conditions at nodes of that tree. We outline several algorithms and investigate their computational efficiency. Applications to energy planning and their interactions with environmental concerns will be outlined, and we will demonstrate how these problems can be implemented in modeling systems.

3b - MC57 - 512A - Modeling and Optimization under Uncertainty Chair: SUVRAJEET SEN [←](#)

Distributionally robust two-stage convex quadratic programming – DAVID MORTON, NORTHWESTERN UNIVERSITY – Nazlican Arslan

Distributionally robust optimization provides a framework for stochastic optimization when the distribution of the random vector is uncertain. The uncertain distribution is assumed to come from an ambiguity set of plausible distributions. The goal is to find a decision that hedges against the worst distribution in the ambiguity set. We focus on a two-stage stochastic program with convex quadratic recourse under distributional ambiguity. We assume uncertainty only in the right-hand side vector. We consider data-driven distributional ambiguity sets based on the Wasserstein distance and an optimal quadratic transport distance. For the latter, we derive a decomposition algorithm and present computational results.

Multi-scale and Multi-stage Stochastic Discrete Optimization and its Applications in Energy and Sustainability – SUVRAJEET SEN, UNIVERSITY OF SOUTHERN CALIFORNIA – Yihang Zhang

One of the major challenges in sustainability arises from the need to coordinate processes of different scales in such a manner that decisions in each scale can be coordinated in a computationally realizable manner, while ensuring that certain metrics of the system (e.g., costs, system reliability, emissions, renewable penetration, etc.) can be maintained within acceptable bounds. The challenges here arise due to several features, such as multi-scale dynamical uncertainty, mixed-binary dynamic decisions, and the interplay among multi-scales. Such collections of intermingling mathematical programming models are central to many applications in energy and sustainability. We will present a decomposition-coordination algorithm which will coordinate multi-scale decisions and uncertainty in such a manner that the system will seek to operate within acceptable bounds as more data becomes available.

Integer Programming for Learning Directed Acyclic Graphs from Non-identifiable Gaussian Models – SIMGE KUCUKYAVUZ, NORTHWESTERN UNIVERSITY – Tong Xu, Armeen Taeb, Ali Shojaie

We study the problem of learning directed acyclic graphs from continuous observational data, generated according to a linear

Gaussian structural equation model. State-of-the-art structure learning methods for this setting have at least one of the following shortcomings: i) they cannot provide optimality guarantees and can suffer from learning sub-optimal models; ii) they rely on the stringent assumption that the noise is homoscedastic, and hence the underlying model is fully identifiable. We overcome these shortcomings and develop a computationally efficient mixed-integer programming framework for learning medium-sized problems that accounts for arbitrary heteroscedastic noise. We present an early stopping criterion under which we can terminate the branch-and-bound procedure to achieve an asymptotically optimal solution and establish the consistency of this approximate solution. In addition, we show via numerical experiments that our method outperforms three state-of-the-art algorithms and is robust to noise heteroscedasticity, whereas the performance of the competing methods deteriorates under strong violations of the identifiability assumption.

3b - TB187 - 512A - Modeling and Algorithms for Stochastic Discrete Optimization Chair:
HAOXIANG YANG

Optimization via Simulation with Discontinuous Recourse Function – ZHUO ZHANG, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN

Staged alert systems have been successfully implemented to minimize the socioeconomic loss while avoiding overwhelming the healthcare system. Optimizing the staged alert system parameters can be considered a challenging two-stage stochastic mixed-integer programming with a discontinuous recourse function. Traditional simulation-based optimization techniques assume continuity and smoothness within objective functions, while other metaheuristics often lack any optimality guarantee. We propose to develop a novel approach that combines Gaussian process-based optimization via simulation techniques with a methodology specifically designed to handle the discontinuity. We prove local optimality results with a convergence guarantee and demonstrate the efficacy of our proposed method by optimizing the stage thresholds in a pandemic staged alert system.

A Distributed Optimization Method for Large-Scale Production Planning Problem – XINGYU LIN, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN

We consider an important type of production planning problems where firms need to decide the production and inventory level for each item, and the transportation between plants during a planning horizon. The objective is to minimize the total costs including the transportation, inventory holding cost, and a penalty cost for unmet demand, under constraints of limited resource and flow balance. In practice, the scale of such problems can be huge for large enterprises with a complex network structure and millions of decision variables and constraints, potentially including integer variables, which is hard for commercial solvers such as Gurobi and CPLEX to directly solve. In this work, we present a solution method for a mixed-integer large-scale production planning problem. The method involves decomposition in terms of items, where the item-based planning subproblems can be solved in parallel. Moreover, we put

the integer requirements into the production decision-making subproblem with only simple capacity constraints, which can be efficiently solved through dynamic programming. To measure the quality of the solutions generated in the process, we compute an upper bound feasible solution and a lower bound estimate based on the current solution at each iteration. Numerical experiments suggest that the algorithm can converge to a good solution with properly chosen starting point and penalty parameter. We also evaluate our method on the basis of real-world data provided by a leading global corporation in China providing information and communications technology (ICT) solutions.

Risk-Aware Security-Constrained Unit Commitment – CHENG GUO, CLEMSON UNIVERSITY – Daniel Bienstock, Yury Dvorkin, Robert Mieth, Jiayi Wang

To better handle real-time load and wind generation volatility in unit commitment, we present an enhancement to the computation of security-constrained unit commitment (SCUC) problem. More specifically, we propose a two-stage optimization model for SCUC, which aims to provide a risk-aware schedule for power generation. Our model features a data-driven uncertainty set based on principal component analysis, which accommodates both load and wind production volatility and captures locational correlation of uncertain data. To solve the model more efficiently, we develop a decomposition algorithm that can handle nonconvex subproblems. Our extensive experiments on NYISO dataset show that the risk-aware model protects the public from potential high costs caused by adverse circumstances.

Beyond Absolute Continuity: a New Class of Dynamic Risk Measures – LUHAO ZHANG, COLUMBIA UNIVERSITY – Rui Gao, Jincheng Yang

The modern theory of risk measures copes with uncertainty by considering multiple probability measures. While it is often assumed that a reference probability measure exists, under which all relevant probability measures are absolutely continuous, there are examples where this assumption does not hold, such as certain distributional robust functionals. In this talk, we introduce a novel class of dynamic risk measures that do not rely on this assumption. We will discuss its convexity, coherence, and time consistency properties.

3b - TC96 - 512A - Theory and Applications of Robust and Stochastic Integer Programming Chair: WEIJUN XIE

Stochastic Programming for Disaster Housing Assistance Planning – YONGJIA SONG, CLEMSON UNIVERSITY – Sheng Yin Chen

In this paper, we propose and study a framework for disaster housing logistics planning under demand uncertainty. Specifically, we utilize the two-stage chance-constrained stochastic programming models to achieve the balance between logistics operational cost and demand fulfillment especially towards extreme disaster scenarios. To do so, we incorporate two operational modalities, one for the ordinary modality and the other for the emergency modality, and the emergency modality is only allowed to be activated for a certain percentage of scenarios that is specified by the decision maker among all scenarios. The set of scenarios is generated according to a re-

gression model for characterizing the disaster housing demand based on a selected number of independent variables, which is trained offline from historical data. Our preliminary numerical results based on a case study on Hurricane Ian have shown the effectiveness of the proposed approach compared to some standard benchmark approaches and provided managerial insights in disaster housing logistics planning.

Dynamic Transmission Line Switching Amidst Wildfire-Prone Conditions – JUAN ALBERTO ESTRADA GARCIA, UNIVERSITY OF MICHIGAN – Ruiwei Jiang

Power grids are vulnerable to wildfire propagation which can trigger extended power loss by failing transmission lines. During dry seasons, environmental conditions increase the risk of wildfires, exposing power grids to failure conditions. Simultaneously, power system operation can induce the ignition and spread of ongoing wildfires. In this work, we propose a multi-stage optimization model to switch transmission lines dynamically to adapt the transmission network topology to respond to the wildfire while considering the trade-off between operational performance and wildfire propagation. The proposed formulation is a multi-stage model with decision-dependent probabilities. We developed a stochastic nested decomposition algorithm to solve this model and present a case study using a 33-bus transmission system.

Improving the Security of United States Elections with Robust Optimization – BRADLEY STURT, UNIVERSITY OF ILLINOIS CHICAGO – Alex Halderman, Braden L Crimmins

For more than a century, election officials across the United States have inspected voting machines before elections using a procedure called Logic and Accuracy Testing (LAT). This procedure consists of election officials casting a test deck of ballots into each voting machine and confirming the machine produces the expected vote total for each candidate. In this work, we bring a scientific perspective to LAT by introducing the first formal approach to designing test decks with rigorous security guarantees. Specifically, we propose using robust optimization to find test decks that are guaranteed to detect any voting machine misconfiguration that would cause votes to be swapped across candidates. Out of all the test decks with this security guarantee, the robust optimization problem yields the test deck with the minimum number of ballots, thereby minimizing implementation costs for election officials. To facilitate deployment at scale, we developed a practical exact algorithm for solving our robust optimization problems based on mixed-integer optimization and the cutting plane method. In partnership with the Michigan Bureau of Elections, we retrospectively applied our robust optimization approach to all 6928 ballot styles from Michigan's November 2022 general election; this retrospective study reveals that the test decks with rigorous security guarantees obtained by our approach require, on average, only 1.2% more ballots than current practice. Our robust optimization approach has since been piloted in real-world elections by the Michigan Bureau of Elections as a low-cost way to improve election security and increase public trust in democratic institutions.

3b - WA58 - 512A - **Stochastic Integer Programming: Models and Algorithms** Chair: JIM LUEDTKE

Master Surgery Scheduling Problem Considering Uncertainty in Time of Stay in Downstream Units; A Matheuristic Solution Approach – ERFANEH NIKZAD, POLYTECHNIQUE MONTRÉAL – Nadia Lahrichi

In this study, we develop a mathematical model for the master surgery scheduling problem, taking into account uncertainties in the length of stay in downstream units such as wards and ICU beds. Hospitals face constraints due to limited ward and ICU bed availability, with daily capacity determined by patient length of stay. By incorporating post-surgery length of stay as stochastic parameters in both the surgical unit and ICU, we reduce the possibility of surgery cancellations due to bed availability issues. Uncertain parameters are defined within a limited set of scenarios, and a two-stage stochastic programming model is proposed to address these uncertainties. Additionally, we develop a Two-phase matheuristic algorithm to efficiently solve the proposed model. In the first phase, an initial solution is generated, followed by a fix and optimize procedure in the second phase to improve the initial solution. The results obtained indicate the efficiency of the proposed solution method.

Probing-Enhanced Stochastic Programming – JEFF LINDEROTH, UNIVERSITY OF WISCONSIN MADISON –

We consider a two-stage stochastic decision problem where the decision-maker has the opportunity to obtain information about the distribution of the random variables X through a set of discrete actions that we refer to as probing. Specifically, probing allows the decision-maker to observe components of a random vector Y that is jointly-distributed with X . We propose a three-stage optimization model for this problem, wherein the first-stage variables select components of Y to observe, and decisions in subsequent stages must be consistent with the obtained information. In the case that X and Y have finite support, Goel and Grossmann gave a mixed-integer programming formulation of this problem whose size is proportional to the square of cardinality of the sample space of the random variables. We propose to solve the model using bounds obtained from an information-based relaxation, combined with a branching scheme that enforces the consistency of decisions with observed information. The branch-and-bound approach can naturally be combined with sampling in order to estimate both lower and upper bounds on the optimal solution value. We demonstrate the scalability of our approach against the exact MIP formulation on instances of a stochastic facility location problem.

The Chance-constrained Stochastic Diversion Path Problem with Sample Average Approximation – WOJIN KIM, UNIVERSITY OF WISCONSIN MADISON – Jim Luedtke, Cole Smith

The diversion path problem is a network interdiction problem in which the leader modifies arc lengths, and the follower finds the shortest path from the source node to the sink node using the modified arc lengths. The leader's goal is to minimize the cost to have the follower choose a special path called the diversion path. When arc lengths are deterministic, this problem is solvable by linear programming. We study the situation where arc lengths are random, and the leader wishes to minimize the cost to assure that the follower follows the diversion path with probability at least a desired target. This chance-constrained stochastic diversion path problem has been studied by Baycik,

Nguyen, and Smith in the special case when arc lengths are independent and normally distributed. We explore the use of sample average approximation(SAA) to approximately solve this problem with a general distribution. We first derive a compact mixed integer programming formulation of the SAA problem, but find that it is difficult to solve due to poor linear programming relaxation quality. We thus explore a method based on a branch-and-cut decomposition algorithm. Computational results will be presented.

Stochastic facility location problem with outsourcing costs – EDUARDO MORENO, UNIVERSIDAD ADOLFO IBANEZ / GOOGLE FRANCE – Ivana Ljubic

Stochastic facility location problems with outsourcing costs (SFLPOC) optimize facility placement and customer assignment under demand uncertainty. Excess demand beyond a facility's capacity incurs outsourcing costs. This work addresses SFLPOC, aiming to minimize overall expected costs (installation, servicing, and outsourcing). We model SFLPOC as a two-stage stochastic program. While prior work focused on specific assumptions or small scenario sets, we present methods suitable for general probability distributions. For discrete scenario sets, we improve upon classic Benders decomposition by exploiting the second-stage subproblem's structure. To handle general distributions, we partition the probability space, enabling the computation of expected values with fewer scenarios. Coupled with Benders cuts, this provides an exact solution method for common distributions (e.g., Bernoulli, Gaussian). Additionally, we introduce a compact formulation specifically for i.i.d. demand distributions, allowing us to solve even continuous distribution problems to optimality. Computational experiments on established benchmarks demonstrate that our compact formulation consistently finds optimal solutions, while the Benders approach provides strong solutions with proven optimality gaps for general distributions, outperforming sample average approximations.

3b - WB178 - 512A - **Recent Advances to Solve Discrete Stochastic Problems** Chair: AYSENUR KARAGOZ

Accelerating Benders decomposition for solving a sequence of sample average approximation problems – HARSHIT KOTHARI, UNIVERSITY OF WISCONSIN MADISON – Jim Luedtke

Sample average approximation (SAA) is a technique to get good approximate solutions to stochastic programs (SP). When applying SAA, it is often useful to solve multiple SAA problems to obtain a confidence interval on the true optimal value of the SP and also get a better solution. We study techniques to accelerate the solution of this sequence of SAA problems, when solving them via Benders decomposition. We exploit similarities in the problem structure, as the problems just differ in the realizations of the random samples. Our extensive computational experiments of large scale problems provide empirical evidence of the improved efficiency of our algorithm. In addition, we also present theoretical results that provide insight into the algorithm's performance.

Scaled cuts for multistage stochastic mixed-integer programs – WARD ROMEIJNDERS, UNIVERSITY OF GRONINGEN –

We consider multistage stochastic mixed-integer programs with general mixed-integer state variables in all time stages. We develop cutting-plane based algorithms that maintain convex polyhedral outer approximations for all expected cost-to-go functions of the problem. We introduce so-called scaled-cuts to iteratively improve these convex polyhedral outer approximations and prove that the outer approximation in the first stage converges uniformly to the first-stage expected cost-to-go function.

Single-Scenario Facet Preservation for Stochastic Mixed-Integer Programs – AYSENUR KARAGOZ, RICE UNIVERSITY – Andrew Schaefer

We consider improving the polyhedral representation of the extensive form of an SMIP. Given a facet-defining valid inequality for a single-scenario version of the SMIP, we provide conditions under which the same inequality remains facet-defining for the extensive form. Our main result gives necessary and sufficient conditions for a facet-defining inequality for a single-scenario version to be facet-defining for the extensive form. We then present several implications, which show that various recourse structures from the literature satisfy these conditions. For example, for an SMIP with simple recourse, any single-scenario facet is also a facet for the extensive form. For more general recourse structures, we provide additional mild necessary and sufficient conditions.

Test sets to compute opportunity cost matrices – JANOSCH ORTMANN, UQAM – Yuchen Ge, Walter Rei

In stochastic programming, scenarios approximate distributions of unknown parameters, but in many applications the number of scenarios required to realistically model the uncertainty makes the optimisation problem numerically intractable. This motivates the scenario reduction problem: by finding a smaller subset of scenarios, reduce the numerical complexity while keeping the error at an acceptable level.

Recently, problem-based scenario reduction methods based on opportunity cost dissimilarities have shown to be effective. In this setting, opportunity cost means the cost of wrongly predicting scenario 1 when actually scenario 2 happens and can be viewed as a measure of how different these two scenarios are with respect to the optimisation problem at hand.

In this talk, I will explain how test sets can be used to efficiently compute such opportunity cost matrices for stochastic integer programs. I will then describe new algorithms, based on ideas from algebraic geometry, to efficiently compute these test sets. Finally I will show promising numerical results.

3b - ThA937 - 512A - **Discrete Stochastic Programming 1** Chair: MARYAM DARYALAL

A Decomposition Method for Network Revenue Management with Dynamic Pricing and Bundles – EMANUELE CONCAS, ENPC – Roberto Wolfler Calvo, Axel Parmentier

This paper introduces a decomposition approach for revenue management within a network context. Our method focuses on dynamically pricing bundles for a network of flights, while optimizing revenue and considering customer choice behavior.

The novelty of our work lies in its complexity. We aim to offer the optimal assortment, selecting bundles from an exponential large set of potential combinations of additional products. Furthermore, we must determine the optimal price for each time within our selling horizon. When combined with the network aspect, we end up with a problem that is very challenging to solve. We address the problem by using dynamic programming decomposition, by breaking down itineraries into individual legs, and solving them one at a time. The non-feasible solution is then reconstructed with a Lagrangian heuristic. Benchmarks in computational experiments demonstrates that our method yields improvements in both upper bound and expected revenue. This highlights the effectiveness of our approach in network revenue management.

Robust Optimality of Bundling I.I.D. Goods Beyond Finite Variance – PIETER KLEER, TILBURG UNIVERSITY –

For selling bundles of many goods with independent and identically distributed (i.i.d.) valuations, we develop a distributionally robust framework, consisting of a two-player game between seller and nature. The seller has only limited knowledge about the value distribution. Nature chooses a revenue-minimizing distribution from all distributions that comply with the limited knowledge, upon which the seller selects a revenue-maximizing mechanism that counters nature's worst-case choice. When the seller knows the mean and variance of valuations, bundling is known to be an asymptotically optimal deterministic mechanism, achieving a normalized revenue close to the mean. Moving beyond this variance assumption, we assume knowledge of the mean absolute deviation (MAD), accommodating more dispersion and heavy-tailed valuations with infinite variance. We show for a large range of MAD values that bundling remains optimal, but the seller can only guarantee a revenue strictly smaller than the mean. Another noteworthy finding is indifference to the order of play, as both the max-min and min-max versions of the problem yield identical values. This contrasts with deterministic mechanisms and the separate sale of goods, where the order of play significantly impacts outcomes. We further underscore the universality of the optimal bundling price by demonstrating its efficacy in optimizing not only absolute revenue but also the absolute regret and ratio objective among all bundling prices

The SAA Method for Two-Stage Stochastic Programs with Endogenous Uncertainty – MARIA CAROLINA BAZOTTE CORGOZINHO, POLYTECHNIQUE MONTREAL – Margarida Carvalho, Thibaut Vidal

Real-world decision-making problems involve Type 1 decision-dependent uncertainty, where the probability distribution of the stochastic process depends on the model decisions. However, few studies focus on two-stage stochastic programs with this type of endogenous uncertainty, and those that do lack general methodologies. We thus propose herein a general method for solving a class of these programs based on the transformation of random variables, a technique widely employed in probability and statistics. The proposed method is tailored to large-scale problems with discrete or continuous endogenous random variables. The random variable transformation allows the use of the sample average approximation (SAA) method, which provides optimality convergence guarantees under certain conditions. We show that, for some classical distribu-

tions, the proposed method reduces to solving mixed-integer linear or convex programs. Finally, we validate this method by applying it to a network design and facility-protection problem, considering distinct decision-dependent distributions for the random variables. Whereas most distributions result in a nonlinear nonconvex deterministic equivalent program, the proposed method solves mixed-integer linear programs in all cases. In addition, it produces attractive performance estimators for the SAA method in a reasonable computational time and outperforms the case in which the endogenous distribution defines a mixed-integer deterministic equivalent.

A Risk-aware Location-Allocation-Pricing Problem with Stochastic Price-Sensitive Demands – MARYAM DARYALAL, HEC MONTRÉAL –

We consider a capacitated location-allocation-pricing problem in a single-commodity supply chain with stochastic price-sensitive demands, where the location, allocation and pricing decisions are made simultaneously. Under a general risk measure representing an arbitrary risk tolerance policy, the problem is modeled as a two-stage stochastic integer program (2SIP). To solve the problem, we adapt an integer L-shaped method for 2SIPs with a translation invariant monotone risk measure and demonstrate its applicability on three common risk measures for moderate, cautious and pessimistic policies. We also introduce two family of problem-specific cuts to enhance the performance of the algorithm. Our computational experiments are conducted to study the effectiveness of the algorithm, along with a comparison between the impact of different policies on the solution of the problem. Our approach for incorporating the risk and the proposed solution framework is general and can be used for other similar discrete optimization problems that make use of risk measures.

3b - ThB124 - 512A - Sequential Decision Making Under Uncertainty Chair: RUI CHEN

Online Decision Making with Nonconvex Local and Convex Global Constraints – RUI CHEN, CORNELL UNIVERSITY –

Motivated by recent progress on online linear programming and online resource allocation, we study the online decision making problem (ODMP) as a natural generalization. In ODMP, a single decision maker undertakes a sequence of decisions over multiple (likely many) time steps. At each time step, the decision maker makes a locally feasible decision based on information available up to that point, without knowledge of the future. The objective is to maximize the total reward accumulated over time while satisfying some global constraints called goal constraints. Decision made at each step results in a vector that represents the contribution of this local decision to the (global) goal constraints. In the online setting, these global goal constraints are soft constraints that can possibly be violated moderately at the end of the time horizon. ODMP has a strong modeling power, allowing discrete and nonlinear local constraints and general convex global constraints (beyond packing and covering). To handle discreteness and nonlinearity in solving ODMP, we propose a simple Fenchel dual-based online algorithm. Under some stochastic input models, we show that our algorithm achieves sublinear constraint violation de-

terministically in meeting the long-term goals, and sublinear regret in expected reward with respect to the optimal offline decisions. Numerical experiments on an online knapsack problem with fairness-over-time constraints and an assortment optimization problem are conducted to demonstrate the potential of our proposed online algorithm.

Value of Stochastic Solution with Right-Hand Side Uncertainty – HAOMING SHEN, UNIVERSITY OF ARKANSAS – Ruiwei Jiang

We revisit the value of stochastic solution (VSS) in the context of distributional ambiguity. When uncertainty arises from the right-hand side of a two-stage stochastic program, we investigate VSS's upper and lower bounds using distributionally robust and optimistic optimization. Because these bounds are NP-hard to compute exactly, we derive tractable yet tight relaxations for them and demonstrate their effectiveness through numerical experiments.

Regularized MIP Model for Optimal Power Flow with Energy Storage Systems and its Applications – DAHYE HAN, GEORGIA INSTITUTE OF TECHNOLOGY – Nan Jiang, Santanu Dey, Weijun Xie

Incorporating energy storage systems (ESS) into power systems has been studied in many recent works, where binary variables are often introduced to model the complementary nature of battery charging and discharging. A conventional approach for these ESS optimization problems is to relax binary variables and convert the problem into a linear program. However, such linear programming relaxation models can yield unrealistic fractional solutions, such as simultaneous charging and discharging. In this paper, we develop a regularized Mixed-Integer Programming (MIP) model for the ESS optimal power flow (OPF) problem. We prove that under mild conditions, the proposed regularized model admits a zero integrality gap with its linear programming relaxation; hence, it can be solved efficiently. By studying the properties of the regularized MIP model, we show that its optimal solution is also near-optimal to the original ESS OPF problem, thereby providing a valid and tight upper bound for the ESS OPF problem. The use of the regularized MIP model allows us to solve two intractable problems: a two-stage stochastic ESS OPF problem and a trilevel min-max-min network contingency problem.

Two-sided assortment optimization: Adaptivity gaps and approximation algorithms – ALFREDO TORRICO, CORNELL UNIVERSITY – Omar El Housni, Ulysse Hennebelle

Choice-based matching platforms have recently proliferated thanks to their application in labor markets, dating, accommodation and ride-sharing. Due to correlated preferences, these platforms face the challenge of reducing choice congestion among popular options who receive more requests than they can handle, which may lead to sub-optimal market outcomes. To address this challenge, we introduce a general framework for two-sided assortment optimization under general choice preferences. The goal is to maximize the expected number of matches by deciding which assortments are displayed to the agents and the order in which they are shown. In this setting, we identify a wide range of policy classes that matching platforms can use in their design. Our goals are: (1) to measure the value that one class of policies has over another one, and (2) to approximately solve the optimization problem itself for a given class. For (1),

we define the adaptivity gap as the worst-case ratio between the optimal values of two different policy classes. We show tight constant gaps between several classes of policies and that the worst policies (relative to the rest) are those who simultaneously show assortments to everyone. For (2), we separately show constant-approximation guarantees for two policy classes which sit in opposite ends of the policy spectrum.

3b - ThC116 - 512A - Advanced algorithms in stochastic mixed-integer programming Chair: KIBA EK KIM

Strengthening Benders cuts via disjunction – KAIWEN FANG, ARIZONA STATE UNIVERSITY – Geunyeong Byeon

In this paper, we explore the utilization of disjunctive cuts to enhance the strength of cuts within a Benders decomposition algorithm. At a broad level, the proposed algorithm leverages integrality information and Benders cuts to generate cuts tightly aligned with the convex hull of the extended formulation. Importantly, this is achieved via iteration of standard Benders subproblem. Specifically, our study highlights that normalization plays a crucial role in eliminating the generation of overly weak cuts. We conduct a comparative analysis to assess the impact of Benders disjunctive cuts under various normalization settings in capacitated facility location problems.

Stochastic Cutting-plane Dynamic integer Programming – HANBIN YANG, CHINA – Haoxiang Yang

Multistage stochastic mixed-integer linear programming (MS-MILP) is a powerful framework for sequential decision-making under uncertainty. Solving MS-MILP with binary state variables traditionally involves iteratively solving MIP subproblems to generate Lagrangian cuts for approximating future value functions. However, this method can be highly time-consuming and computationally demanding, especially with complex subproblems. This study introduces an innovative LP-based Stochastic Dual Dynamic Programming (SDDP)-type algorithm specifically for MS-MILP problems with binary state variables. This approach eliminates the need to solve MIP subproblems. Using a geometric perspective of the connection between the convex envelope of a value function and the convex hull of its feasible region, which uses the cutting-planes tree algorithm to refine the LP relaxation of lifted subproblems' feasible region and generate Benders' cuts to approximate the value functions, leading to a more accurate approximation of the value function. A key achievement of this research is demonstrating the finite convergence of this algorithm, making it a more practical and efficient solution for complex MS-MILP problems. The complexity of solving a particular case of the problem, where uncertainty in the subproblems is eliminated from the constraints through reformulation, has been significantly reduced. This modification enables the efficient resolution of problems with extended time horizons and numerous realizations at each stage.

Decomposition Algorithm for Two-Stage Distributionally Robust Optimization over 1-Wasserstein Balls: Enhancing Performance and Scalability – GEUNYEONG BYEON, ARIZONA STATE UNIVERSITY – Kaiwen Fang, Kibaek Kim

We present a cutting-plane-based decomposition algorithm along with several algorithmic enhancements designed to address two-stage distributionally robust optimization problems over 1-Wasserstein balls. Through a numerical analysis on a facility location problem with demand uncertainty, we highlight the potential benefits of employing a particular distance metric for enhanced out-of-sample performance, while also examining its computational trade-offs. Furthermore, we explore alternative approaches aimed at enhancing the scalability of our proposed algorithm.

3c - MB108 - 511D - Applications of Robust and Distributionally Robust Optimization Chair:
 ZHEN YANG 

Budget-Driven Multi-Period Hub Location: A Robust Time Series Approach – ZHI CHEN, THE CHINESE UNIVERSITY OF HONG KONG – Jie Hu, Shuming Wang

We study the multi-period hub location problem with uncertain periodic demands. In particular, we construct a nested ambiguity set that characterizes uncertain periodic demands via a general multivariate time series model, and to ensure stable periodic cost flows, we propose to constrain each expected periodic cost within a budget while maximizing the robustness level (i.e., the size) of the ambiguity set. Statistically, the nested ambiguity set ensures the model's solution to enjoy finite-sample performance guarantees under certain regularity conditions on the underlying process of the stochastic demand. Computationally, the uncapacitated model can be efficiently solved via a bisection search algorithm. Numerical experiments demonstrate the attractiveness and competitiveness of our proposed modeling and solution approaches.

Network Flow Models for Robust Binary Optimization with Selective Adaptability – IAN YIHANG ZHU, NUS BUSINESS SCHOOL – Merve Bodur, Timothy Chan

Adaptive robust optimization problems have received significant attention in recent years, but remain notoriously difficult to solve when recourse decisions are discrete in nature. In this paper, we propose new reformulation techniques for adaptive robust binary optimization problems (ARBO) with objective uncertainty. Our main contribution revolves around a collection of exact and approximate network flow reformulations for the ARBO problem, which we develop by building upon ideas from the decision diagram literature.

Robust Optimization with Moment-Dispersion Ambiguity – LI CHEN, THE UNIVERSITY OF SYDNEY – Chenyi Fu, Fan Si, Melvyn Sim, Peng Xiong

Robust optimization presents a compelling methodology for optimization under uncertainty, providing a practical, ambiguity-averse evaluation of risk when the probability distribution is encapsulated by an ambiguity set. We introduce the moment-dispersion ambiguity set, an improvement on the moment-based set, enabling separate characterization of a random variable's central location, dispersion, and support. To describe dispersion, we define the dispersion characteristic function, capturing complex attributes like sub-Gaussian and asymmetric dispersion, and its associated dispersion characteristic set, which serves as the input format for representing dispersion

ambiguity in algebraic modeling tools. We devise a process for constructing and integrating ambiguity sets, showcasing their modeling flexibility. In particular, we introduce the independence propensity hyper-parameter to foster joint ambiguity set creation for multiple random variables, enhancing our model's real-world applicability and facilitating varying inter-dependence characterization without needing a correlation matrix. For ambiguous risk assessment over moment-dispersion ambiguity sets, we develop safe tractable approximations for assessing entropic risks linked with affine and convex piecewise affine cost functions, accommodating varying risk tolerances. Lastly, we substantiate our approach with two numerical case studies involving appointment scheduling and portfolio optimization. By adjusting the independence propensity hyper-parameter, we illustrate how our model can deliver robust yet less conservative solutions compared to existing moment-based robust optimization models and sample-based approaches, which is particularly useful when only marginal information is accessible to the decision-maker.

Wasserstein Regularization for 0-1 Loss – ZHEN YANG, THE UNIVERSITY OF TEXAS AT AUSTIN – Rui Gao

Wasserstein distributionally robust optimization (DRO) seeks robust solutions by safeguarding against data variations within a specified Wasserstein ball. While its regularization effects and generalization capabilities have been extensively explored for continuous losses, directly applying these insights to the 0-1 loss presents significant challenges. This study investigates this issue within the context of linear classification, establishing a connection between Wasserstein DRO with the 0-1 loss and a novel regularization framework. In this framework, the regularization term is defined as a polynomial function of the Wasserstein ball's radius and the data density near the decision boundary. We propose a radius selection rule that guarantees finite-sample performance. Our findings highlight a fundamental distinction in radius selection between 0-1 and continuous losses: a smaller radius, contrary to the conventional root-n rule, is often sufficient in practical scenarios. Numerical experiments validate the efficiency of our suggested radius selection rule.

3c - TA99 - 510D - Robust Optimization and Machine Learning 1 Chair: ARAS SELVI 

Platform Design for the First Mile of Commodity Supply Chains – SERGIO CAMELO, STANFORD UNIVERSITY – Dan Iancu, Joann De Zegher

We propose a data-driven platform that provides traceability to the first mile of agricultural supply chains by coordinating the transactions of farmers and intermediaries. We model unique aspects of the supply chain, including pre-existing informal relationships between farmers and intermediaries, and we develop algorithms to solve real-world instances. We test the results on data from the palm oil supply chain and show the platform's potential to reduce costs and increase farmers' welfare.

An MILP-Based Solution Scheme for Factored and Robust Factored Markov Decision Processes – MAN CHUNG YUE, THE UNIVERSITY OF HONG KONG – Wolfram Wiesemann, Huikang Liu
 Factored Markov decision processes (MDPs) are a prominent

paradigm within the artificial intelligence community for modeling and solving large-scale MDPs whose rewards and dynamics decompose into smaller, loosely interacting components. Through the use of dynamic Bayesian networks and context-specific independence, factored MDPs can achieve an exponential reduction in the state space of an MDP and thus scale to problem sizes that are beyond the reach of classical MDP algorithms. However, factored MDPs are typically solved using custom-designed algorithms that can require meticulous implementations and considerable fine-tuning. In this paper, we propose a mathematical programming approach to solving factored MDPs. In contrast to existing solution schemes, our approach leverages off-the-shelf solvers, which allows for a streamlined implementation and maintenance; it effectively capitalizes on the factored structure present in both state and action spaces; and it readily extends to the largely unexplored class of robust factored MDPs, whose transition kernels are only known to reside in a pre-specified ambiguity set. Our numerical experiments demonstrate the potential of our approach.

Wasserstein Adversarial Logistic Regression with Synthetic Data – ARAS SELVI, IMPERIAL COLLEGE LONDON – Eleonora Kreacic, Mohsen Ghassemi, Vamsi Potluru, Tucker Balch, Manuela Veloso

Empirical risk minimization often fails to provide robustness against adversarial attacks in test data, causing poor out-of-sample performance. Adversarial training (AT) has thus emerged as the de facto standard for hedging against such attacks. Although robust against adversarial attacks, AT does not address robustness against distributional ambiguity and is prone to overfitting. In order to address this, we study adversarial training of logistic regression (LR) over a Wasserstein ambiguity set around the empirical distribution. Furthermore, we develop a framework to effectively leverage synthetic data in distributionally robust AT in order to reduce the conservatism of distributional robustness. Focusing on the resulting distributionally robust adversarial LR with synthetic data, we analyze the complexity and properties of the underlying optimization problem, develop tractable approximation algorithms, and demonstrate that our method consistently outperforms benchmark models on standard real-world benchmark datasets.

End-to-end Conditional Robust Optimization – ABHILASH CHENREDDY, HEC MONTRÉAL – Erick Delage

The field of Contextual Optimization (CO) integrates machine learning and optimization to solve decision making problems under uncertainty. Recently, a risk sensitive variant of CO, known as Conditional Robust Optimization (CRO), combines uncertainty quantification with robust optimization in order to promote safety and reliability in high stake applications. Exploiting modern differentiable optimization methods, we propose a novel end-to-end approach to train a CRO model in a way that accounts for both the empirical risk of the prescribed decisions and the quality of conditional coverage of the contextual uncertainty set that supports them. While guarantees of success for the latter objective are impossible to obtain from the point of view of conformal prediction theory, high quality conditional coverage is achieved empirically by ingeniously employing a logistic regression differentiable layer within the calculation of coverage quality in our training loss. We show that the proposed training algorithms produce decisions that

outperform the traditional estimate then optimize approaches.

3c - TA170 - 511E - Recent Advances in the Theory of Stochastic and Robust Optimization Chair: HAOMING SHEN

Robust Data-driven Prescriptiveness Optimization – ANGELOS GEORGHIOU, UNIVERSITY OF CYPRUS – Mehran Poursoltani, Erick Delage

The abundance of data has led to the emergence of a variety of optimization techniques that attempt to leverage available side information to provide more anticipative decisions. The wide range of methods and contexts of application have motivated the design of a universal unitless measure of performance known as the coefficient of prescriptiveness. This coefficient was designed to quantify both the quality of contextual decisions compared to a reference one and the prescriptive power of side information. To identify policies that maximize the former in a data-driven context, this paper introduces a distributionally robust contextual optimization model where the coefficient of prescriptiveness substitutes for the classical empirical risk minimization objective. We present a bisection algorithm to solve this model, which relies on solving a series of linear programs when the distributional ambiguity set has an appropriate nested form and polyhedral structure. Studying a contextual shortest path problem, we evaluate the robustness of the resulting policies against alternative methods when the out-of-sample dataset is subject to varying amounts of distribution shift.

On the Adversarial Robustness of Benjamini-Hochberg – LOUIS CHEN, NAVAL POSTGRADUATE SCHOOL – Matan Seri, Roberto Szechtman

The Benjamini-Hochberg (BH) procedure is widely used to control the false detection rate (FDR) in large-scale hypothesis testing. Although this control may be distributionally robust, we investigate its adversarial robustness to perturbations of the p-values. In particular, we present a class of data-perturbation algorithms, discuss when BH does and does not exhibit adversarial robustness, and perform computational experiments. With our algorithms, we demonstrate that it is possible for BH's control to be badly broken with relatively few (even just one) p-value perturbation(s), and provide non-asymptotic guarantees on the expected adversarial-adjustment to FDR. Our technical analysis involves reframing the BH procedure as a "balls into bins" process and drawing a connection to generalized ballot problems that facilitates an information-theoretic approach for deriving non-asymptotic lower bounds. This study demonstrates an instance of how distributional robustness may come at the price of adversarial robustness.

Fourth-order Marginal Moment Model: Reformulations and Applications – LI XIAOBO, NATIONAL UNIVERSITY OF SINGAPORE – Mingyang Fu

This paper investigates the bounds on the expectation of combinatorial optimization given moment information for each individual random variable. A popular approach to solving this problem, known as the marginal moment model (MMM), is to reformulate it as a semidefinite program (SDP). In this paper, we investigate the structure of MMM with up to fourth-order

marginal moments and reformulate them as second-order cone programs (SOCP). Additionally, we establish that this SOCP formulation is equivalent to a convex optimization problem over the convex hull of the feasible region of the original combinatorial optimization, presenting closed-form expressions for both the objective function and its derivative. These reformulations enable more efficient computation of the bounds and persistency value. In addition, we explore two types of ambiguity sets characterized by incomplete moment information. We further discuss the relationship between the aforementioned MMMs and another widely-used model, the marginal distribution model (MDM). Beyond bounding the worst-case expectations, our approaches can be modified to bound the worst-case conditional value at risk (CVaR) of the combinatorial optimization.

Building on this theoretical advancement, we explore two applications. First, we consider the project crashing problem, wherein both the means and variances of activity durations can be controlled through effort. We demonstrate that the distributionally robust project crashing problem, incorporating up to fourth-order moment information, can be reformulated as either an SOCP or a convex minimization over a simple polytope. Numerical analysis reveals that MMM with fourth moment information yields tighter bounds on expected delays and requires a significantly smaller budget than the mean-variance model for a fixed delay guarantee. Second, we apply our reformulations to solve the distributionally robust newsvendor problem with moment information, extending the well-known Scarf’s model. We derive several new closed-form solutions and explore how the optimal order quantities depend on skewness and kurtosis. Numerically, we show that incorporating additional moment information can lead to better performance, especially in the high service level regime.

3c - TA938 - 511D - **Robust and Distributionally Robust Optimization 1** Chair: MATTEO PETRIS

Cut-Generation Algorithms for Uncertain Unit Commitment under Mixed-Information Models – VITOR LUIZ PINTO DE PINA FERREIRA, ÉCOLE DES PONTS PARISTECH – Vincent Leclère, Ali Hellani, Anna Robert, Marco Marchese, Nouha Dkhili, Sami Ghazouani

The classical Unit Commitment problem is a MILP that can optimize thermal unit schedules to minimize costs and/or emissions while meeting demand and operational constraints inside and in-between stages (e.g. minimum uptime constraints). Integrating renewable energy introduces intermittency, leading to Uncertain Unit Commitment (UUC) models that necessitate efficient computational methods. Our collaboration with TotalEnergies focuses on an isolated industrial microgrid with renewable, thermal, and energy storage resources. Addressing uncertainty involves encoding risk attitudes, with approaches ranging from risk neutral—minimizing the expected objective—to robust—minimizing worst-case—, the latter of which is our focus. Furthermore, information models dictate the dependency of decisions on random events. As an approximation, we investigate models where different variables have different dependencies. In our proposal, binary decisions over the entire horizon are made at the first time step and continu-

ous decisions are made as they need to be executed, ensuring tractability and feasibility. Efficiently solving UUC problems demands specialized techniques. We propose cut-based decomposition methods, including adaptations of Stochastic and Robust Dual Dynamic Programming algorithms. We assess algorithmic performance and solution quality through numerical experiments.

Distributionally robust optimization using optimal transport for Gaussian mixture models – ZUKUI LI, UNIVERSITY OF ALBERTA –

Distributionally robust optimization (DRO) is currently a hot topic in literature due to its ability to provide robust solutions in the face of uncertain parameters with unknown probability distributions. The core idea behind DRO is to find a solution that remains resilient even under the worst-case scenario by constructing an ambiguity set of probability distributions around the nominal distribution. This nominal distribution is typically estimated from sampled data.

Among the methods used to construct ambiguity sets, metric-based approaches are popular. These methods involve defining an ambiguity set comprising all candidate distributions that are within a certain distance from the nominal distribution, as prescribed by a chosen probability distance metric. The Wasserstein distance is a well-studied metric for DRO. However, the conventional formulation is only manageable under specific assumptions on the constraint functions and may lead to overly cautious solutions, ignoring any existing distributional characteristics of the samples, such as multimodality.

To overcome the limitations of traditional approaches, our work proposes a variant of the optimal transport problem between Gaussian mixture models (GMM) to construct the ambiguity set for DRO. In this variant, optimal transport is performed between discrete measures, with supports represented by the Gaussian components of the GMM. In our method, the nominal distribution is estimated as a GMM from the sampled data, which often exhibits multimodality in practical scenarios. We also derive a tractable formulation for the distributionally robust chance-constrained problem using the GMM-based ambiguity set, approximating the probabilistic constraints with Conditional Value-at-Risk (CVaR). Case studies demonstrate that our proposed model outperforms traditional Wasserstein DRO methods when dealing with multimodal distributed uncertainty.

Robust approaches for the Kidney Exchange Problem – MATTEO PETRIS, ESSEC BUSINESS SCHOOL – Claudia Archetti, Ivana Ljubic

The kidney Exchange Problem (KEP) can be defined on a directed weighted graph where the vertices represent the altruistic donors and the incompatible patient-donor pairs and the arcs represent the possible transplants. The arc weights encode the medical benefit which is a measure of the quality of the associated transplant. In such a graph, kidney exchanges can be modelled as circuits starting and ending in an incompatible pair or as paths starting at an altruistic donor. Due to medical reasons, circuits and paths are subject to length constraints. The aim of the KEP is to determine a set of disjoint kidney exchanges of maximal medical benefit or maximal cardinality (all weights equal to one). In this work, we consider two types of uncertainty in the KEP which stem from the estimation of the medical benefit (weights of the arcs) and from the failure

of a transplantation (existence of an arc). We model both uncertainties by means of uncertainty sets with constant budget. The robust approach entails finding the best KEP solution in the worst-case scenario within the uncertainty set. We modelled the robust counterpart via a max-min formulation which is defined on exponentially-many variables associated with the circuits and paths. We propose two exact approaches to solve it. The first one exploits the well-known result of Bertsimas and Sim, whereas the second one is based on a reformulation of the robust counterpart to a single-level problem. Both solution approaches rely on a Branch-Price-and-Cut algorithm where the exponentially-many variables are dynamically generated. The computational experiments prove the efficiency of our approaches.

3c - TB100 - 511D - **Robust Optimization and Machine Learning 2** Chair: MENG MENG LI ↔

Bayesian Risk-Averse Reinforcement Learning – YUHAO WANG, GEORGIA INSTITUTE OF TECHNOLOGY – Enlu Zhou

Reinforcement Learning (RL) has become a powerful tool in various fields such as autonomous driving and robotics in recent years. Within RL, an agent continuously interacts with its environment by choosing actions, observing rewards, and transitioning to new states. However, interacting with the actual physical environment can be costly or even unfeasible, which has led to the development of offline RL. In offline RL, the agent operates within a simulated training environment that is constructed based on previously observed data to approximate the real environment where the policy will ultimately be applied.

Directly implementing the policy learned from the training environment in the real environment can be risky, because of model misspecification, which is caused by the lack of historical data. To account for this model misspecification, we adopt a formulation of Bayesian risk MDP (BRMDP) with infinite horizon, which uses Bayesian posterior to estimate the transition model and impose a risk measure to account for the model uncertainty. We derive the asymptotic normality that characterizes the difference between the Bayesian risk value function and the original value function under true unknown distribution. The result indicates the Bayesian risk-averse approach tends to pessimistically underestimate the original value function and such a gap increases with stronger risk-aversion but decreases with more available data.

Robust Regret Markov Decision Processes – CHIN PANG HO, CITY UNIVERSITY OF HONG KONG – Yuqi Zha, Zhi Chen, Wolfram Wiesemann

Markov decision processes (MDPs) are one of the standard frameworks for reinforcement learning and sequential decision-making problems under uncertainty. However, a practical limitation of standard MDPs is their assumption of having precise knowledge of model parameters, such as transition probabilities and rewards. This assumption becomes problematic in real-world scenarios where the effect of these estimation errors in these parameters can accumulate over time, leading to output policies that are sensitive to errors and may fail catastrophically when deployed. To address this challenge, robust MDPs have emerged as a popular alternative. Robust MDPs aim to optimize worst-case performance by considering the ambiguity

in the model parameters. While robust MDPs offer reliable policies even with limited data, they are often overly conservative. In this presentation, we propose to adopt the regret decision criterion to tackle this issue. We introduce a novel framework called robust regret MDPs, which specifically optimize stepwise regret under model ambiguity. By incorporating this criterion, we aim to strike a balance between policy reliability and the overly conservative nature of worst-case performances. The proposed robust regret MDPs can be solved by the value iteration method. Moreover, we present novel algorithms for efficiently computing the associated Bellman updates for large instances.

Distributionally Robust Path Integral Control – DUO ZHOU, UNIVERSITY OF ILLINOIS URBANA CHAMPAIGN – HYNK PARK, GRANI A HANASUSANTO, TAKASHI TANAKA

We consider a continuous-time continuous-space stochastic optimal control problem, where the controller lacks exact knowledge of the underlying diffusion process, relying instead on a finite set of historical disturbance trajectories. In situations where data collection is limited, the controller synthesized from empirical data may exhibit poor performance. To address this issue, we introduce a novel approach named Distributionally Robust Path Integral (DRPI). The proposed method employs distributionally robust optimization (DRO) to robustify the resulting policy against the unknown diffusion process. Notably, the DRPI scheme shows similarities with risk-sensitive control, which enables us to utilize the path integral control (PIC) framework as an efficient solution scheme. We derive theoretical performance guarantees for the DRPI scheme, which closely aligns with selecting a risk parameter in risk-sensitive control. We validate the efficacy of our scheme and showcase its superiority when compared to risk-neutral PIC policies in the absence of the true diffusion process.

Towards Optimal Offline Reinforcement Learning – MENG MENG LI, EPFL – Daniel Kuhn, Tobias Sutter

We study offline reinforcement learning problems with a long-run average reward objective. The state-action pairs generated by any fixed behavioral policy thus follow a Markov chain, and the empirical state-action-next-state distribution satisfies a large deviations principle. We use the rate function of this large deviations principle to construct an uncertainty set for the unknown true state-action-next-state distribution. We also construct a distribution shift transformation that maps any distribution in this uncertainty set to a state-action-next-state distribution of the Markov chain generated by a fixed evaluation policy, which may differ from the behavioral policy. We prove that the worst-case average reward of the evaluation policy with respect to all distributions in the shifted uncertainty set provides, in a rigorous statistical sense, the least conservative estimator for the average reward under the unknown true distribution. This guarantee is available even if one has only access to one single trajectory of serially correlated state-action pairs. The emerging robust optimization problem can be viewed as a robust Markov decision process with a non-rectangular uncertainty set. We develop an efficient first-order algorithm for solving such problems. Numerical experiments show that our methods compare favorably against state-of-the-art methods.

Superior and Light Pareto Robust Optimization for IMRT Treatment Planning – HOUA MAHMOUDZADEH, UNIVERSITY OF WATERLOO – Fahimeh Rahimi, Danielle Ripsman, Hossein Abouee Mehrizi

Robust optimization (RO) is a well-known methodology used in applications where the worst-case scenario of uncertainty must be prevented. A drawback of this approach is that it does not optimize for non-worst-case scenarios and may result in over-conservative solutions. Pareto robust optimization (PRO) is a recent methodology that can further optimize for non-worst-case scenarios and generate less conservative robust solutions when a non-worst-case scenario occurs. In this paper, we extend the theory of Pareto robust optimization to enhance its applicability to practical settings where expert knowledge on likely uncertain scenarios exists. We define and characterize superior PRO solutions and develop algorithms that can partition the set of RO solutions to determine the Pareto robust frontier of superior PRO solutions. We demonstrate how this Pareto frontier can aid decision-makers in determining superior solutions, based on their expert knowledge of the application. To further reduce the conservatism of the PRO solutions, we also introduce the concept of light PRO solutions that slightly sacrifice the worst-case performance to improve non-worst-case behavior while maintaining the robustness of the solution guaranteed. We quantify the gain in non-worst-case scenarios as a function of worst-case loss, discuss the theoretical properties of this gain function, and analyze the trade-off between worst-case optimality and non-worst-case performance. We illustrate the application of the proposed approach to the radiation therapy treatment planning for breast cancer using realistic patient data sets. The results of both a basic and a clinical case study confirm that the proposed approach can significantly improve the quality of the solution in non-worst-case scenarios while maintaining robustness and limiting the optimality error of the worst-case outcome within a reasonable threshold.

Dealing With Uncertainty Over Time When Optimizing Industrial Decarbonization Pathways – JUSTIN STARREVELD, UNIVERSITY OF AMSTERDAM – Dick Den Hertog

In our research we look to optimize long-term strategic decisions regarding the production, transportation and storage of hydrogen for an industrial cluster in the Netherlands. As the long-term future is highly uncertain, there is much parametric uncertainty regarding future developments in technology, economics and policymaking. As a first step we apply Robustness Analysis to examine the robustness of the nominal solution, where we extend the methodology to a multistage adaptive optimization setting. As a second step, we apply our novel method “Robust Optimization by Iterative Scenario Sampling and Statistical Testing” to obtain a more robust solution.

A Holistic Framework for Decarbonization at OCP – PERIKLIS PETRIDIS, MIT – Dimitris Bertsimas, Vassilis Digalakis

In light of rapidly decreasing costs of renewable energy technologies, companies are increasingly looking to integrate renewable energy sources into their energy mix. In the last decade, the costs of solar, wind, and battery energy storage systems

have fallen by 90%, 70%, and 85% respectively, making clean energy projects increasingly cost-effective compared to fossil fuel-based energy, especially for large-scale operations with long-term planning horizons. However, the intermittent nature of renewable energy sources, such as solar and wind, poses significant reliability challenges for energy systems, especially when coupled with the need to meet growing energy demands. In this work, we collaborate with OCP, one of the world’s largest exporters of phosphate rock and fertilizers, to achieve their ambitious goal of decarbonizing their energy mix by 2027, while expanding their operations to meet a projected energy demand of approximately 12 TWh per annum by 2030, which is four times their current energy consumption and nearly a quarter of Morocco’s current electricity production. We develop a holistic, tractable, and robust framework for decarbonization at OCP, which addresses the challenges of integrating solar and wind energy sources over a long-term planning horizon. This framework consists of two main components: (i) a robust strategic capacity expansion planning model that determines the up-front capital expenditures (CAPEX) for renewable energy installations and battery placement decisions, and (ii) an adaptive robust economic dispatch model that provides operational policies on an hourly basis for energy transmission, generation, and procurement, and also serves to validate the robustness of the strategic decisions. Our framework is designed to be utilized by a broad range of stakeholders in the renewable energy sector, with a focus on facilitating bankable feasibility studies, including detailed modeling of the Levelized Cost of Electricity (LCOE), strategic CAPEX planning, and validation of robustness guarantees, akin to P50/P95/P99 exceedance probabilities used by lenders. After simulating the final robust strategic decisions, we achieve a 97% clean energy utilization starting in 2027.

Robust Facility Location in Disaster Preparation for Earthquakes with Aftershocks – MINAKSHI PUNAM MANDAL, ESSEC BUSINESS SCHOOL – Laurent Alfandari, Ivana Ljubic

Earthquakes are one of the deadliest natural disasters that can cause catastrophic damages, leading to loss of life and property and displacing thousands of people. Often, most earthquakes are followed by major and minor aftershocks. The aftershocks could also lead to other secondary disasters like tsunamis, floods, avalanches, building collapses, etc. Thus, it is crucial to take into account the damages caused by aftershocks in the planning process. In this paper, we study an uncapacitated facility location problem for storing relief materials in the event of such a disaster. We adopt a two-stage budgeted-robustness approach. We consider a major earthquake in the first stage, followed by aftershocks in the second stage. We propose several Mixed-Integer Linear Programming formulations for our problem and employ Branch-and-Cut methods to solve two of them. We analyze the performance of our models on synthetically generated instances and provide a case study on Turkey which is highly prone to earthquakes. Finally, we provide managerial insights regarding, in particular, the relevance of modeling the second-stage aftershocks and its impact on optimal locations.

Reinforcement Learning – JOSE BLANCHET, STANFORD UNIVERSITY – Shengbo Wang, Nian Si, Zhengyuan Zhou

Motivated by the need for a robust policy in the face of environment shifts between training and the deployment, we contribute to the theoretical foundation of distributionally robust reinforcement learning (DRRL). This is accomplished through a comprehensive modeling framework centered around distributionally robust Markov decision processes (DRMDPs). This framework obliges the decision maker to choose an optimal policy under the worst-case distributional shift orchestrated by an adversary. By unifying and extending existing formulations, we rigorously construct DRMDPs that embraces various modeling attributes for both the decision maker and the adversary. These attributes include adaptability granularity, exploring history-dependent, Markov, and Markov time-homogeneous decision maker and adversary dynamics. Additionally, we delve into the flexibility of shifts induced by the adversary, examining SA and S-rectangularity. Within this DRMDP framework, we investigate conditions for the existence or absence of the dynamic programming principle (DPP). From an algorithmic standpoint, the existence of DPP holds significant implications, as the vast majority of existing data and computationally efficiency RL algorithms are reliant on the DPP. To study its existence, we comprehensively examine combinations of controller and adversary attributes, providing streamlined proofs grounded in a unified methodology. We also offer counterexamples for settings in which a DPP with full generality is absent.

Regularization for Adversarial Robust Learning – JIE WANG, GEORGIA INSTITUTE OF TECHNOLOGY – Rui Gao, Xie Yao

Despite the growing prevalence of artificial neural networks in real-world applications, their vulnerability to adversarial attacks remains to be a significant concern, which motivates us to investigate the robustness of machine learning models. While various heuristics aim to optimize the distributionally robust risk using the ∞ -Wasserstein metric, such a notion of robustness frequently encounters computation intractability. To tackle the computational challenge, we develop a novel approach to adversarial training that integrates f-divergence regularization into the distributionally robust risk function. This regularization brings a notable improvement in computation compared with the original formulation. We develop stochastic gradient methods with biased gradient oracles to solve this problem efficiently. Theoretical convergence analysis reveals that this algorithm has near-optimal sample complexity. Moreover, we establish the regularization effects and demonstrate this formulation is asymptotic equivalence to a regularized empirical risk minimization (ERM) framework, by considering various scaling regimes of the regularization parameter η and robustness level ρ . These regimes yield gradient norm regularization, variance regularization, or a smoothed gradient norm regularization that interpolates between these extremes. We numerically validate our proposed method in supervised learning and reinforcement learning applications and showcase its state-of-the-art performance against various adversarial attacks.

Distributionally Robust Affine Markov Decision Processes – KARTHYEK MURTHY, SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

– Divya Padmanabhan

Markov Decision Processes (MDPs) serve as a powerful model for sequential decision-making under uncertainty. Despite their prominence, a key challenge in employing MDPs for large-scale planning and control applications is the significant estimation uncertainty present in the model parameters when they are estimated from limited data. To mitigate its impact, we consider a robust approach set up to find policies yielding the highest average reward achievable uniformly over any state-transition model from an uncertainty set. Our primary contribution is in exhibiting such distributionally robust MDP models for which (i) the robust value function is simply an affine function of the present state, and as a result, (ii) optimal policies can be found by solving lower-dimensional recursions without suffering the curse of dimensionality. We achieve this for state-transition and reward models with an affine dependence on the present state-action pair, and action spaces that are linearly constrained based on the present state in a naturally decomposable manner. As a consequence, these robust affine MDP models become a novel and expressive addition to the limited collection of MDP instances known to be simultaneously immune to both estimation uncertainty and dimension-related challenges.

Enhancing Data-Driven Distributionally Robust Optimization: Jointly Incorporating First-Order Moment and Wasserstein Distance – CAIHUA CHEN, CHCHEN@NJU.EDU.CN – Yinyu Ye, Tianxiao Wang

Distribution distance and moment information are two widely used methods for characterizing distribution properties. In this research, we study the distributionally robust optimization problems where the ambiguity sets encompass both the Wasserstein distance and first-order moment information. These ambiguity sets offer a more effective approach for leveraging sample data, thereby mitigating conservativeness. To harness the advantages of the finite-sum and asymmetrical structure inherent in such problems, we propose a novel variance-reduced asymmetrical primal-dual algorithm for efficiently solving the associated dual problem. Additionally, for a specific class of problems, we exploit the regularization structure to achieve enhanced theoretical results. Through extensive numerical investigations involving three distinct problem domains (namely, machine learning, portfolio selection, and the newsvendor problem), we compare the performance of our proposed model against that of an ambiguity set based solely on the Wasserstein distance. The results show the substantial improvement in out-of-sample performance achieved by incorporating first-order information.

3c - WB125 - 511E - **Robust Optimization and Machine Learning 4** Chair: MENG MENG LI [↩](#)

First-order data-driven distributionally robust optimization – SHIMRIT SHTERN, TECHNION ISRAEL INSTITUTE OF TECHNOLOGY –

During the era of big data, Sample Average Approximation (SAA) became prominent as a data-driven technique for approximating stochastic optimization problems. This is partly because SAA lends itself to efficient solutions through first-order methods. Nevertheless, when the size of the dataset is limited, SAA is susceptible to overfitting. Recently, distributionally robust optimization (DRO) based on Wasserstein-

type ambiguity sets has emerged as an alternative data-driven method capable of mitigating overfitting. However, it generally entails a higher computational overhead compared to SAA. In this work, we aim tackle the computational roadblock of applying these data-driven DROs by demonstrating their compatibility with first-order methods. Specifically, we show that one- and two-stage optimization problems whose feasibility does not depend on the data realizations are amenable to first-order methods, in both the offline and online settings. Furthermore, we introduce novel probabilistic guarantees for implementing the online version of Wasserstein-based DRO. In cases where solution feasibility depends on the data realizations, we propose a first-order methods which relies on approximating the feasible set. We validate the effectiveness of our suggested methods through numerical experiments.

New algorithm for the difference of convex functions optimization problem – AHARON BEN TAL, TECHNION ISRAEL INSTITUTE OF TECHNOLOGY – Lubov Tetraushvili

Maximizing the difference of 2 convex functions over a convex feasible set (the so called DCA problem) is a hard problem. There is very large number of publications for solving this problem. Most of them are variations of the widely used DCA algorithm. The performance of this algorithm, for reaching a good approximation of a global solution, depends crucially on the choice of its starting point. We introduce a new algorithm (NDC) which is notoriously different then DCA. A major effort in NDC is the generation of a good starting point. This is obtained by using COMAX, an algorithm for maximizing globally a constrained convex problem. In phase, one of COMAX a starting point is derived by approximating the objective function with a specific quadratic function (Taylor expansion around the MINIMIZER of the max-convex function) and then approximating the convex feasible set by an ellipsoid. The result is a, so called, "hidden Convexity problem" which can be easily solved globally. The optimal solution is a basis for obtaining a good starting point for a standard phase 2 gradient ascent algorithm converging to a candidate for the global maximum. The performance of the NDC algorithm is tested numerically, and the results are compared to the performance of the classical DCA.

Towards Watermarking Tabular Datasets – VAMSI POTLURU, J.P. MORGAN AI RESEARCH – Tabular datasets play an important role for many industries including financial, healthcare, and insurance among others involving machine learning tasks. We propose two novel approaches (with distortion and distortion-free) for watermarking tabular datasets arising from generative models inspired by techniques for text LLMs. Experiments are shown to validate our approaches on various datasets and are robust to various adversarial attacks.

Online MDP with Transition Prototypes: A Robust Adaptive Approach – SHUO SUN, UNIVERSITY OF CALIFORNIA, BERKELEY – Meng Qi, Zuo Jun Max Shen

In robust Markov Decision Process (MDP), an ambiguity set is constructed on the collected data. However, in many cases, data arrives as a stream, leading to the question of how to adaptively construct the ambiguity set as we collect more data. In this work, we consider a setting where we need to identify

the real state-action transition probability from a finite set of candidate prototypes. We construct an online adaptive ambiguity set by updating our belief over the candidates. We prove the efficiency of this adaptive ambiguity set by providing regret analysis for the policy given by the solving the robust optimization problem with this adaptive ambiguity set. We also provide numerical results to demonstrate the empirical performances.

3c - WB939 - 511D - Robust and Distributionally Robust Optimization 2 Chair: FABRICIO OLIVEIRA 

Robust Optimization Under Controllable Uncertainty – EVA LEY, TU BRAUNSCHWEIG – Rebecca Marx, Maximilian Merkert, Tim Niemann, Sebastian Stiller

Applications for optimization with uncertain data in practice often feature a possibility to reduce the uncertainty at a given query cost, e.g., by conducting measurements, surveys, or paying a third party in advance to limit the deviations. To model this type of applications, we introduce the concept of optimization problems under controllable uncertainty (OCU). For an OCU we assume the uncertain cost parameters to lie in bounded, closed intervals. The optimizer can shrink each of these intervals around a certain value called hedging point, possibly reducing it to a single point. Depending on whether the hedging points are known in advance or not, different types of OCU arise. Moreover, the models may differ with respect to when the narrowing down, the underlying optimization, and/or the revelation of true data take place. In the talk, we discuss two example problem settings - one with known and one with unknown hedging points - in more detail, where we handle the remaining uncertainty by the paradigm of robust optimization. For both cases, we give conditions under which a single-level reformulation is possible. Throughout the talk, we use shortest-path problems as underlying optimization problem for illustrating specifics of and phenomena arising in OCU.

Clustering-based scenario reduction for distributionally robust optimization with approximation guarantees – KEVIN MARTIN AIGNER, FRIEDRICH ALEXANDER UNIVERSITY ERLANGEN NÜRNBERG – Sebastian Denzler, Frauke Liers, Sebastian Pokutta, Kartikey Sharma

Stochastic and (distributionally) robust optimization problems often become computationally challenging as the number of scenarios increases. Therefore, scenario reduction is a key technique for reducing the problem size of uncertain optimization programs. We introduce a dimension reduction method for distributionally robust optimization based on the clustering of the scenario set. We prove quantifiable approximation bounds by appropriately projecting the original ambiguity set onto the reduced set of scenarios. Our methodology applies to both continuous and discrete random variables. Numerical experiments on mixed-integer benchmark instances yield significant reductions in solution time with a small approximation error.

Distributionally Robust Optimization Approaches for Resilience and Low Carbon Enhancement of Distribution Power Systems – TIANXIAO WANG, NANJING UNIVERSITY – Caihua

Chen, Houcai Shen

Resilience and low carbon considerations are key factors in assessing the performance of distribution power systems. In this study, we investigate the integration of microgrids into the main grid, which offer the ability to effectively handle power-disrupting incidents through islanding capabilities and facilitate the integration of renewable energy sources. Our focus lies in defining resilience and proposing a multi-stage distributionally robust optimization model that encompasses the stages of preparation, response, and restoration. Additionally, we extend our model to include the cost of carbon emissions as part of the objective. To address the problem at hand, we formulate the model as a mixed-integer linear programming problem, ensuring the adherence of chance constraints to the worst-case distribution within a novel ambiguity set, which combines the Wasserstein distance and first-order moment information. Numerical experiments on the IEEE 34-bus and 118-bus test systems demonstrate the significant improvements achieved by the incorporation of microgrids.

Streamlining Emergency Response: A K-Adaptable model and a Column-and-Constraint-Generation Algorithm – FABRICIO OLIVEIRA, AALTO UNIVERSITY – Paula Weller

Emergency response refers to the systematic response to an unexpected, disruptive occurrence such as a natural disaster. The response aims to mitigate the consequences of the occurrence by providing the affected region with the necessary supplies. A critical factor for a successful response is its timely execution, but the unpredictable nature of disasters often prevents quick reactionary measures. Preallocating the supplies before the disaster takes place allows for a faster response, but requires more overall resources because the time and place of the disaster are not yet known. This gives rise to a trade-off between how quickly a response plan is executed and how precisely it targets the affected areas. Aiming to capture the dynamics of this trade-off, we develop a K-adjustable robust model, which allows a maximum of K second-stage decisions, i.e., response plans. This mitigates tractability issues and allows the decision-maker to seamlessly navigate the gap between the readiness of a proactive yet rigid response and the accuracy of a reactive yet highly adjustable one. The approaches we consider to solve the K-adaptable model are threefold: Approximately, via a partition-and-bound method, and optimally via a branch-and-bound method as well as a static robust reformulation in combination with a column-and-constraint generation algorithm. In a computational study, we compare and contrast the different solution approaches and assess their potential.

3c - ThA174 - 511D - Innovative RO Techniques in Transportation and Location Science Chair:
IHSAN YANIKOGLU 

Tractable smart predict and optimize in the non-realizable case – OMAR BENNOUNA, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Jiawei Zhang, Asuman Ozdaglar, Saurabh Amin

We study the predict-then-optimize framework approach, a common application of machine learning, which entails estimating unknown parameters of a linear optimization problem and then solving the optimization task with these estimations.

For example, consider an energy allocation problem when the energy cost in different areas is uncertain. Despite the absence of precise energy cost values at the time of problem-solving, machine learning models are employed to predict these costs, and the resulting optimization problem, which consists for example of minimizing energy costs while meeting some minimal requirements, is solved using state-of-the-art optimization algorithms. We focus on the unrealizable setting, i.e. when the used hypothesis set (predictor function class) does not contain the ground truth value. In this case, there is no known algorithm that solves this predict-then-optimize problem. We provide a tractable algorithm which successfully finds an optimal solution in this setting.

Robust Lane Covering Problem – IHSAN YANIKOGLU, ASSOC. PROF. –

We study a logistics network where shippers collaborate and bundle their shipment requests in order to negotiate better rates with a common carrier. In this setting, shippers are able to identify collaborative routes with decreased overall empty truck movements. After the optimal routes that minimize total cost of covering all the shippers' demand are determined, this cost is allocated among the shippers. Our goal is to devise cost allocation mechanisms that ensure the sustainability of the collaboration. We first develop cost allocation mechanisms with well-known properties from the cooperative game theory literature, such as budget balance, stability and cross monotonicity. Next, we define a set of new properties, such as a guaranteed discount from the stand alone cost for each shipper, desirable in our setting and propose several cost allocation schemes that could lead to implementable solutions. We also perform a computational study on randomly generated and real-life data to derive insights on the performance of the developed allocation schemes.

Distributionally Robust Optimization with Multimodal Decision-Dependent Ambiguity Sets – BESTE BASCIFTCI, UNIVERSITY OF IOWA – Xian Yu

We consider a two-stage distributionally robust optimization (DRO) model with multimodal uncertainty, where both the mode probabilities and uncertainty distributions could be affected by the first-stage decisions. To address this setting, we propose a generic framework by introducing a ϕ -divergence based ambiguity set to characterize the decision-dependent mode probabilities and further consider both moment-based and Wasserstein distance-based ambiguity sets to characterize the uncertainty distribution under each mode. We identify two special ϕ -divergence examples (variation distance and χ^2 -distance) and provide specific forms of decision dependence relationships under which we can derive tractable reformulations. Furthermore, we investigate the benefits of considering multimodality in a DRO model compared to a single-modal counterpart through an analytical analysis. We provide a computational study over the facility location problem to illustrate our results, which demonstrate that omission of multimodality and decision-dependent uncertainties within DRO frameworks result in inadequately performing solutions with worse in-sample and out-of-sample performances under various settings.

Electric Bus Fleet Scheduling under Travel Time and Energy Consumption Uncertainty – FARZAD AVISHAN, HEC MONTRÉAL –

The public transportation system is experiencing a substantial shift due to the rapid expansion of electromobility infrastructure and operations. This transformation is anticipated to contribute to decarbonizing and promoting environmental sustainability significantly. Among the most pressing planning issues in this area is the optimization of operational and strategic costs associated with electric fleets, which has recently garnered the attention of researchers. This paper investigates the scheduling and procurement problem of electric fleets under travel time and energy consumption uncertainty. A novel mixed-integer linear programming model is proposed, which determines the number of buses required to cover all trips, yields the schedule of the trips, and creates bus charging plans. The robust optimization paradigm is employed to address uncertainty, and a new budget uncertainty set is introduced to control the robustness of the solution. The efficiency of the model is evaluated through an extensive Monte Carlo simulation. Additionally, a case study is conducted on the off-campus college transport network at Binghamton University to demonstrate the real-world applicability of the model. The numerical results have shown that ignoring uncertainty can lead to schedules where up to 48% of the trips are affected, which are either delayed or missed. The proposed approach can also be applied to other transportation networks with similar characteristics and uncertainties.

3c - ThA277 - 511E - **Optimization - Machine Learning Interfaces** Chair: CHAITHANYA BANDI ←

Equivalent reformulation of multistage robust optimization under time-varying exogenous uncertainty with an application to optimal bidding in energy markets – MIRAI TANAKA, THE INSTITUTE OF STATISTICAL MATHEMATICS – Kenji Ono, Takufumi Yoshida, Akiko Takeda

We consider multistage robust linear optimization problems with time-varying exogenous ellipsoidal uncertainty sets. We show the problems can be equivalently reformulated into single-level second-order cone optimization problems. We also show a case study on optimal bidding to allocate electricity resources to day-ahead and real-time energy markets.

General Polyhedral Approximation of Two-Stage Robust Linear Programming – TIM NIE-MANN, TU BRAUNSCHWEIG – Lukas Grunau, Sebastian Stiller

We consider two-stage robust linear programs with uncertain righthand side. We develop a General Polyhedral Approximation (GPA), in which the uncertainty set U is substituted by a finite set of polytopes derived from the vertex set of an arbitrary polytope that dominates U . We analyse and computationally test the performance of GPA for the budgeted uncertainty set U (with m rows). El Housni et al. showed that for budgeted uncertainty affine policies are best possible approximations. In practice calculating affine policies typically requires prohibitive running times. Therefore an approximation of U by a single simplex has been proposed by Ben-Tal et al.. GPA maintains the low practical running times of the simplex based approach while improving the quality of approximation by a constant factor. The generality of our method allows to use any polytope dominating U (including the simplex). We

provide a family of polytopes that allows for a trade-off between running time and approximation factor. The previous simplex based approach reaches a threshold at the square root of m after which it is not better than a quasi nominal solution. Before this threshold, GPA significantly improves the approximation factor. After the threshold, it is the first fast method to outperform the quasi nominal solution.

Multi-purchase behavior: Modeling, estimation, and optimization – DEEKSHA SINHA, META

We study the problem of modeling purchase of multiple products and utilizing it to display optimized recommendations for online retailers and e-commerce platforms. Rich modeling of users and fast computation of optimal products to display given these models can lead to significantly higher revenues and simultaneously enhance the user experience. We present a parsimonious multi-purchase family of choice models called the BundleMVL-K family, and develop a binary search based iterative strategy that efficiently computes optimized recommendations for this model. We establish the hardness of computing optimal recommendation sets, and derive several structural properties of the optimal solution that aid in speeding up computation. This is one of the first attempts at operationalizing multi-purchase class of choice models. We show one of the first quantitative links between modeling multiple purchase behavior and revenue gains. The efficacy of our modeling and optimization techniques compared to competing solutions is shown using several real world datasets on multiple metrics such as model fitness, expected revenue gains and run-time reductions. For example, the expected revenue benefit of taking multiple purchases into account is observed to be 5% in relative terms for the Ta Feng and UCI shopping datasets, when compared to the MNL model for instances with 1500 products. Additionally, across 6 real world datasets, the test log-likelihood fits of our models are on average 17% better in relative terms. Our work contributes to the study multi-purchase decisions, analyzing consumer demand and the retailers optimization problem. The simplicity of our models and the iterative nature of our optimization technique allows practitioners meet stringent computational constraints while increasing their revenues in practical recommendation applications at scale, especially in e-commerce platforms and other marketplaces.

A Robust Approach to Classification and Truth-Validation: Incorporating Human and Large Language Model Decision Making – CHAITHANYA BANDI, NUS BUSINESS SCHOOL –

The advent of generative AI and foundational models has opened up promising avenues to enhance the efficiency of many service systems. However, their wider application has been stymied by the persistent presence of model errors and inaccuracies. These errors represent an intrinsic characteristic of current AI models, and not just a transient hurdle. In this talk, we present a Robust Optimization based approach that embraces this reality and seeks to optimize performance within this context. We present a Robust optimization-based algorithmic framework and develop the Fast Algorithm for Classification and Truth-validation (FACT), which aims to address the critical need for rigorous truth-validation in service systems. FACT is designed to leverage the decision-making competencies of both humans and foundational Large Language Models (LLMs) through a dynamic task routing strategy, in-

formed by task complexity, projected cost, and the capacity of the decision-making entity. We present our results from a pilot implementation at a customer service portal of an online education company, which demonstrates significant efficiency benefits. This approach offers a promising trajectory for future research and practical implementations, especially in the context of systems where AI and human decision-making must be effectively integrated.

3c - ThB296 - 511D - **Distributionally Robust Optimization: Theory Meets Applications**
Chair: RUIWEI JIANG 

Two-Stage Distributionally Robust Optimization for Service Region Design in Crowdsourced Delivery – ALIAA ALNAGGAR, TORONTO METROPOLITAN UNIVERSITY – Merve Bodur

We consider a service region planning problem faced by a crowdsourced delivery platform where drivers are commuters who are willing to deviate from their original routes to make a delivery in exchange for a compensation. The availability of drivers and service demand are uncertain, and it is difficult to estimate their probability distributions due to the absence of service history. To mitigate the effects of data ambiguity, we propose a two-stage distributionally robust optimization (DRO) model. The first stage selects which nodes to offer service in, while the second stage matches drivers and orders after uncertainty is realized. We derive an exact reformulation of the DRO model and develop a monolithic approximation based on a convex relaxation of the subproblem. We further strengthen the proposed approximation by a set of valid inequalities inspired by the linearization reformulation technique. The benefit of the proposed approach for service region design in crowdsourced delivery is demonstrated through numerical experiments.

Reliable Facility Location with Wasserstein Ambiguity – KANGLIN LIU, BEIJING JIAOTONG UNIVERSITY – Ruiwei Jiang, Weijun Xie, Zhi Hai Zhang
Facility disruptions can severely impact the reliability and operational cost of many service systems. In this paper, we consider a reliable uncapacitated facility location problem, in which opened facilities may suffer random failure due to, e.g., natural disasters. In reality, past observations of such failure are often limited, making it challenging to estimate the probability distribution. In this paper, we assume that this distribution is ambiguous and it lives in a Wasserstein ball centered at an empirical distribution. Accordingly, we consider a two-stage model that minimizes the fixed location cost plus the expected assignment cost, with regard to the worst-case probability distribution in the Wasserstein ball. This approach implicitly considers many more failure distributions than the empirical one, leading to a “distributionally robust” solution with better out-of-sample performance. Although two-stage distributionally robust models are NP-hard to solve in general, we exploit the problem structure to derive a compact reformulation based on a mixed-integer linear program. In addition, this derivation produces the worst-case failure distribution of the facilities. To solve this model, we design two branch-and-cut algorithms, both of which significantly outperform commercial solvers in computation. Finally, we conduct extensive numerical studies on standard test instances in the literature. The results demon-

strate the effectiveness and potential value of our approach.

Distributionally Robust Optimization of Probability of Exceedance – HOANG NAM NGUYEN, GEORGE WASHINGTON UNIVERSITY –

We study a novel class of fractional distributionally robust optimization, namely P-DRO, which maximizes the worst-case probability of exceeding of a reward-risk ratio. We consider two types of ambiguity sets, moment-based and statistical distance-based and two types of support, uncertain probabilities and continuum of realizations. Under specific conditions of the reward-risk functions, we develop new convex reformulations for each type of ambiguity sets.

Smooth Uncertainty Sets: Dependence of Uncertain Parameters via a Simple Polyhedral Set – NOAM GOLDBERG, BAR ILAN UNIVERSITY – Michael Poss, Shimrit Shtern

We propose a novel polyhedral uncertainty set for robust optimization. This set imposes bounds on the difference of pairs of uncertain parameters. Intuitively, this is because some uncertain parameters may not differ by more than some given absolute amounts. Such bounds may be dictated by the underlying physics of the problem and may be easily expressed by domain experts. Compared with other methods that may model the relationship of uncertain parameters by linear correlation, smooth uncertainty may only require bounds on the difference of uncertain parameters. Although not required to define our sets, we derive probability guarantees that improve as a function of the magnitude of positive correlation between uncertain parameter values. Finally, we develop compact reformulation techniques for specially structured problems and a row generation algorithm.

3c - FA133 - 511D - **Robust Optimization meets other domains** Chair: AHMADREZA MARANDI 

Robust Appointment Scheduling with General Convex Uncertainty Sets – JUDITH BRUGMAN, TILBURG UNIVERSITY – Johan Van Leeuwen, Dick Den Hertog

The efficient management of appointments is crucial in enhancing high overall operational efficiency and service quality in many sectors, with healthcare as leading example. The Appointment Scheduling Problem (ASP) considers a finite number of customers with uncertain service times served consecutively by a single server, and aims at finding the schedule that minimizes the weighted costs of customers’ waiting time and server’s idle and overtime. We address the ASP from a Robust Optimization (RO) perspective, assuming that the uncertain service times fall within a given uncertainty set and considering the worst-case costs over this set. This problem can be formulated as a complex RO problem that is convex in the uncertain parameters, and we use several techniques, such as the Reformulation Perspectification Technique, to solve it. As a result, we propose a robust framework that can handle large problem instances and accommodates general polyhedral and ellipsoidal uncertainty sets. Numerically solving the ASP with uncertain service times is notoriously hard, and previous studies using stochastic programming methods could only deal with small problem instances. Our novel RO method beats this curse of dimensionality, and can obtain optimal robust schedules for over 50 customers and good approximate solutions for

over 100 customers. For specific choices of uncertainty sets, we discuss patterns in these robust schedules, often the well-known dome-shape.

Connections and Reformulations between Robust and Bilevel Optimization – JOHANNES THÜRAUF, UNIVERSITY TRIER – Marc Goerigk, Jan-nis Kurtz, Martin Schmidt

Robust and bilevel optimization share the common feature that they involve a certain multilevel structure. Hence, although they model something rather different when used in practice, they seem to have a similar mathematical structure. In this talk, we analyze the connections between different types of robust problems (strictly robust problems with and without decision-dependence of their uncertainty sets, min-max-regret problems, and two-stage robust problems) as well as of bilevel problems (optimistic problems, pessimistic problems, and robust bilevel problems). It turns out that bilevel optimization seems to be more general in the sense that for most types of robust problems, one can find proper reformulations as bilevel problems but not necessarily the other way around. We hope that these results pave the way for a stronger connection between the two fields, in particular to use both theory and algorithms from one field in the other and vice versa.

LP-based approximations for disjoint bilinear and two-stage adjustable robust optimization. – AYOUB FOUSSOUL, COLUMBIA UNIVERSITY – Omar El Housni, Vineet Goyal

We consider the class of disjoint bilinear programs $\max\{x^T y \mid x \in X, y \in Y\}$ where X and Y are packing polytopes. We present an $O(\log \log m_1 / \log m_1 \cdot \log \log m_2 / \log m_2)$ -approximation algorithm for this problem where m_1 and m_2 are the number of packing constraints in X and Y respectively. In particular, we show that there exists a near-optimal solution (x, y) such that x and y are “near-integral”. We give an LP relaxation of the problem from which we obtain the near-optimal near-integral solution via randomized rounding. As an application of our techniques, we present a tight approximation for the two-stage adjustable robust optimization problem with covering constraints and right-hand side uncertainty where the separation problem is a bilinear optimization problem. In particular, based on the ideas above, we give an LP restriction of the two-stage problem that is an $O(\log n / \log \log n \cdot \log L / \log \log L)$ -approximation where L is the number of constraints in the uncertainty set. This significantly improves over state-of-the-art approximation bounds known for this problem. Furthermore, we show that our LP restriction gives a feasible affine policy for the two-stage robust problem with the same (or better) objective value. As a consequence, affine policies give an $O(\log n / \log \log n \cdot \log L / \log \log L)$ -approximation of the two-stage problem, significantly generalizing the previously known bounds on their performance.

Quadratic Optimization Through the Lens of Adjustable Robust Optimization – AHMADREZA MARANDI, EINDHOVEN UNIVERSITY OF TECHNOLOGY – Abbas Khademi

Quadratic optimization (QO) has been studied extensively in the literature due to its applicability in many practical problems. While practical, it is known that QO problems are generally NP-hard. So, researchers developed many approximation methods to find good solutions. In this talk, we go beyond

the norm and analyze QO problems using robust optimization techniques. To this end, we first show that any QO problem can be reformulated as a disjoint bi-convex QO problem. Then, we provide an equivalent adjustable robust optimization (ARO) reformulation and leverage the methods available in the literature on ARO to approximate this reformulation. More specifically, we show that using a so-called decision rule technique to approximate the ARO reformulation is interpreted as using a linearization-relaxation technique on its bi-convex reformulation problem. Additionally, we design an algorithm that can find a close-to-optimal solution based on our new reformulations. Our numerical results demonstrate the efficiency of our algorithm, particularly for large-sized instances, compared with the off-the-shelf solvers.

3c - FA140 - 511E - **Robust Optimization for Operations Problems** Chair: ZHENYU HU 

Robust topology optimization using the Bernstein approximation – MIGUEL CARRASCO, UNIVERSIDAD DE LOS ANDES – Julio López, Alfredo Canelas

In this work, we present a novel approach to topology optimization for robust truss structure design. Trusses are two- or three-dimensional mechanical structures consisting of a collection of nodes connected by bars, typically made of linear elastic, isotropic, and homogeneous materials. These structures are engineered to withstand external loads, which often include both stochastic and deterministic forces applied to the nodes.

Our method considers both deterministic and random loadings, aiming to minimize compliance (a measure of structural energy) while adhering to constraints on volume and failure probability. Managing failure probability is often computationally demanding, but we address it using the Bernstein approximation, resulting in a model that is a linear conic programming problem. Additionally, we present a more efficient problem reformulation with small semidefinite constraints. To demonstrate the practicality of our approach, we provide solutions to several examples of truss topology optimization.

Optimized Dimensionality Reduction for Moment-based Distributionally Robust Optimization – KAI PAN, THE HONG KONG POLYTECHNIC UNIVERSITY – Shiyi Jiang, Jianqiang Cheng, Zuo Jun Max Shen

Moment-based distributionally robust optimization (DRO) provides an optimization framework to integrate statistical information with traditional optimization approaches. Under this framework, one assumes that the underlying joint distribution of random parameters runs in a distributional ambiguity set constructed by moment information and makes decisions against the worst-case distribution within the set. Although most moment-based DRO problems can be reformulated as semidefinite programming (SDP) problems that can be solved in polynomial time, solving high-dimensional SDPs is still time-consuming. Unlike existing approximation approaches that first reduce the dimensionality of random parameters and then solve the approximated SDPs, we propose an optimized dimensionality reduction (ODR) approach. We first show that the ranks of the matrices in the SDP reformulations are small, by which we are then motivated to inte-

grate the dimensionality reduction of random parameters with the subsequent optimization problems. Such integration enables two outer and one inner approximations of the original problem, all of which are low-dimensional SDPs that can be solved efficiently, providing two lower bounds and one upper bound correspondingly. More importantly, these approximations can theoretically achieve the optimal value of the original high-dimensional SDPs. As these approximations are nonconvex SDPs, we develop modified Alternating Direction Method of Multipliers (ADMM) algorithms to solve them efficiently. We demonstrate the effectiveness of our proposed ODR approach and algorithm in solving multiproduct newsvendor and production-transportation problems. Numerical results show significant advantages of our approach on the computational time and solution quality over the three best possible benchmark approaches. Our approach can obtain an optimal or near-optimal (mostly within 0.1%) solution and reduce the computational time by up to three orders of magnitude.

Robust Workforce Management with Crowdsourced Delivery – CHUN CHENG, DALIAN UNIVERSITY OF TECHNOLOGY – Melvyn Sim, Yue Zhao

We investigate how crowdsourced delivery platforms with both contracted and ad-hoc couriers can effectively manage their workforce to meet delivery demands amidst uncertainties. Our objective is to minimize the hiring costs of contracted couriers and the crowdsourcing costs of ad-hoc couriers while considering the uncertain availability and behavior of the latter. Due to the complication of calibrating these uncertainties through data-driven approaches, we instead introduce a basic reduced information model to estimate the upper bound of the crowdsourcing cost and a generalized reduced information model to obtain a tighter bound. Subsequently, we formulate a robust satisficing model associated with the generalized reduced information model and show that a binary search algorithm can tackle the model exactly by solving a modest number of convex optimization problems. Our numerical tests using Solomon's data sets show that reduced information models provide decent approximations for practical delivery scenarios. Simulation tests further demonstrate that the robust satisficing model has better out-of-sample performance than the empirical optimization model that minimizes the total cost under historical scenarios. Our framework provides a practical tool for decision-making that enables platforms to effectively manage their workforce resources and balance their cost objectives with service quality requirements.

Robust Capacity Planning with General Upgrading – ZHENYU HU, NATIONAL UNIVERSITY OF SINGAPORE – Zhaowei Hao, Long He, Jun Jiang

General upgrading is a strategy by which a firm can upgrade a customer with any higher-end product whenever a lower-end product is out of stock. In this paper, we consider the capacity planning problem to decide the initial capacity for multiple products to maximize the expected total profit when general upgrading is allowed. We formulate the problem as a two-stage distributionally robust optimization (DRO) model under the commonly employed ambiguity set with marginal mean and variance information. To obtain an exact reformulation as a second-order cone program (SOCP) that is directly solvable, one needs to characterize the extreme points of the dual of the second-stage problem. To this end, we first show

that the dual of the second-stage problem of general upgrading can be equivalently reformulated as an economic lot-sizing problem with bounded inventory constraints. We then derive a binary extended formulation for the extreme points of the dual polyhedron based on the characterization via a shortest path network, which enables a polynomially solvable SOCP reformulation. Our characterization of the extreme points can also be used when partial correlation information is incorporated, and in this case, it leads to a tractable semidefinite program. Our reformulation of the second-stage problem connects various problems studied separately in the literature, such as appointment scheduling and economic lot-sizing problems. Our extensive numerical studies show that our DRO solution performs best with limited training data or in highly uncertain environments with non-stationary demand distributions. Using a real data set from a cosmetic company, we demonstrate that our DRO model yields a higher mean profit and lower variability than the sample average approximation.

3c - FB104 - 511D - Data-driven Robust Optimization Chair: JUSTIN DUMOCHELLE

Conformal Inverse Optimization – BO LIN, UNIVERSITY OF TORONTO – Erick Delage, Timothy Chan

Inverse optimization has been increasingly used to estimate unknown parameters in an optimization model based on decision data. We show that such a point estimation is insufficient in a prescriptive setting where the estimated parameters are used to prescribe new decisions. The prescribed decisions may be low-quality and misaligned with human intuition and thus are unlikely to be adopted. To tackle this challenge, we propose conformal inverse optimization, which seeks to learn an uncertainty set for the unknown parameters and then solve a robust optimization model to prescribe new decisions. Under mild assumptions, we show that the suggested decisions can achieve bounded out-of-sample optimality gaps, as evaluated using both the ground-truth parameters and the decision maker's perception of the unknown parameters. Our method demonstrates strong empirical performance compared to classic inverse optimization.

Learning Decision-Focused Uncertainty Sets in Robust Optimization – IRINA WANG, PRINCETON UNIVERSITY –

We propose a data-driven technique to automatically learn the uncertainty sets in robust optimization. Our method reshapes the uncertainty sets by minimizing the expected performance across a family of problems subject to guaranteeing constraint satisfaction. Our approach is very flexible and can learn a wide variety of uncertainty sets while preserving tractability. We solve the constrained learning problem using a stochastic augmented Lagrangian method that relies on differentiating the solutions of the robust optimization problems with respect to the parameters of the uncertainty set. Due to the nonsmooth and nonconvex nature of the augmented Lagrangian function, we apply the nonsmooth conservative implicit function theorem to establish convergence to a critical point, which is a feasible solution of the constrained problem under mild assumptions. Using empirical process theory, we show finite-sample probabilistic guarantees of constraint satisfaction for the resulting solutions. Numerical experiments show that our method out-

performs traditional approaches in robust and distributionally robust optimization in terms of out-of-sample performance and constraint satisfaction guarantees.

Neur2RO: Neural Two-Stage Robust Optimization – JUSTIN DUMOCHELLE, UNIVERSITY OF TORONTO –

Robust optimization provides a mathematical framework for modeling and solving decision-making problems under worst-case uncertainty. This work addresses two-stage robust optimization (2RO) problems (also called adjustable robust optimization), wherein first-stage and second-stage decisions are made before and after uncertainty is realized, respectively. This results in a nested min-max-min optimization problem which is extremely challenging computationally, especially when the decisions are discrete. We propose Neur2RO, an efficient machine learning-driven instantiation of column-and-constraint generation (CCG), a classical iterative algorithm for 2RO. Specifically, we learn to estimate the value function of the second-stage problem via a novel neural network architecture that is easy to optimize over by design. Embedding our neural network into CCG yields high-quality solutions quickly as evidenced by experiments on two 2RO benchmarks, knapsack and capital budgeting. For knapsack, Neur2RO finds solutions that are within roughly 2% of the best-known values in a few seconds compared to the three hours of the state-of-the-art exact branch-and-price algorithm; for larger and more complex instances, Neur2RO finds even better solutions. For capital budgeting, Neur2RO outperforms three variants of the k-adaptability algorithm, particularly on the largest instances, with a 5 to 10-fold reduction in solution time.

3c - FB122 - 511E - Optimization with Marginals and Moments Chair: KARTHIK NATARAJAN

Distributionally robust optimization through the lens of submodularity – KARTHIK NATARAJAN, SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN – Divya Padmanabhan, Arjun Ramachandra

Distributionally robust optimization is used to solve decision making problems under adversarial uncertainty where the distribution of the uncertainty is itself ambiguous. In this paper, we identify a class of these instances that is solvable in polynomial time by viewing it through the lens of submodularity. We discuss connections to the multimarginal optimal transport problem and the generalized moment problem by bridging ideas from convexity in continuous optimization to submodularity in discrete optimization.

When submodularity meets pairwise independence – ARJUN RAMACHANDRA, IIM BANGALORE – Karthik Natarajan

Submodularity provides a natural context to model a large class of discrete optimization problems including but not limited to influence maximization, mechanism design, resource allocation and several machine learning problems. As a set functional property, submodularity models the notion of diminishing returns in the discrete space. Theoretically, it has intrigued scientists due to strong structural similarity with both convex and concave functions in the continuous space, which has been exploited to derive approximation guarantees for deter-

ministic submodular optimization problems, using continuous extensions. Our work, however, approaches submodular optimization from the lens of distributional robustness which seeks to evaluate or approximate the worst-case expected value of a submodular set function (subjected to random inputs) over a set of joint distributions satisfying some assumptions. Even with univariate information (marginal probabilities of the random inputs) alone, evaluating this optimal expected value is known to be an NP-complete problem. Existing approaches tackle the hardness by approximating the optimal expected value, assuming the random inputs to be independent. This notion is formalized by the concept of correlation gap which quantifies how much we “lose” in the expectation of the function by ignoring the correlation structure of the random set and assuming independence instead. For monotone submodular set functions, it was shown that the correlation gap is upper bounded by $e/(e-1)$ in Agrawal et.al. (2012). In reality, however, more complex notions of randomness are often encountered, such as when weak correlations coexist with higher-order dependencies. Inspired by the need to incorporate more realistic notions of randomness and driven by the curiosity to understand the interplay between functional properties and randomness, we study the behaviour of monotone submodular set functions with pairwise independent random input. We show that in this setting, the $e/(e-1)$ bound on the correlation gap can be improved to $4/3$ (and that it is tight) in the following cases: (a) for small size of random inputs with general marginal probabilities (b) for general size of random inputs with small marginal probabilities and (c) for a specific submodular function (whose expectation is the probability that the chosen input set is non-empty) for general size of random inputs with general marginal probabilities. For rank functions of k-uniform matroids, we show that the ratio can be further improved to $4k/(4k-1)$ for general size of random inputs with identical probabilities. Applications in distributionally robust optimization and mechanism design are demonstrated. Our results illustrate a fundamental difference in the behavior of submodular functions under weaker notions of independence with potential ramifications in improving existing algorithmic approximations.

Convex Optimization for Bundle Size Pricing Problem – HAILONG SUN, ANTAI COLLEGE OF ECONOMICS AND MANAGEMENT, SHANGHAI JIAO TONG UNIVERSITY – Li Xiaobo, Chung Piaw Teo

We study the bundle size pricing (BSP) problem in which a monopolist sells bundles of products to customers and the price of each bundle depends only on the size (number of items) of the bundle. Although this pricing mechanism is attractive in practice, finding optimal bundle prices is difficult because it involves characterizing distributions of the maximum partial sums of order statistics. In this paper, we propose to solve the BSP problem under the cross moment model (CMM), which is constructed using only the first and second moments of customer valuations. Correlations between valuations of bundles are captured by the covariance matrix. We show that the BSP problem under this model is convex and can be efficiently solved using off-the-shelf solvers. Our approach is flexible in optimizing prices for any given bundle size. Numerical results show that it performs very well compared with state-of-the-art heuristics. This provides a unified and efficient approach to solve the BSP problem under various distributions and di-

mensions. We also provide a few insights regarding the BSP problem and CMM. First, the BSP problem can be converted into a multichoice pricing problem with correlations of valuations, and it is crucial to capture such correlations to construct good approximate solutions. Second, using only moment information in the way of CMM can be reasonable for constructing a good approximation to the BSP problem.

Progresses in Modeling and solving distributional robust optimization problems – HE SIMAI, SHANGHAI JIAO TONG UNIVERSITY –

Distributional robust optimization(DRO) focusing on making efficient and robust decision with limited data or information. The two key questions are, how to formulate the uncertainty set of the unknown distribution, and how to solve the corresponding model efficiently. In this talk, we present our recent progresses on how to utilize limited information efficiently to build sharper uncertainty set to achieve more robust and accurate decisions, and the corresponding solution approach.

Firstly, for single dimensional moment based DRO models, we establish the value of distribution shape information by showing its influence on model accuracy, identify the optimal solution structure by reverse convex optimization theory, and propose numerical approaches for solving such models. Secondly, for heavy tail or light tail distributions, general moments (expectation of general function of the random variable) like entropy function, moment generation function, log-function, fractional power function, etc, are widely used in other research areas like statistics. We construct a general approach for solving single dimensional DRO problems with general moments, by reducing the primal-dual complementary-slackness condition into determinant of matrices similar to Vandermonde matrices, which is much easier to analyze and solve than the original condition. We further illustrate how to utilize such an approach to gain analytical/semi analytical solutions to various problems. Thirdly, we report our recent progresses on extending our works to high dimensional scenarios. Lastly, we show correlation information can greatly enhance the model accuracy of data driven DRO models using Wasserstein distance, and ease the curse-of-dimensionality of such models.

3c - FB184 - 511A - Exact methods for robust nonlinear optimization Chair: DICK DEN HERTOOG 

RPT for robust nonlinear optimization problems – DANIQUE DE MOOR, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Dick Den Hertog, Thororis Koukouvinos, Dimitris Bertsimas, Jianzhe Zhen In this talk, we address robust optimization problems involving robust nonlinear constraints, that are defined as sums of products of linear times concave functions with respect to the uncertain parameters. We introduce the Reformulation-Perspectification Technique, which allows us to obtain a convex relaxation of such problems and subsequently solve them to optimality. We provide multiple examples of robust optimization problems falling in that class. Further, we extend the applicability of our approach to two-stage adaptive robust optimization problems.

Reformulation-Perspectification Technique

for problems with conic uncertainty sets – DICK DEN HERTOOG, UNIVERSITY OF AMSTERDAM – Dimitris Bertsimas, Danique De Moor, Thororis Koukouvinos, Jianzhe Zhen

In this talk, we study robust nonlinear optimization problems involving sum of linear times convex (SLC) functions as well as uncertainty sets that are defined by conic constraints belonging to one of the five basic cones, that is, linear cone, second order cone, power cone, exponential cone, and semidefinite cone. By using the Reformulation Perspectification Technique, we can obtain a convex relaxation by forming the perspective of each function and linearizing all product terms of uncertain parameters with newly introduced uncertain parameters. To further tighten the approximation, we can pairwise multiply the conic constraints. In this paper, we analyze all possibilities of multiplying conic constraints. Moreover, in case of an exponential cone we generate valid inequalities that can be used to further strengthen the approximation and in case of a power cone we generate additional valid inequalities.

An exact method for a class of robust nonlinear optimization problems – THORORIS KOUKOUVINOS, OPERATIONS RESEARCH CENTER, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, UNITED STATES – Dimitris Bertsimas, Danique De Moor, Dick Den Hertog, Jianzhe Zhen

We introduce a novel exact approach for solving optimization problems with robust nonlinear constraints, that are sums of products of linear times concave (SLC) functions with respect to the uncertain parameters. Our approach synergizes a cutting set method with reformulation-perspectification techniques and branch and bound. Numerical experiments on a robust convex geometric optimization problem and a robust linear optimization problem with data uncertainty and implementation error show that our approach can solve robust nonlinear problems that cannot be solved by existing methods in the literature. Moreover, a numerical experiment on a lot-sizing problem on a network demonstrates the efficiency of our method for two-stage ARO problems.

3d - MB167 - 513A - Reinforcement Learning in Industry Chair: DANIEL JIANG 

Adaptive Experimentation at Scale: A Computational Framework – ETHAN CHE, COLUMBIA BUSINESS SCHOOL – Daniel Jiang, Hongseok Namkoong, Jimmy Wang

Experimentation is the basis of scientific decision-making. Adaptive experimentation (e.g., bandits) can provide substantial statistical power, but existing algorithms are designed by theoreticians to achieve analytic guarantees and notoriously difficult to implement and maintain in practice. We propose a new paradigm of adaptive experimentation centered around computational tools. Motivated by problem instances involving large batches, delayed feedback, and a small number of opportunities to reallocate sampling effort, we design near-optimal, easy-to-deploy adaptive experiments algorithms that can flexibly deal with any batch size. Our main observation is that normal approximations universal in statistics can guide the design of scalable adaptive experimentation methods. Our preliminary results show large performance gains against state-of-the-art methods (e.g., Thompson sampling).

Exploration Incentives in Model-Based Reinforcement Learning – ALEC KOPPEL, AI RESEARCH, JP MORGAN CHASE –

Reinforcement Learning (RL) is a form of stochastic adaptive control in which one seeks to estimate parameters of a controller only from data, and has gained popularity in recent years. However, technological applications of RL are often hindered astronomical sample complexity demanded by their training. Model-based reinforcement learning is known to provide a practically sample efficient approach; however, its performance certificates in terms of Bayesian regret often require restrictive Gaussian assumptions, and may fail to distinguish between vastly different performance in sparse or dense reward settings. Motivated by these gaps, we propose a way to make MBRL, namely, Posterior Sampling combined with Model-Predictive Control (MPC), computationally efficient for mixture distributions based a novel application of integral probability metrics and kernelized Stein discrepancy. Then, we build upon this insight to pose a new exploration incentive called Stein Information Gain, which permits us to come up with a variant of information-directed sampling (IDS) whose exploration incentive is evaluable in closed-form. Bayesian and information-theoretic regret bounds of the proposed algorithms are presented. Finally, experimental validation on some environments from OpenAI Gym and Deepmind Control Suite illuminates the merits of the proposed methodologies in the sparse-reward setting.

RL and ADP for Rideshare Operations – ZHIWEI (TONY) QIN, LYFT –

Dispatch is at the center of the rideshare operations. In this talk, we will first revisit the core developments in reinforcement learning (RL) for rideshare dispatch and the extensions to online learning and joint optimization with repositioning. We will subsequently look at in parallel an approximate DP approach to the same problem and discuss their interesting connections, which offer an alternative lens to interpret the RL approach.

Deep Reinforcement Learning for Recommender Systems and Beyond – ZHEQING ZHU, META AI –

Current recommender systems predominantly employ supervised learning algorithms, which often fail to optimize for long-term user engagement. This short-sighted approach highlights the significance of sequential recommender systems, designed to make decisions that extend beyond immediate user responses. To maximize cumulative positive feedback, these systems must balance exploration—probing users for insightful feedback to inform future recommendations—and the strategic selection of items that pave the way for more successful future interactions.

However, the dynamic nature of user behavior and evolving social trends present additional challenges, demanding that sequential recommender systems operate effectively in non-stationary environments. This necessitates a more discerning exploration strategy, focusing on gathering enduring insights rather than ephemeral information.

In this talk, we introduce several key advancements in sequential recommender systems and related production systems. These contributions are centered around sequential decision making and scalable and intelligent exploration techniques, accommodating both immediate and delayed user feedback, while

adeptly adjusting to non-stationary contexts. We provide empirical evidence demonstrating that our proposed algorithms outperforms existing methods in various benchmarks and real-world systems.

3d - MC783 - 515B - Multi-stage Stochastic Programming and Reinforcement learning 1

Chair: ELIZABETH GLISTA 

Recent Developments in the MPI-SPPY Solver Library for Scalable Stochastic Programming – JEAN PAUL WATSON, LAWRENCE LIVERMORE NATIONAL LABORATORY – David Woodruff, Bernard Knueven

The open source mpi-sppy library is a Python package that is focused on providing scalable decomposition implementations of progressive hedging (PH) and related algorithms for general stochastic programs. Originally built on the Pyomo algebraic modeling language (AML), mpi-sppy has been used to solve million-scenario problem instances on HPC platforms, and emphasizes (but is not limited to) multi-stage mixed-integer models. In this talk, we provide an overview of recent capability developments in mpi-sppy, including (1) automatic computation of the PH penalty parameter, (2) warm-starting of PH solves based on solutions to previously solved instances, (3) alternative linearizations of the PH quadratic penalty objective terms, (4) advanced variable bounds-tightening methods, and (5) AML-independent extensions to mpi-sppy (allowing, e.g., AMPL or GAMS base models).

Resampling Stochastic Gradient Descent Cheaply for Efficient Uncertainty Quantification – ZITONG WANG, COLUMBIA UNIVERSITY –

Stochastic gradient descent (SGD) or stochastic approximation has been widely used in model training and stochastic optimization. While there is a huge literature on analyzing its convergence, inference on the obtained solutions from SGD has only been recently studied, yet it is important due to the growing need for uncertainty quantification. We investigate two computationally cheap resampling-based methods to construct confidence intervals for SGD solutions. One uses multiple, but few, SGDs in parallel via resampling with replacement from the data, and another operates this in an online fashion. Our methods can be regarded as enhancements of established bootstrap schemes to substantially reduce the computation effort in terms of resampling requirements while bypassing the intricate mixing conditions in existing batching methods. We achieve these via a recent so-called cheap bootstrap idea and Berry-Esseen-type bound for SGD.

Stochastic Capacity Expansion Planning considering Power Flow and Unit Commitment – ELIZABETH GLISTA, LAWRENCE LIVERMORE NATIONAL LABORATORY (LLNL) –

Planning for the future of the power grid is of growing importance due to a variety of concerns: increasing demand for electric power, a shift toward weather-dependent sources of power generation, and increasing uncertainty due to climate risks. In this setting of high variability and uncertainty, multi-stage stochastic capacity expansion planning (CEP) can be used to inform long-term transmission, generation, and storage investment in a power system. While various stochastic CEP models exist that can co-optimize investments in a network under un-

certainty, many of these models take a simplified version of the power network, ignoring related power flow (PF) and unit commitment (UC) constraints, under the assumption that these short-term constraints are unimportant for long-term planning. In this talk, we present a two-stage CEP model that can include more realistic PF and UC constraints. Using the open-source Electrical Grid Research and Engineering Tools (EGRET) library, we show how EGRET's modular suite of power system tools, including PF and UC models, can be leveraged in a two-stage stochastic CEP model that is then solved via Progressive Hedging, and we discuss extensions to a multi-stage setting. Finally, we demonstrate the performance of our CEP model formulations on a large-scale synthetic-but-realistic California test case with over 8,000 buses and up to a year of representative future days. By comparing CEP results with and without UC and PF representations, we discuss how the solution changes when we consider these additional constraints.

3d - TA53 - 515B - Approximation of Large Scale Structured Markov Decision Processes in Operations Applications Chair: SELVAPRABU NADARAJAH

Combinatorial Fluid Models for Routing – ANDRE AUGUSTO CIRE, UNIVERSITY OF TORONTO – Adam Diamant

We investigate a dynamic scheduling framework to assign and route agents to satisfy geographically dispersed stochastic demand. Our problem is motivated by a home healthcare application where health practitioners (HPs) visit their patients on daily tours. We model the decision of which patients to assign to HPs as a discrete-time Markov decision process, and propose an approximate linear programming framework due to the curse of dimensionality and the complex combinatorial structure associated with an HP's daily travel. Specifically, we propose a one-step policy improvement approach where the basis policies are derived from a fluid approximation of the system, leveraging the tour structure to enhance the approximation quality. We study the asymptotic behavior of our fluid system and the resulting allocation and routing optimization model, which can be reduced to a two-stage stochastic program. Our solution methodology is compared to existing policies in a discrete-event simulation using data from a Canadian home care agency.

Structure Aware Lagrangian Relaxations for Weakly Coupled Markov Decision Processes – SELVAPRABU NADARAJAH, UNIVERSITY OF ILLINOIS CHICAGO – Andre Augusto Cire

Weakly coupled Markov decision processes (WDPs) represent a class of multi-stage decision making models with structure that facilitates computationally tractable Lagrangian methods. WDPs "couple" smaller component Markov decision processes via linking constraints. We investigate the value of new structure aware Lagrangian models that lie between two extremes: standard Lagrangian relaxation, which satisfies linking constraints in expectation, and feasible models that satisfy linking constraints exactly. Our investigation is based on a new combinatorial perspective for finding approximations. This perspective deepens the theoretical understanding of WDP approximations and provides approximation guarantees. We also show that our theory translates to significant improvements in ex-

isting feasible models, and enables new structured Lagrangian models of substantially smaller size.

Unit Commitment without Commitment: A Dynamic Framework for Managing an Integrated Energy System Under Uncertainty – DAVID BROWN, DUKE UNIVERSITY –

Though variability and uncertainty have always posed challenges for power systems, the increasing use of renewable energy sources has exacerbated these issues. At a vertically integrated utility, the system operator manages many generation units – renewable and otherwise – and storage units to ensure that the total energy production matches contemporaneous demand. Current industry practice at these utilities involves solving "unit commitment" and "economic dispatch" optimization problems to choose production plans: these models, while complex, do not explicitly incorporate uncertainty. In this paper, we develop a dynamic framework to help system operators manage production under uncertainty. We formulate the problem as a stochastic dynamic program and use Lagrangian methods to decompose the system across units. The Lagrangian model relaxes the demand-matching constraint and introduces stochastic Lagrange multipliers that can be interpreted as prices representing the varying marginal value of energy production; each unit is then operated to maximize its own expected "profit" given these uncertain prices. These unit-specific value functions are then used to incorporate longer-term effects in dispatch decisions. The unit-specific value functions also provide a way to value generation and storage units in an uncertain environment. We develop relevant theory and demonstrate this dynamic framework using data from the Duke Energy Carolinas and Progress systems. Our numerical experiments demonstrate that this dynamic approach is computationally feasible at an industrial scale and can improve on current practice. Specifically, our results suggest that this dynamic approach can reduce operational costs by about 2% on average in the present Duke Energy system and, in a "future" system with increased solar and storage capacity, can reduce operational costs by 4-5% on average. Perhaps more strikingly, this dynamic approach, on average, performs within 0.2-0.3% of production plans based on perfect foresight about future net demands.

Faster Approximate Dynamic Programming by Freezing Slow States – DANIEL JIANG, META – Yijia Wang

We consider infinite horizon Markov decision processes (MDPs) with fast-slow structure, meaning that certain parts of the state space move "fast" (and in a sense, are more influential) while other parts transition more "slowly." Such structure is common in real-world problems where sequential decisions need to be made at high frequencies, yet information that varies at a slower timescale also influences the optimal policy. Examples include: (1) service allocation for a multi-class queue with (slowly varying) stochastic costs, (2) a restless multi-armed bandit with an environmental state, and (3) energy demand response, where both day-ahead and real-time prices play a role in the firm's revenue. Models that fully capture these problems often result in MDPs with large state spaces and large effective time horizons (due to frequent decisions), rendering them computationally intractable. We propose an approximate dynamic programming algorithmic framework based on the idea of "freezing" the slow states, solving a set of sim-

pler finite-horizon MDPs (the lower-level MDPs), and applying value iteration (VI) to an auxiliary MDP that transitions on a slower timescale (the upper-level MDP). We also extend the technique to a function approximation setting, where a feature-based linear architecture is used. On the theoretical side, we analyze the regret incurred by each variant of our frozen-state approach. Finally, we give empirical evidence that the frozen-state approach generates effective policies using just a fraction of the computational cost, while illustrating that simply omitting slow states from the decision modeling is often not a viable heuristic.

3d - TB22 - 515B - **Stochastic Dual Dynamic Programming : theory and applications**

Chair: JOAQUIM DIAS GARCIA

Duality, upper bounds, and policy quality in Multistage Stochastic Programming – BERNARDO FREITAS PAULO DA COSTA, FUNDAÇÃO GETÚLIO VARGAS – Vincent Leclère

The standard SDDP algorithm produces lower estimates of the value functions at each stage, which in turn induce a policy. In the risk-neutral setting, the cost of such a policy can be estimated by Monte-Carlo, which also yields an upper bound for the optimal value of the problem. In the risk-averse case, due to the variable weighting of scenarios, such elementary estimates are not easily available. A natural alternative is to construct upper bounds for the value functions themselves. This can be performed either with inner approximations or through the dual problem, which both require an estimate of the Lipschitz constant of the value functions. We show that the dual method yields policies with guaranteed upper bounds on the risk-adjusted costs, similarly to policies induced by inner approximations. This contrasts to the policies constructed from lower bounds, which could have poor quality even for a good lower bound. We further show that an appropriate relaxation of the dual method is equivalent to the inner approximations, which provides an interesting link between the upper bound algorithms.

Data-Driven Stochastic Dual Dynamic Programming: Performance Guarantees and Regularization Schemes – GRANI A HANASUSANTO, UNIVERSITY OF ILLINOIS URBANA CHAMPAIGN – Hynk Park, Zhuangzhuang Jia

We propose a data-driven scheme for multistage stochastic linear programming with Markovian random parameters by extending the stochastic dual dynamic programming (SDDP) algorithm. In our data-driven setting, only a finite number of historical trajectories are available. The proposed SDDP scheme evaluates the cost-to-go functions only at the observed sample points, where the conditional expectations are estimated empirically using kernel regression. The scheme thus avoids the construction of scenario trees, which may incur exponential time complexity during the backward induction step. However, if the training data is sparse, the resulting SDDP algorithm exhibits a high optimistic bias that gives rise to poor out-of-sample performances. To mitigate the small sample effects, we adopt ideas from the distributionally robust optimization (DRO), which replaces the empirical conditional expectation in the cost-to-go function with a worst-case conditional expectation over a polyhedral ambiguity set. We derive the theoretical

out-of-sample performance guarantee of the data-driven SDDP scheme and show that the dependence of its sample complexity on the number of time stages is merely polynomial. Finally, we validate the effectiveness of the proposed algorithm and demonstrate its superiority over existing data-driven schemes through extensive numerical experiments.

Multistage stochastic inventory optimization with contracting – ANDY PHILPOTT, THE UNIVERSITY OF AUCKLAND – Dominic Keehan, Andrew Mason

We consider a multistage stochastic production planning problem over a 12 month planning horizon faced by New Zealand's largest dairy products manufacturer. The sales price for each product follows a random process. In each month, the company decides a mix of products to make, what products to sell immediately, and what products to store for later sales. In the absence of contracting we present a separation result that can be used to determine a production plan independent of a sales/inventory policy. When the company arranges forward sales contracts, the optimal production and sales policies become coupled. We present some numerical results for policies computed using SDDP and compare them with policies derived from model predictive control.

A Multicut Approach to Compute Upper Bounds for Risk-Averse SDDP – JOAQUIM DIAS GARCIA, PSR –

Stochastic Dual Dynamic Programming (SDDP) is a widely used and fundamental algorithm for solving multistage stochastic optimization problems. Although SDDP has been frequently applied to solve risk-averse models with the Conditional Value-at-Risk (CVaR), it is known that the estimation of upper bounds is a methodological challenge, and many methods are computationally intensive. In practice, this leaves most SDDP implementations without a practical and clear stopping criterion. In this paper, we propose using the information already contained in a multicut formulation of SDDP to solve this problem with a simple and computationally efficient methodology.

The multicut version of SDDP, in contrast with the typical average cut, preserves the information about which scenarios give rise to the worst costs, thus contributing to the CVaR value. We use this fact to modify the standard sampling method on the forward step so the average of multiple paths approximates the nested CVaR cost. We highlight that minimal changes are required in the SDDP algorithm and there is no additional computational burden for a fixed number of iterations.

We present multiple case studies to empirically demonstrate the effectiveness of the method. First, we use a small hydrothermal dispatch test case, in which we can write the deterministic equivalent of the entire scenario tree to show that the method perfectly computes the correct objective values. Then, we present results using a standard approximation of the Brazilian operation problem and a real hydrothermal dispatch case based on data from Colombia. Our numerical experiments showed that this method consistently calculates upper bounds higher than lower bounds for those risk-averse problems and that lower bounds are improved thanks to the better exploration of the scenarios tree.

3d - TC940 - 515B - **Multi-stage Stochastic Programming and Reinforcement learning 1**
Chair: GIULIA PEDRIELLI

A Low-Rank ADMM Splitting Approach for Semidefinite Programming – QIUSHI HAN, UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN – Chenxi Li, Zhenwei Lin, Caihua Chen, Qi Deng, Dongdong Ge, Huikang Liu, Yinyu Ye

We introduce a new first-order method for solving general semidefinite programming problems, based on the alternating direction method of multipliers (ADMM) and a matrix-splitting technique. Our algorithm has an advantage over the Burer-Monteiro approach as it only involves much easier quadratically regularized subproblems in each iteration. For a linear objective, the subproblems are well-conditioned quadratic programs that can be efficiently solved by the standard conjugate gradient method. We show that the ADMM algorithm achieves sublinear or linear convergence rates to the KKT solutions under different conditions. Building on this theoretical development, we present LoRADS, a new solver for linear SDP based on the Low-Rank ADMM Splitting approach. LoRADS incorporates several strategies that significantly increase its efficiency. Firstly, it initiates with a warm-start phase that uses the Burer-Monteiro approach. Moreover, motivated by the SDP low-rank theory [So et al. 2008], LoRADS chooses an initial rank of logarithmic order and then employs a dynamic approach to increase the rank. Numerical experiments indicate that LoRADS exhibits promising performance on various SDP problems. A noteworthy achievement of LoRADS is its successful solving of a matrix completion problem with 15,694,167 constraints and a matrix variable of size $40,000 \times 40,000$ in 351 seconds.

Scaling Bayesian optimization with Multi-agent Rollout – GIULIA PEDRIELLI, ARIZONA STATE UNIVERSITY –

Solving black-box global optimization problems efficiently across domains remains challenging especially for large scale optimization problems. Bayesian optimization has obtained important success as a black box optimization technique based on surrogates, but it still suffers when applied to large scale heterogeneous landscapes. Recent approaches have proposed non-myopic approximations and partitioning of the input domain into subregions to prioritize regions that capture important areas of the solution space. We propose a Multi Agent Rollout formulation of Bayesian optimization (MAroBO) that partitions the input domain among finite set of agents for distributed sampling. In addition to selecting candidate samples from their respective sub regions, these agents also influence each other in partitioning the sub regions. Consequently, a portion of the function is optimized by these agents which prioritize candidate samples that don't undermine exploration in favor of a single step greedy exploitation. The efficacy of MAroBO is demonstrated on synthetic test functions.

3d - WA148 - 515B - **Advances in Approximate Dynamic Programming with Energy Applications** Chair: STEPHAN MEISEL

Network dual reoptimization policies and

bounds for managing energy real options – ALESSIO TRIVELLA, UNIVERSITY OF TWENTE – Selvaprabu Nadarajah

Energy operations and investments are managed with real option models that embed timing and/or switching decisions, where the former and latter types of decisions change irreversibly and reversibly, respectively, the status of the real asset. Optimizing the flexibility in these models requires solving a finite-horizon Markov decision process with finite endogenous states but a high-dimensional and continuous exogenous state space, which is typically intractable and requires heuristic approaches. Least squares Monte Carlo (LSM) is a popular approximate dynamic programming method for this purpose while a known forecast-based reoptimization heuristic (FRH) is not well suited to handle timing decisions. We develop a network dual reoptimization framework that leverages network flow algorithms and employs a novel class of partial information relaxations to overcome this shortcoming. Our framework provides both a new policy and a dual bound and we establish that: (i) the decisions taken by our policy are equivalent to solving a scenario-based two-stage stochastic program but computationally more efficient since scenarios are solved independently; and (ii) our dual bound is tighter than the known penalized perfect information bound. Numerical experiments on two energy real options dealing with commodity production and vehicle fleet upgrade show that our policy outperforms FRH by 10-20% and LSM by to 2-5%, while our dual bound improves by less than 1%.

Reinforcement Learning for Energy-Efficient Job Shop Scheduling with Large Structured Action Spaces – WOUTER VAN HEESWIJK, UNIVERSITEIT TWENTE –

Over the past years, reinforcement learning has gained increasing traction in the operations research community, with novel solution methods proposed in computer science finding their way to solve sequential planning and scheduling problems. Although many methods have been proposed to handle large state spaces, combinatorial optimization problems are also well-known for their extremely large decision spaces. Despite their size, such decision spaces often display spatiotemporal structures and desirable properties such as convexity, which are often exploited through mathematical programming. This research proposes a novel algorithmic pipeline, in which local decision neighborhoods are generated and formulated as mathematical programs. These neighborhoods are subsequently explored based on Q-values, utilizing both mathematical programming and simulated annealing to identify improving actions. The proposed pipeline is applied on a stylized variant of a job shop scheduling problem, in which energy consumption is required to be balanced. Performance is shown to match or outperform state-of-the-art embedding methods and scales beyond enumerable decision spaces.

Stochastic Approximation for Upper Bound Optimization in Markov Decision Processes – NEGAR SOHEILI, UNIVERSITY OF ILLINOIS AT CHICAGO – Selvaprabu Nadarajah

Markov decision processes (MDPs) with small endogenous states are prevalent in options pricing, investment planning, and operations. Solving these problems is difficult because of the presence of a high-dimensional exogenous state driven by a stochastic process. Computing an upper bound of the

optimal policy of such MDPs is important for assessing the quality of heuristic policies. Approximate linear programming (ALP) and pathwise optimization (PO) are two established approaches that minimize an upper bound on the optimal policy value. However, they often lead to challenging semi-infinite linear programs. PO is known to provide better upper bounds than ALP. We propose a new ALP relaxation formulation based on surrogate relaxation using exogenous state information. The upper bound computed by this model is better than that of ALP and converges to the MDP optimal policy value under appropriately sampled basis functions. In addition, we show a relationship between the upper bounds of PO and our ALP relaxation and also compare these two models more broadly. We develop first-order methods to solve our ALP relaxation with convergence guarantees and show that optimized upper bounds for MDPs with small endogenous states can be computed more efficiently than using ALP.

Reinforcement Learning or Policy Sampling: An Online Coupling Policy with an Energy Storage Application – STEPHAN MEISEL, UNIVERSITY OF TWENTE –

We propose a new approach for solving episodic sequential decision problems that occur under unknown uncertainty. The approach leverages concurrent real-time simulation to couple online reinforcement learning with a multi-armed bandit type of online learning approach that we introduce under the name policy sampling. We show the theoretical performance guarantees of the proposed coupling policy, and illustrate the new method by applying it to an energy storage management problem.

3d - WB176 - 515B - **Advances in Stochastic Programming** Chair: GUANGHUI LAN ↔

Optimal L-p Risk Minimization – GUANGHUI LAN, GEORGIA INSTITUTE OF TECHNOLOGY – Zhichao Jia, Jimmy Zhang

Convex risk measures play a foundational role in the area of stochastic optimization. However, in contrast to risk neutral models, their applications are still limited due to the lack of efficient solution methods. In particular, the mean L-p semi-deviation is a classic risk minimization model, but its solution is highly challenging due to the composition of concave-convex functions and the lack of uniform Lipschitz continuity. In this talk, we discuss some progresses on the design of efficient algorithms for l-p risk minimization, including a novel lifting reformulation to handle the concave-convex composition, and a new stochastic approximation method to address the non-Lipschitz continuity. We establish an upper bound on the sample complexity associated with this approach and show that this bound is not improvable for L-p risk minimization in general.

Epi-Consistency of Multistage Stochastic Optimization Problems – DOMINIC KEEHAN, UNIVERSITY OF AUCKLAND – Johannes Royset

We study the statistical consistency of sample average approximations of infinite-horizon multistage stochastic optimization problems. Existing results establish the consistency of minima using arguments which rely on the stage cost functions being bounded. When this does not hold, a different approach is required. To this end we develop an existence and consistency result for the approximation of fixed-point problems in met-

ric spaces which can be applied to infinite-horizon problems. Utilising the Attouch-Wets distance on a suitable space of approximating functions and results on the epi-convergence of expectation functions with varying measures and integrands, we show how epi-consistency can be assured when the stage cost functions are unbounded. The arguments are also applied to finite-horizon multistage stochastic optimization problems. A number of examples are considered.

Three-stage Stochastic Programming is As Easy As Classical Stochastic Optimization – YIFAN HU, EPFL – Jie Wang, Xie Yao, Andreas Krause, Daniel Kuhn

Multistage stochastic programming poses significant challenges due to the curse of dimensionality. In this talk, we study three-stage problems, often perceived as more challenging than the classical stochastic optimization. Remarkably, we develop a randomized gradient-based algorithm that achieves the same complexity bound as classical stochastic optimization. This unexpected outcome challenges conventional wisdom regarding the difficulty of three-stage problems, suggesting that prevailing beliefs may not always hold true. The development of such a randomized gradient-based algorithm opens up exciting directions for advancing algorithmic design in multistage stochastic programming.

Risk averse and distributionally robust modeling of multistage stochastic programming – ALEXANDER SHAPIRO, GEORGIA INSTITUTE OF TECHNOLOGY –

It is known that there is a duality relation between risk-averse and distributionally robust approaches to stochastic programming. An extension from the static to dynamic (multistage) settings is not straightforward. This involves such basic concepts as conditional counterparts of risk-averse/distributionally robust functionals, dynamic equations and time consistency of nested formulations of multistage stochastic programs. In this talk we present a point of view on these topics in the frameworks of Stochastic Programming, Stochastic Optimal Control and MDP.

3d - WC328 - 515B - **Advances in Multi-stage Stochastic Programming and Reinforcement learning** Chair: YIZHE HUANG ↔

Policy Mirror Descent Inherently Explores Action Space – YAN LI, GEORGIA INSTITUTE OF TECHNOLOGY –

Explicit exploration in the action space was assumed to be indispensable for online policy gradient methods to avoid a drastic degradation in sample complexity, for solving general reinforcement learning problems over finite state and action spaces. We show that this is not needed by introducing new policy evaluation operators used in the policy mirror descent update. An $O(1/\epsilon^2)$ sample complexity is established. Notably, the evaluation operator uses no explicit exploration in estimating the state-action value function.

Nonconvex Landscape of Policy Optimization for A Class of Finite Horizon Markov Decision Processes and Applications in Operations Models – XIN CHEN, GEORGIA INSTITUTE OF TECHNOLOGY –

We explore policy gradient methods for finite horizon Markov Decision Processes (MDPs) with continuous state and action spaces. Policy gradient methods for MDPs do not converge to global optimal solutions in general due to the non-convexity of the objective functions. We identify several easily verifiable conditions to guarantee the global Kurdyka-Lojasiewicz (KL) condition for the objectives of policy gradient optimization problems for a class of MDPs. This allows to establish that the policy gradient optimization problems can be solved by first-order methods with a sample complexity in the order of $1/\epsilon$ and polynomial at the planning horizon length to attain ϵ -global optimal solutions. Our results find applications in a host of operations models including the stochastic cash balance problem and multi-period inventory system with Markov-modulated demand, giving the first sample complexity results in the literature.

Adversarially Robust Optimal Control with Causal Transport Distance – YIZHE HUANG, UNIVERSITY OF TEXAS AT AUSTIN – Rui Gao

We investigate stochastic optimal control problems in the presence of ambiguous disturbance distributions. To address this dynamic uncertainty, we propose a minimax distributionally robust formulation based on the ∞ -causal transport distance. We develop a nested dynamic programming reformulation for this problem. To tackle the dynamic programming problem involving continuous states and controls, we devise a stochastic dual dynamic programming algorithm. Numerical experiments on energy planning demonstrate the effectiveness of our approach.

3d - ThA246 - 515B - Statistical Learning, Optimization and Stochastic Programming for Reinforcement Learning Chair: MING YIN

HyperAgent: Advancing Scalable Exploration through Fast Uncertainty Estimation in RL – YINGRU LI, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN – Xu Jiawei, Lei Han, Zhi Quan Luo
We introduce HyperAgent, a scalable reinforcement learning (RL) algorithm that advances data-efficient exploration through fast incremental uncertainty estimation. HyperAgent leverages the hypermodel framework to efficiently approximate the posterior distribution of action-value functions, providing a simple yet stronger alternative to the epsilon-greedy exploration strategy and integrating seamlessly with established deep RL frameworks. Additionally, we extend this approach with GPT-derived feature embeddings in GPT-HyperAgent, enhancing uncertainty-aware reward modeling in contextual bandit problems with natural language input.

Theoretical analysis confirms that, with logarithmic per-step computational complexity, HyperAgent’s performance matches exact Thompson sampling (TS) in linear contextual bandits and Randomized Least-Square Value Iteration (RLSVI) in tabular RL environments. Fundamental to this analysis is the first probability tool for sequential random projection.

Empirical results support our theoretical findings, demonstrating HyperAgent’s robust performance in large-scale deep RL benchmarks and real-world safety-critical applications, such as automated content moderation with human feedback.

Variance-Dependent Regret Bounds for Non-stationary Linear Bandits – DONGRUO ZHOU, INDIANA UNIVERSITY –

We investigate the non-stationary stochastic linear bandit problem where the reward distribution evolves each round. Existing algorithms characterize the non-stationarity by the total variation budget B_K , which is the summation of the change of the consecutive feature vectors of the linear bandits over K rounds. However, such a quantity only measures the non-stationarity with respect to the expectation of the reward distribution, which makes existing algorithms sub-optimal under the general non-stationary distribution setting. In this work, we propose algorithms that utilize the variance of the reward distribution as well as the B_K , and show that they can achieve tighter regret upper bounds. Specifically, we introduce two novel algorithms that address cases where the variance information of the rewards is known and unknown, respectively. Notably, when the total variance V_K is much smaller than K , our algorithms outperform previous state-of-the-art results on non-stationary stochastic linear bandits under different settings. Experimental evaluations further validate the superior performance of our proposed algorithms over existing works.

Provable and Practical: Efficient Exploration in Reinforcement Learning via Langevin Monte Carlo – HAQUE ISHFAQ, MCGILL UNIVERSITY / MILA – Qingfeng Lan, Pan Xu, A. Rupam Mahmood, Doina Precup, Anima Anandkumar, Kamyar Aziz-zadenesheli

We present a scalable and effective exploration strategy based on Thompson sampling for reinforcement learning (RL). One of the key shortcomings of existing Thompson sampling algorithms is the need to perform a Gaussian approximation of the posterior distribution, which is not a good surrogate in most practical settings. We instead directly sample the Q function from its posterior distribution, by using Langevin Monte Carlo, an efficient type of Markov Chain Monte Carlo (MCMC) method. Our method only needs to perform noisy gradient descent updates to learn the exact posterior distribution of the Q function, which makes our approach easy to deploy in deep RL. We provide a rigorous theoretical analysis for the proposed method and demonstrate that, in the linear Markov decision process (linear MDP) setting, it has a sub-linear regret bound. We apply this approach to deep RL, by using Adam optimizer to perform gradient updates. Our approach achieves better or similar results compared with state-of-the-art deep RL algorithms on several challenging exploration tasks from the Atari57 suite.

Rethinking of Deep RL: Learning the Target Network in Function Space – MING YIN, PRINCETON UNIVERSITY – Kavosh Asadi, Shoham Sabach, Yao Liu, Mengdi Wang, Rasool Fakoor

We focus on the task of learning the value function in the approximate reinforcement learning (RL) setting. Existing algorithms solve this task by updating a pair of online and target networks while ensuring that the parameters of these two networks are equivalent. We propose Lookahead-Replicate (LR), a new value-function approximation algorithm that is agnostic to this parameter-space equivalence. Instead, the algorithm is designed to maintain an equivalence between the two networks in the function space, which is obtained by employing a new

target-network update. We show that LR leads to a convergent behavior in learning the value function. We also present empirical results demonstrating that LR-based updates significantly improve the performance of deep RL on the Atari benchmark.

3d - ThB186 - 515B - **Efficient RL: Adaptive Experimentation and Estimation** Chair: IMON BANERJEE ↩

Stochastic Programming Without Assuming a Probability Distribution – DAVID WOODRUFF, UC DAVIS –

The presentation will begin by describing smooth bootstrap applied to confidence intervals for the optimality gap associated with a solution to a stochastic program. This extends work where stochastic programming is used based on a sample from an unknown probability distribution. If there is time, the second part of the talk will describe a software architecture for computing solutions with upper and lower bounds in an asynchronous parallel environment.

Bayesian Reinforcement Learning with Limited Evaluations and Nonlinear Objectives – SOMAYEH MOAZENI, STEVENS INSTITUTE OF TECHNOLOGY –

This talk studies Bayesian Reinforcement Learning problems where the number of evaluations is constrained. We consider a versatile framework where (1) the utility function can exhibit nonlinearity with respect to the actions and (2) each action is defined by a set of (continuous or categorical) features within a solution space constrained by equalities and inequalities. This framework accommodates a wide array of practical scenarios across diverse application domains. We present efficient methods for computing a forward-looking measurement policy, known as the knowledge-gradient (KG) policy or this sequential learning problem setting. We evaluate the performance of our computational approach and optimal learning policy against popular benchmark policies from the sequential Bayesian ranking and selection and Bayesian Bandits.

Planning Adaptive Experiments with Model-Predictive Control – ETHAN CHE, COLUMBIA BUSINESS SCHOOL – Jimmy Wang, Daniel Jiang, Hongseok Namkoong

Implementing adaptive experimentation methods in the real world often encounters a multitude of operational difficulties, including batched/delayed feedback, non-stationary environments, and constraints on treatment allocations. To improve the flexibility of adaptive experimentation, we propose a Bayesian, optimization-based framework founded on model-predictive control (MPC) for the linear contextual bandit setting. While we focus on simple regret minimization, the framework can flexibly incorporate multiple objectives along with constraints, batches, personalized and non-personalized policies, as well as predictions of future context arrivals. Most importantly, it maintains this flexibility while guaranteeing improvement over non-adaptive A/B testing across all time horizons, and outperforms standard policies such as Thompson Sampling in an empirical benchmark study with non-stationary rewards. Overall, this framework offers a way to guide adaptive designs across the varied demands of modern large-scale experiments.

Sample efficient estimation of the transition kernels of controlled Markov chains – IMON BANERJEE, NORTHWESTERN UNIVERSITY –

We will present estimation bounds on non-parametric estimates of the transition kernels of Controlled Markov chains (CMC's). CMC's are a natural choice for modelling various industrial and medical processes, and are also relevant to reinforcement learning (RL). Therefore, learning the transition dynamics of CMC's in a sample efficient manner is an important question. We will attempt to answer this question when the underlying state-control space is both finite and infinite. Under finiteness, we will develop a Probably Approximately Correct (PAC) bound for estimation error. We will explore the additional challenges when the state-control space is infinite, and tackle them using techniques from adaptive estimation. At the end, we will also posit some open questions.

3d - ThC255 - 515B - **Adaptive stochastic optimization with applications in energy system operations and transition** Chair: ANDY SUN ↩

Design of Optimized Proximal Splitting Algorithms – PETER BARKLEY, NAVAL POSTGRADUATE SCHOOL –

We present a novel semi-definite programming method which allows us to design custom proximal splitting algorithms for solving convex composite minimization problems. These algorithms can be tailored to the compute architecture and problem in a way that minimizes the expected time required to converge. We present convergence rates for these algorithms derived using the PEP framework, which also allows us to find the optimal step size for the algorithm.

A New Risk Assessment and Management Paradigm in Electricity Markets – PANAGIOTIS ANDRIANESIS, BOSTON UNIVERSITY – Alex Wasilkoff, Michael Caramanis

We present results of an ARPA-E funded project under the PERFORM - Performance-based Energy Resource Feedback, Optimization, and Risk Management - program, which develops innovative management systems that represent the relative delivery risk of each asset, like wind farms, and balance the collective risk of all assets across the grid. We propose a New Risk Assessment and Management Paradigm designed to overhaul Electricity Markets by efficiently addressing uncertainty in the forthcoming massive renewable generation and electrification of fossil fuel reliant energy uses. We present methodologies constituting a risk-driven paradigm to achieve higher adoption of stochastic resources and a more efficient and reliable system operation, and we provide proof of concept on the Southwest Power Pool territory.

Bundle Method Approaches for Capacity Expansion Planning with Large Network and Temporal Scales – THOMAS LEE, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Andy Sun

Renewable electricity systems underpin global decarbonization efforts. For computational tractability, capacity expansion models often employ simplifications of temporal or network details. Yet the value of new resource investments, such as storage or transmission, are affected by both temporal and

spatial coupling. Improving accuracy motivates algorithmic improvements. Bundle method approaches are examined to solve nodal-resolution capacity expansion models with reasonable temporal resolution. Various practical implementation improvements are considered. Case studies are conducted on realistic large-scale US networks.

3d - FA182 - 515B - Advances at the intersection of reinforcement learning and stochastic optimization Chair: SAEED GHADIMI ↩

A Recurrent Reinforcement Learning Strategy for Online Scheduling of a State Task Network Under Uncertainty – DANIEL RANGEL MARTINEZ, UNIVERSITY OF WATERLOO – Luis Ricardez-Sandoval

Deep Reinforcement Learning (DRL) has gained interest for optimal scheduling in industrial facilities. The aim is to train an agent to find the optimal sequence of decisions in the performance of a task, e.g., a short-term schedule for manufacturing production. Key features of DRL agents for scheduling applications under uncertainty are the unique method to explore solutions in complex systems, fast adaptation to unexpected deviations in the parameters, and their ability to provide immediate responses. DRL has just emerged in the field of short-term scheduling optimization, showing good capabilities in creating schedules through different techniques. In the literature, the usual approach for the decision-making models in DRL is the Markov Decision Process (MDP) which uses information from the present time to take the consecutive action in the schedule. Nevertheless, MDPs might not be attractive to model the decision process in scheduling because of the evident correlation between events in scheduling. In this work we proposed the use of a Partially Observable MDP (POMDP) that considers a set of consecutive events to make a decision.

In this work, a methodology to develop a RL agent that acts as an online scheduler on a State Task Network (STN) subject to uncertainty is presented. Proximal Policy Optimization (PPO) is used to train a hybrid agent capable to provide multiple actions [1]. To account for the partial observability of the environment, the agent receives multiple states and correlates them using a Recurrent Neural Network (RNN). The agent produces a hybrid action at every time interval: one that starts a task in the network and the other that specifies the amount of product to be initialized. The hybrid agent was trained to maximize production in a given horizon. To the authors' knowledge, the use of a hybrid RNN agent has not been considered to address scheduling in STNs under uncertainty.

The hybrid agent designs a schedule considering the objective, uncertainty, and constraints of the system by using a reward shaping strategy. The DRL framework was tested on two case studies involving the online scheduling of two well-known STNs [2], [3]. Parametric uncertainty was incorporated in both cases to show the agent's adaptability to different circumstances. Results showed that the agent extracted useful knowledge from the instances to create attractive scheduling decisions.

[1] Z. Fan, R. Su, W. Zhang, and Y. Yu, "Hybrid Actor-Critic Reinforcement Learning in Parameterized Action Space." arXiv, May 30, 2019. Accessed: Jan. 18, 2024. [Online]. Available: <http://arxiv.org/abs/1903.01344>

[2] E. Kondili, C. C. Pantelides, and R. W. H. Sargent, "A general algorithm for short-term scheduling of batch operations—I. MILP formulation," *Comput. Chem. Eng.*, vol. 17, no. 2, Art. no. 2, Feb. 1993, doi: 10.1016/0098-1354(93)80015-F.

[3] L. G. Papageorgiou and C. C. Pantelides, "Optimal Campaign Planning/Scheduling of Multipurpose Batch/Semicontinuous Plants. 2. A Mathematical Decomposition Approach," *Ind. Eng. Chem. Res.*, vol. 35, no. 2, pp. 510–529, Jan. 1996, doi: 10.1021/ie950082d.

A Unified Framework for Sequential Decision Problems: From Dynamic Programming/RL to Stochastic Programming – WARREN POWELL, PRINCETON UNIVERSITY/OPTIMAL DYNAMICS –

"Stochastic programming" is typically associated with two (or multi) stage math programs using scenario trees, or methods based on Benders decomposition. These methods are treated as competitive with approximate dynamic programming/reinforcement learning and model predictive control, and are distinct from recently popular methods from "AI" based on deep neural networks. I refer to these (and many other methods) as the "jungle of stochastic optimization." I will pull these (and other) methods together in a single universal framework that can be used to model any sequential decision problem. Every method that has been proposed in the academic literature (or used in practice) can be organized into one of four (meta) classes of policies. I will highlight the power of parameterized policies, which are easily the most widely used methods in practice. I will then describe the concept of a parametric cost function approximation which is simply a parameterized deterministic approximation which has to be properly tuned in a simulator (but is sometimes tuned in the field). While the academic literature has focused on stochastic lookahead policies (using scenario trees) and methods based on Bellman's equation (which includes Benders decomposition, approximate dynamic programming and reinforcement learning), industry almost exclusively uses different forms of parameterized policies, often applied in an ad hoc fashion. Parameterized policies are far more practical, but introduce the challenges of designing the parameterization, and tuning the parameters. I will close by arguing that if the stochastic programming community shifts its focus to these classes of policies, our work will be much more useful in practice.

Inducing Group Fairness in Real-World Pipelines through Uncertainty Protections – SWATI GUPTA, MIT – Mehrdad Ghadiri

When the data used for optimization and decision-making is uncertain or random, optimum solutions can have a disparate impact on different demographic groups. This impact arises from variations in the probability distributions of data corresponding to different groups and the interactions of our estimates about these with optimization. In this talk, we will present a framework to tackle such disparities using "uncertainty protections" on different groups of people. We will discuss provable strategies to mitigate the resultant disparities, depending on the properties of the notion of fairness used. We will augment the theory with promising computational results on (i) school allocation with differentially private data and (ii) inventory management with stochastic demands.

A Functional Model Method for Nonconvex

Nonsmooth Conditional Stochastic Optimization – ANDRZEJ RUSZCZYNSKI, RUTGERS UNIVERSITY – Shangzhe Yang

We consider stochastic optimization problems involving an expected value of a nonlinear function of a base random vector and a conditional expectation of another function depending on the base random vector, a dependent random vector, and the decision variables. We call such problems conditional stochastic optimization problems. They arise in many applications, such as uplift modeling, reinforcement learning, and contextual optimization. We propose a specialized single time-scale stochastic method for nonconvex constrained conditional stochastic optimization problems with a Lipschitz smooth outer function and a generalized differentiable inner function. In the method, we approximate the inner conditional expectation with a rich parametric model whose mean squared error satisfies a stochastic version of a Lojasiewicz condition. The model is used by an inner learning algorithm. The main feature of our approach is that unbiased stochastic estimates of the directions used by the method can be generated with one observation from the joint distribution per iteration, which makes it applicable to real-time learning. The directions, however, are not gradients or subgradients of any overall objective function. We prove the convergence of the method with probability one, using the method of differential inclusions and a specially designed Lyapunov function, involving a stochastic generalization of the Bregman distance. Finally, a numerical illustration demonstrates the viability of our approach.

3d - FB109 - 515B - **Advances in multi-stage stochastic programming** Chair: BERNARDO PAGNONCELLI

Multi-Stage Stochastic Programming for Integrated Hurricane Evacuation and Logistics Planning – YONGJIA SONG, CLEMSON UNIVERSITY – Sudhan Bhattarai

We study an integrated hurricane relief logistics and evacuation planning problem, integrating hurricane evacuation and relief item pre-positioning operations that are typically treated separately. We propose a fully adaptive multi-stage stochastic programming (MSSP) model and solution approaches based on two-stage stochastic programming (2SSP). Utilizing historical forecast errors modeled using the auto-regressive model of order one, we generate hurricane scenarios and approximate the hurricane process as a Markov chain, and each Markovian state is characterized by the hurricane's location and intensity attributes. We conduct a comprehensive numerical experiment based on case studies motivated by Hurricane Florence and Hurricane Ian. Through the computational results, we demonstrate the value of fully adaptive policies given by the MSSP model over static ones given by the 2SSP model in terms of the out-of-sample performance. By conducting an extensive sensitivity analysis, we offer insights into how the value of fully adaptive policies varies in comparison to static ones with key problem parameters.

Multistage Chance-Constrained Programming – HAMED RAHIMIAN, CLEMSON UNIVERSITY

In this talk, we study multistage chance-constrained programming. We consider two possibilities for the chance constraint:

(i) enforced statically, just at the end of the planning horizon and (ii) enforced dynamically over time. For both cases, we provide decomposition-based computational approaches. We illustrate preliminary computational results on a stylized inventory problem. Moreover, we discuss the value of a multi-stage policy compared to a policy obtained from a two-stage multiperiod model.

Interpretable Vaccine Administration and Inventory Replenishment Policies via Smooth-in-Expectation Decision Rules – MERVE BODUR, UNIVERSITY OF EDINBURGH – Zoha Sherkat Masoumi, Jim Luedtke

The effective management of vaccine vials and inventory is imperative for ensuring widespread immunization coverage. We aim to address the challenges associated with this problem, including the need for interpretable policies. We propose a Markov decision process model, which offers flexibility to accommodate various operational aspects such as patient queues, clinic early closure considerations, and the requirement to serve every patient until exhausting vaccine inventory. For developing interpretable policies, we employ smooth-in-expectation decision rules, a recently proposed approach for multistage stochastic programs with mixed-integer recourse decisions and very many stages (e.g., hundreds). Leveraging these decision rules, we formulate and optimize the vaccine administration policies via a problem-specific flowchart design, alongside vial ordering decisions, facilitating interpretability and adaptability in real-world healthcare settings. We implement a batch stochastic gradient descent algorithm to solve the optimization problem, where for the initialization step we use the solution of a two-stage stochastic programming model we propose as an approximation to the multistage problem. Through extensive numerical experiments, we demonstrate the efficiency of the proposed approach and highlight the efficacy of various policies, including those considering patient queues.

Solving multistage equilibrium problems: the Krusell-Smith model – BERNARDO PAGNONCELLI, SKEMA BUSINESS SCHOOL – Mateus Waga, Davi Valladão, Andy Philpott

In this study, we consider multistage problems involving multiple agents, commonly recognized as stochastic dynamic games. Tackling such problems poses a formidable challenge, particularly in real-world scenarios where a multitude of agents are at play. We present a general formulation of the problem, and we focus on an incomplete market, heterogeneous agent model with aggregate uncertainty, known as the Krusell-Smith model. We first show the collusive solution is different from the Nash equilibrium one in a simple, three-period case. Our numerical exploration commences by discussing various strategies traditionally employed to derive equilibrium solutions for this complex problem. We begin with the moment approach, initially proposed by Krusell-Smith in '98. We then introduce a novel strategy that distinguishes itself by not assuming a predefined function form for the agents' capital dynamics. Instead, our approach involves constructing a Markov Chain at each iteration, based on the states visited. Furthermore, we conduct a comparative analysis of our solutions against those obtained through conventional approaches, providing insights into the relative strengths and weaknesses of each methodology.

3e - MB254 - 514B - **Data-driven optimization in urban systems** Chair: MANXI WU ↔

Network Flow Problems with Electric Vehicles

– HARIPRIYA PULYASSARY, CORNELL UNIVERSITY – Kostas Kollias, Aaron Schild, David Shmoys, Manxi Wu

Electric vehicle (EV) adoption in long-distance logistics faces challenges such as range anxiety and uneven distribution of charging stations. Two pivotal questions emerge: How can EVs be efficiently routed in a charging network considering range limits, charging speeds and prices? And, can the existing charging infrastructure sustain the increasing demand for EVs in long-distance logistics? This paper addresses these questions by introducing a novel theoretical and computational framework to study the EV network flow problems. We present an EV network flow model that incorporates range constraints and nonlinear charging rates, and identify conditions under which polynomial-time solutions can be obtained for optimal single EV routing, maximum flow, and minimum-cost flow problems. Our findings provide insights for optimizing EV routing in logistics, ensuring an efficient and sustainable future.

Pricing shared rides – JULIA YAN, UNIVERSITY OF BRITISH COLUMBIA – Chiwei Yan, Yifan Shen

Shared rides, which pool individual riders into a single vehicle, are essential for mitigating congestion and promoting more sustainable urban transportation. However, major ridesharing platforms have long struggled to maintain a healthy and profitable shared rides product. To understand why shared rides have struggled, we analyze procedures commonly used in practice to set static prices for shared rides, and discuss their pitfalls. We then propose a pricing policy that is adaptive to matching outcomes, dubbed match-based pricing, which varies prices depending on whether a rider is dispatched alone or to what extent she is matched with another rider. Analysis on a single origin-destination setting reveals that match-based pricing is both profit-maximizing and altruistic, simultaneously improving cost efficiency (i.e., the fraction of cost saved by shared rides relative to individual rides) and reducing rider payments relative to the optimal static pricing policy. These theoretical results are validated on a large-scale simulation with hundreds of origin-destinations from Chicago ridesharing data. The improvements in efficiency and reductions in payments are especially noticeable when costs are high and demand density is low, enabling healthy operations where they have historically been most challenging.

A Generative Learning Approach for Data-Driven Robust Optimization – ARON BRENNER, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Jennifer Sun, Rahman Khorramfar, Saurabh Amin

Increasingly, data-driven and robust optimization approaches are being leveraged to optimize short and long-term planning decisions in energy systems under uncertain supply and demand. In particular, effective decision-making in future decarbonized energy systems necessitates modeling supply-side uncertainties, such as solar and wind availability, at a high spatio-temporal resolution. To this end, risk-averse decision makers require high-fidelity models that accurately characterize tradeoffs between nominal and worst-case operational outcomes and consequently yield decisions that are simultaneously low-cost and robust to future uncertainties. However, existing approaches for two-stage optimization in energy applications

trade off realism for tractability in representing both second-stage uncertainties and decision-making, potentially overestimating incurred costs and tipping the balance of planning towards undue risk aversion. Here, we propose a deep learning-based approach to risk-averse two-stage optimization that aims to reduce this conservative bias. Specifically, we propose utilizing a variational autoencoder with a monotonic decoder trained to generate realistic high-dimensional realizations of energy supply and demand that enter as uncertain parameters in the second stage. We then solve the risk-averse two-stage model using a constraint generation approach whose subproblem is solved in each iteration by maximizing second-stage costs - a function of generated supply-demand data - with respect to the autoencoder latent variable. Finally, we demonstrate the efficacy of our approach in a case study that considers a zonal representation of the New England generation and transmission system and compare against benchmark optimization approaches.

Information design for spatial resource allocation – MANXI WU, CORNELL UNIVERSITY – Ozan Candogan

In this paper, we study platforms where resources and jobs are spatially distributed, and resources have the flexibility to strategically move to different locations for better payoffs. The price of the service at each location depends on the number of resources present and the market size, which is modeled as a random state. Our focus is on how the platform can utilize information about the underlying state to influence resource repositioning decisions and ultimately increase commission revenues. We establish that in many practically relevant settings a simple monotone partitional information disclosure policy is optimal. This policy reveals state realizations below a threshold and above a second (higher) threshold, and pools all states in between and maps them to a unique signal realization. We also provide algorithmic approaches for obtaining (near-)optimal information structures that are monotone partitional in general settings.

3e - MC177 - 514C - **Learning for Inventory** Chair: YAQI XIE ↔

Hindsight Learning for MDPs with Exogenous Inputs – SEAN SINCLAIR, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Felipe Frujeri, Ching An Cheng, Luke Marshall, Hugo Barbalho, Jingling Li, Jennifer Neville, Ishai Menache, Adith Swaminathan

Many resource management problems require sequential decision-making under uncertainty, where the only uncertainty affecting the decision outcomes are exogenous variables outside the control of the decision-maker. We model these problems as Exo-MDPs (Markov Decision Processes with Exogenous Inputs) and design a class of data-efficient algorithms for them termed Hindsight Learning (HL). Our HL algorithms achieve data efficiency by leveraging a key insight: having samples of the exogenous variables, past decisions can be revisited in hindsight to infer counterfactual consequences that can accelerate policy improvements. We compare HL against classic baselines in the multi-secretary, airline revenue management, and newsvendor problems. We also scale our algorithms to a business-critical cloud resource management problem - allocating Virtual Machines (VMs) to physical machines, and simulate their performance with real datasets from a large pub-

lic cloud provider. We find that HL algorithms outperform domain-specific heuristics, as well as state-of-the-art reinforcement learning methods.

VC Theory for Inventory Policies – YAQI XIE, UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS – Will Ma, Linwei Xin

Advances in computational power and AI have increased interest in reinforcement learning approaches to inventory management. This paper provides a theoretical foundation for these approaches and investigates the benefits of restricting to policy structures that are well-established by decades of inventory theory. In particular, we prove generalization guarantees for learning several well-known classes of inventory policies, including base-stock and (s, S) policies, by leveraging the celebrated Vapnik-Chervonenkis (VC) theory. We apply the concepts of the Pseudo-dimension and Fat-shattering dimension from VC theory to determine the generalizability of inventory policies, that is, the difference between an inventory policy's performance on training data and its expected performance on unseen data. We focus on a classical setting without contexts, but allow for an arbitrary distribution over demand sequences and do not make any assumptions such as independence over time. We corroborate our supervised learning results using numerical simulations. Managerially, our theory and simulations translate to the following insights. First, there is a principle of "learning less is more" in inventory management: depending on the amount of data available, it may be beneficial to restrict oneself to a simpler, albeit suboptimal, class of inventory policies to minimize overfitting errors. Second, the number of parameters in a policy class may not be the correct measure of overfitting error: in fact, the class of policies defined by T time-varying base-stock levels exhibits a generalization error comparable to that of the two-parameter (s, S) policy class. Finally, our research suggests situations in which it could be beneficial to incorporate the concepts of base-stock and inventory position into black-box learning machines, instead of having these machines directly learn the order quantity actions.

Costly Data Collection in the Newsvendor Problem – ZIJIN ZHANG, UNIVERSITY OF MICHIGAN ANN ARBOR –

In recent years, data-driven algorithms and policies have been developed and widely used in many applications. While some applications can utilize the preexisting data, there are many cases where collecting data is very expensive and time-consuming. Previous research has mainly focused on improving decision-making with data, leaving a gap in understanding the optimal quantity and quality of data needed. Our paper studies the data-driven variant of one of the most widely used operations models, the newsvendor model, in which the retailer is selling products with uncertain demand. Effective ordering decisions can only be made when the retailer collects insightful data to learn about the demand distribution. However, the accuracy of data can vary, and collecting this data is costly due to the sample size (quantity) and sample noise (quality). We aim to understand how the quantity and quality of data impact the retailer's newsvendor profit and to develop policies for data collection to enhance profit margins. In this paper, we provide a novel denoised-SAA approach that minimizes the in-sample newsvendor cost given a noisy dataset, and prove several bounds of the cost margin on the quantity and qual-

ity. Based on our in-depth theoretical analysis, we develop a series of data collection policies with analytical performance guarantees. In an extensive set of computational experiments, we show that these policies perform well in realistic settings.

Learning to Order for Inventory Systems with Lost Sales and Uncertain Supplies – JIASHUO JIANG, HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY – Chen Boxiao, Jiawei Zhang, Zhengyuan Zhou

We consider a stochastic lost-sales inventory control system with a lead time L over a planning horizon T . Supply is uncertain, and is a function of the order quantity (due to random yield/capacity, etc). We aim to minimize the T -period cost, a problem that is known to be computationally intractable even under known distributions of demand and supply. In this paper, we assume that both the demand and supply distributions are unknown and develop a computationally efficient online learning algorithm. We show that our algorithm achieves a regret (i.e. the performance gap between the cost of our algorithm and that of an optimal policy over T periods) of $O(L + \sqrt{T})$ when $L \geq \log(T)$. We do so by 1) showing our algorithm cost is higher by at most $O(L + \sqrt{T})$ for any $L \geq 0$ compared to an optimal constant-order policy under complete information (a well-known and widely-used algorithm) and 2) leveraging its known performance guarantee from the existing literature. To the best of our knowledge, a finite-sample $O(\sqrt{T})$ (and polynomial in L) regret bound when benchmarked against an optimal policy is not known before in the online inventory control literature. A key challenge in this learning problem is that both demand and supply data can be censored; hence only truncated values are observable. We circumvent this challenge by showing that the data generated under an order quantity q_2 allows us to simulate the performance of not only q_2 but also q_1 for all q_1

3e - TA134 - 514C - **Robust or Go Bust: Tales in Distributionally Robust and Data-Driven Optimization** Chair: WOLFRAM WIESEMANN [↩](#)

Distributionally Robust Linear Quadratic Control – DAN IANCU, STANFORD GRADUATE SCHOOL OF BUSINESS – Bahar Taskesen, Cagil Kocyigit, Daniel Kuhn

The Linear-Quadratic-Gaussian (LQG) problem is a fundamental control paradigm that has been studied and applied in various fields such as engineering, computer science, economics, and neuroscience. It involves controlling a system with linear dynamics affected by additive noise and with imperfect state observations, with the goal of minimizing a quadratic cost function depending on the state and control variables. In this work, we consider a generalization of the discrete-time, finite-horizon LQG problem where the noise distributions are unknown and belong to Wasserstein ambiguity sets centered at nominal (Gaussian) distributions. The objective is to minimize a worst-case cost across all distributions in the ambiguity set, including non-Gaussian distributions. Despite the added complexity, we prove that a control policy that is linear in the observations is optimal, as in the classic LQG problem. We propose a numerical solution method that efficiently characterizes this optimal control policy. Our method uses the Frank-Wolfe algorithm to identify the least-favorable distribu-

tions within the Wasserstein ambiguity sets and computes the controller's optimal policy using Kalman filter estimation under these distributions.

Randomized Assortment Optimization – WOLFRAM WIESEMAN, IMPERIAL COLLEGE LONDON – Zhengchao Wang, Heikki Peura

When a firm selects an assortment of products to offer to customers, it uses a choice model to anticipate their probability of purchasing each product. In practice, the estimation of these models is subject to statistical errors, which may lead to significantly suboptimal assortment decisions. Recent work has addressed this issue using robust optimization, where the true parameter values are assumed unknown and the firm chooses an assortment that maximizes its worst-case expected revenues over an uncertainty set of likely parameter values, thus mitigating estimation errors. In this talk, we introduce the concept of randomization into the robust assortment optimization literature. We show that the standard approach of deterministically selecting a single assortment to offer is not always optimal in the robust assortment optimization problem. Instead, the firm can improve its worst-case expected revenues by selecting an assortment randomly according to a prudently designed probability distribution. We demonstrate this potential benefit of randomization both theoretically in an abstract problem formulation as well as empirically across three popular choice models: the multinomial logit model, the Markov chain model, and the preference ranking model. We show how an optimal randomization strategy can be determined exactly and heuristically. Besides the superior in-sample performance of randomized assortments, we demonstrate improved out-of-sample performance in a data-driven setting that combines estimation with optimization. Our results suggest that more general versions of the assortment optimization problem—incorporating business constraints, more flexible choice models and/or more general uncertainty sets—tend to be more receptive to the benefits of randomization.

Wasserstein Distributionally Robust Optimization with Heterogeneous Data Sources – DANIEL KUHN, EPFL – Yves Rychener, Adrian Esteban Perez, Juan M. Morales

We study decision problems under uncertainty, where the decision-maker has access to K data sources that carry biased information about the underlying risk factors. The biases are measured by the mismatch between the risk factor distribution and the K data-generating distributions with respect to an optimal transport (OT) distance. In this situation the decision-maker can exploit the information contained in the biased samples by solving a distributionally robust optimization (DRO) problem, where the ambiguity set is defined as the intersection of K OT neighborhoods, each of which is centered at the empirical distribution on the samples generated by a biased data source. We show that if the decision-maker has a prior belief about the biases, then the out-of-sample performance of the DRO solution can improve with K ; irrespective of the magnitude of the biases. We also show that, under standard convexity assumptions, the proposed DRO problem is computationally tractable if either K or the dimension of the risk factors is kept constant.

Learning Optimal and Fair Policies for Online Allocation of Scarce Societal Resources

from Data Collected in Deployment – CAGIL KOCYIGIT, UNIVERSITY OF LUXEMBOURG –

We study the problem of allocating scarce societal resources of different types (e.g., permanent housing, deceased donor kidneys for transplantation, ventilators) to heterogeneous allocatees on a waitlist (e.g., people experiencing homelessness, individuals suffering from end-stage renal disease, Covid-19 patients) based on their observed covariates. We leverage administrative data collected in deployment to design an online policy that maximizes expected outcomes while satisfying budget constraints, in the long run. Our proposed policy waitlists each individual for the resource maximizing the difference between their estimated mean treatment outcome and the estimated resource dual-price or, roughly, the opportunity cost of using the resource. Resources are then allocated as they arrive, in a first-come first-serve fashion. We demonstrate that our data-driven policy almost surely asymptotically achieves the expected outcome of the optimal out-of-sample policy under mild technical assumptions. We extend our framework to incorporate various fairness constraints. We evaluate the performance of our approach on the problem of designing policies for allocating scarce housing resources to people experiencing homelessness in Los Angeles based on data from the homeless management information system. In particular, we show that our policies simultaneously improve exit rates from homelessness and enhance fairness compared to historical allocations.

3e - TA158 - 512C - Statistical and Computational Aspects of Robust Estimation Chair: SOROOSH SHAFIEE

Outlier-Robust Wasserstein DRO – SOROOSH SHAFIEE, CORNELL –

Distributionally robust optimization (DRO) is an effective approach for data-driven decision-making in the presence of uncertainty. Geometric uncertainty due to sampling or localized perturbations of data points is captured by Wasserstein DRO (WDRO), which seeks to learn a model that performs uniformly well over a Wasserstein ball centered around the observed data distribution. However, WDRO fails to account for non-geometric perturbations such as adversarial outliers, which can greatly distort the Wasserstein distance measurement and impede the learned model. We address this gap by proposing a novel outlier-robust WDRO framework for decision-making under both geometric (Wasserstein) perturbations and non-geometric (total variation (TV)) contamination that allows an ϵ -fraction of data to be arbitrarily corrupted. We design an uncertainty set using a certain robust Wasserstein ball that accounts for both perturbation types and derive minimax optimal excess risk bounds for this procedure that explicitly capture the Wasserstein and TV risks. We prove a strong duality result that enables tractable convex reformulations and efficient computation of our outlier-robust WDRO problem. When the loss function depends only on low-dimensional features of the data, we eliminate certain dimension dependencies from the risk bounds that are unavoidable in the general setting. Finally, we present experiments validating our theory on standard regression and classification tasks.

Coherent Wasserstein Distributionally Robust Optimization – JONATHAN LI, UNIVERSITY OF OTTAWA – Tiantian Mao

Wasserstein distributionally robust optimization (DRO), as a key paradigm in data-driven optimization, aims to provide decisions with out-of-sample performance guarantees by optimizing against a set of distributions close to the empirical one in Wasserstein p -distance, where $p \in [1, \infty]$. The set, known as the Wasserstein ball, becomes less conservative as $p \rightarrow \infty$, containing distributions more heavily concentrated around the data, i.e., more data-driven. Despite this attractive feature of higher order p , the type-1 Wasserstein distance, i.e., $p = 1$, remains most popular to date. The advantages of the type-1 distance are threefold: from a modelling perspective, it accommodates richer families of distributions; computationally, it leads to Wasserstein DRO problems with favourable complexity, often solvable via linear programs; and statistically, it ensures the best rate of convergence for finite-sample guarantees. We introduce a novel alternative, coherent Wasserstein metrics, achieving flexibility comparable to type- p metrics for reducing conservatism, while inheriting the advantages of the type-1 distance. These metrics are derived from risk-averse formulations of the optimal transport problem with coherent risk measures. We show that distributionally robust optimization problems using coherent Wasserstein metrics, termed coherent Wasserstein distributionally robust optimization (CWDRO), can retain key properties of type-1 Wasserstein DRO, while potentially achieving better utilization of data.

Universal Gradient Descent Ascent Method for Smooth Minimax Optimization – TAOLI ZHENG, THE CHINESE UNIVERSITY OF HONG KONG – Jiajin Li, Man Cho So

Smooth minimax optimization has garnered significant attention over the past decade, owing to its widespread applications in machine learning. Considerable research has focused on tailored algorithms for smooth minimax problems, addressing specific structural conditions like convexity or concavity of primal and dual functions, as well as Polyak-Łojasiewicz (PL) and Kurdyka-Łojasiewicz (KL) conditions. However, verifying these conditions is challenging in practice, motivating our pursuit of universal algorithms for smooth minimax problems. To answer this questions, we propose a novel universally applicable single-loop algorithm, the doubly smoothed optimistic gradient descent ascent method (DS-OGDA), which can be applied to convex-concave, nonconvex-concave, convex-nonconcave, and nonconvex-nonconcave problems with one-sided KL properties. With a same set of parameters, DS-OGDA can achieve convergence with $\mathcal{O}(\epsilon^{-4})$ complexity for nonconvex minimax problems and $\mathcal{O}(\epsilon^{-2})$ complexity for convex-concave minimax problems. Sharper (even optimal) iteration complexity can be achieved with known KL exponent or in convex-concave scenarios. Specifically, under the one-sided KL condition with exponent $\theta \in (0, 1)$, DS-OGDA converges with $\mathcal{O}(\epsilon^{-2 \max\{2\theta, 1\}})$ complexity. Additionally, DS-OGDA attains $\mathcal{O}(\epsilon^{-1})$ complexity in convex-concave settings. They all match the corresponding best results in the literature. We further demonstrate the universality of DS-OGDA through numerical experiments on a class of distributionally robust optimization problems.

The out-of-sample prediction error of the square-root lasso and related estimators – CYNTHIA RUSH, COLUMBIA UNIVERSITY –

We study the classical problem of predicting an outcome variable, Y , using a linear combination of a d -dimensional covariate vector, X . We are interested in linear predictors whose coefficients

solve: $\inf_{\beta} (E[(Y - \langle \beta, X \rangle)^r])^{1/r} + \delta \|\beta\|$, where $r \geq 1$ and $\delta \geq 0$ is a regularization parameter. We provide conditions under which linear predictors based on these estimators minimize the worst-case prediction error over a ball of distributions determined by a type of max-sliced Wasserstein metric. A detailed analysis of the statistical properties of this metric yields a simple recommendation for the choice of regularization parameter. The suggested order of δ , after a suitable normalization of the covariates, is typically d/n , up to logarithmic factors. Our recommendation is computationally straightforward to implement, pivotal, has provable out-of-sample performance guarantees, and does not rely on sparsity assumptions about the true data generating process.

3e - TB73 - 514C - **Data-driven Methods for Contextual Stochastic Optimization** Chair: MENG QI

Robust End-to-End Learning under Endogenous Uncertainty – RARES CRISTIAN, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Pavithra Harsha, Perakis Georgios, Brian Quanz

The decisions one takes can often affect the outcome observed. How can we effectively learn to make decisions when there are no ground-truth counterfactual observations? We propose a robust end-to-end learning approach to the contextual stochastic optimization problem under decision-dependent uncertainty. We tackle this problem by constructing uncertainty sets over the space of ML models and present efficient algorithms to solve these inherently non-convex optimization problems. We computationally test the proposed approach on multi-item pricing and assortment problems where demand is affected by cross-item complementary and supplementary effects. Moreover, we show the proposed approach outperforms traditional methods by more than 20%.

Robust Actionable Prescriptive Analytics – MINGLONG ZHOU, FUDAN UNIVERSITY – Li Chen, Xun Zhang, Long Zhao, Melvyn Sim

We propose a new robust actionable prescriptive analytics framework that leverages past data and side information to minimize a risk-based objective function under distributional ambiguity. Our framework aims to find a policy that directly transforms the side information into implementable decisions. Specifically, we focus on developing actionable response policies that offer the benefits of interpretability and implementability. To address the potential issue of overfitting to empirical data, we adopt a data-driven robust satisficing approach that effectively handles uncertainty. We tackle the computational challenge for linear optimization models with recourse by developing a new tractable safe approximation for robust constraints, accommodating bilinear uncertainty and general norm-based uncertainty sets. Additionally, we introduce a bi-affine recourse adaptation to enhance the quality of the approximation. Furthermore, we present a localized robust satisficing model that efficiently solves combinatorial optimization problems with tree-based static policies. Finally, we demonstrate the practical application of our framework through a simulation case study on risk-minimizing portfolio optimization using past returns as side information. We also provide a simulation case study on how the framework can be applied to obtain an interpretable policy for allocating taxis to different demand

regions in response to weather information.

Integrated Conditional Estimation-Optimization

– MENG QI, CORNELL UNIVERSITY – Paul Grigas, Zuo Jun Shen

Many real-world optimization problems involve uncertain parameters with probability distributions that can be estimated using contextual feature information. In contrast to the standard approach of first estimating the distribution of uncertain parameters and then optimizing the objective based on the estimation, we propose an integrated conditional estimation-optimization (ICEO) framework that estimates the underlying conditional distribution of the random parameter while considering the structure of the optimization problem. We directly model the relationship between the conditional distribution of the random parameter and the contextual features, and then estimate the probabilistic model with an objective that aligns with the downstream optimization problem. We show that our ICEO approach is asymptotically consistent under moderate regularity conditions and further provide finite performance guarantees in the form of generalization bounds. Computationally, performing estimation with the ICEO approach is a non-convex and often non-differentiable optimization problem. We propose a general methodology for approximating the potentially non-differentiable mapping from estimated conditional distribution to the optimal decision by a differentiable function, which greatly improves the performance of gradient-based algorithms applied to the non-convex problem. We also provide a polynomial optimization solution approach in the semi-algebraic case. Numerical experiments are also conducted to show the empirical success of our approach in different situations including with limited data samples and model mismatches.

Conformal Contextual Optimization with a Smart Predict-then-Optimize Method – PAUL GRIGAS, UC BERKELEY – Hyunki Im

In contextual optimization, a decision-maker utilizes a machine learning (ML) model for predicting uncertain parameters of a downstream optimization model based on contextual features. We study an extension of contextual stochastic linear optimization (CSLO) that, in contrast to most of the existing literature, involves inequality constraints that depend on the uncertain parameters that are predicted by the ML model. Building on previous work that develops the “Smart Predict-then-Optimize (SPO)” loss and its tractable SPO+ surrogate loss in the case of a known deterministic feasible region, and work that develops robust variants of contextual optimization using conformal prediction methods, we propose a “Conformal Smart Predict-then-Optimize (CSPO)” approach for addressing uncertainty in the constraints. Specifically, we first propose the CSPO loss – a direct extension of the SPO loss – that measures the decision error, or regret, induced by following a robust predict-then-optimize approach that uses a conformal prediction method to produce an uncertainty set. We then propose a convex surrogate loss function called the CSPO+ loss – a direct extension of the SPO+ loss – to tractably train a prediction model in our CSLO setting. To effectively train the model, we employ a data filtering procedure to address issues of infeasibility whenever the predicted uncertainty set fails to cover the true parameter values. Furthermore, we utilize importance sampling when training the model to address the distribution shift induced by the data filtering step. Theoretically, we es-

tablish statistical consistency of the CSPO+ loss relative to the CSPO loss and provide finite-sample convergence guarantees for the CSPO loss with importance sampling under mild assumptions. Experimentally, we demonstrate strong performance of the CSPO+ loss on several different CSLO problem classes.

3e - TC198 - 514C - Learning and Verification of Continuous Optimizers Chair: BRANDON AMOS

Data-Driven Performance Guarantees for Classical and Learned Optimizers – RAJIV SAMBHARYA, PRINCETON UNIVERSITY – Bartolomeo Stellato

We introduce a data-driven approach to analyze the performance of continuous optimization algorithms using generalization guarantees from statistical learning theory. We study classical and learned optimizers to solve families of parametric optimization problems. We build generalization guarantees for classical optimizers, using a sample convergence bound, and for learned optimizers, using the Probably Approximately Correct (PAC)-Bayes framework. To train learned optimizers, we use a gradient-based algorithm to directly minimize the PAC-Bayes upper bound. Numerical experiments in signal processing, control, and meta-learning showcase the ability of our framework to provide strong generalization guarantees for both classical and learned optimizers given a fixed budget of iterations. For classical optimizers, our bounds are much tighter than those that worst-case guarantees provide. Meanwhile, our generalization bounds for learned optimizers outperform the empirical outcomes observed in their non-learned counterparts.

Verification of First-Order Methods for Parametric Quadratic Optimization – VINIT RANJAN, PRINCETON UNIVERSITY – Bartolomeo Stellato

We introduce a numerical framework to verify the finite step convergence of first-order methods for parametric convex quadratic optimization. We formulate the verification problem as a mathematical optimization problem where we maximize a performance metric (e.g., fixed-point residual at the last iteration) subject to constraints representing proximal algorithm steps (e.g., linear system solutions, projections, or gradient steps). Our framework is highly modular because we encode a wide range of proximal algorithms as variations of two primitive steps: affine steps and element-wise maximum steps. Compared to standard convergence analysis and performance estimation techniques, we can explicitly quantify the effects of warm-starting by directly representing the sets where the initial iterates and parameters live. We show that the verification problem is NP-hard, and we construct strong semidefinite programming relaxations using various constraint tightening techniques. Numerical examples in nonnegative least squares, network utility maximization, Lasso, and optimal control show a significant reduction in pessimism of our framework compared to standard worst-case convergence analysis techniques.

Amortized optimization for optimal transport and LLM attacks – BRANDON AMOS, META –

Amortized optimization methods provide fast solvers by predicting approximate solutions to optimization problems. This talk covers two recent advancements using amortiza-

tion to significantly speed up the solvers of non-trivial optimization problems arising in the fields of optimal transport (OT) and large language model (LLM) attacks. Computational optimal transport problems may involve solving three nested optimization problems, each of which amortization can help with: 1) the solution map from the measures to the primal/dual OT solution (Meta OT: <https://arxiv.org/abs/2206.05262>), 2) the computation of the c-transform or Fenchel conjugate (amortized conjugates: <https://arxiv.org/abs/2210.12153>), and 3) the computation of geodesics and Lagrangian (minimum-action) paths/costs (Lagrangian OT: <https://openreview.net/pdf?id=myb0FKB8C9>). Adding amortization to the standard solvers in these OT settings significantly improves the runtime and deployment time of the methods. These faster amortized solutions to the Fenchel conjugate and geodesic/Lagrangian paths are of potential more general interest in other settings bottlenecked by numerical solutions to them. Beyond these optimal transport applications, we will also discuss the prompt optimization problems arising in adversarial attacks on LLMs (AdvPrompter: <https://arxiv.org/abs/2404.16873>). Here, amortization enables us to attain state-of-the-art results on the standard AdvBench dataset, that also transfer to closed-source black-box LLM APIs. The fast amortized predictions then enable us to generate a synthetic dataset of adversarial examples which an LLM can be fine-tuned on to make it more robust against jailbreaking attacks while maintaining performance.

3e - WA941 - 514C - **Data-driven optimization**
1 Chair: SAEID ABBASIPARIZI

Stochastic gradient methods with adaptive learning rates – ADRIAN RIEKERT, UNIVERSITY OF MÜNSTER – Steffen Dereich, Arnulf Jentzen

It is known that the standard stochastic gradient descent (SGD) method, but also more sophisticated adaptive stochastic optimization methods such as the Adam optimizer can in general not converge if the learning rates do not tend to zero. In practice, one often deals with this problem by employing a deterministic decreasing learning rate schedule. Nevertheless, the optimal learning rates highly depend on the considered optimization problem and are generally difficult to find. In this talk we propose a learning-rate-adaptive approach that updates the learning rate dynamically during the training process depending on the past behavior of the loss function. For a particular class of strongly convex target functions, using results from martingale and Markov chain theory, we prove convergence of the SGD method for such adaptive learning rate sequences. Additionally, we illustrate the performance of the proposed approach for several optimization problems arising in scientific machine learning such as data-driven solutions of partial differential equations.

A Tractable Online Learning Algorithm for the Multinomial Logit Contextual Bandit – PRIYANK AGRAWAL, COLUMBIA UNIVERSITY – Theja Tulabandhula, Vashist Avadhanula

The talk will be primarily on the contextual variant of the MNL-Bandit problem. More specifically, we consider a dynamic set optimization problem, where in every round, a decision maker offers a subset (assortment) of products to a consumer and observes their response. Consumers purchase products to maximize their utility. We assume that a set of at-

tributes describes the products, and the mean utility of a product is linear in the values of these attributes. We model consumer choice behavior using the widely used Multinomial Logit (MNL) model and consider the decision maker's problem of dynamically learning the model parameters while optimizing cumulative revenue over the selling horizon T . Though this problem has recently attracted considerable attention, many existing methods often involve solving an intractable non-convex optimization problem, and their theoretical performance guarantees depend on a problem-dependent parameter, which could be prohibitively large. In particular, existing algorithms for this problem have regret bounded by $O(\sqrt{\kappa d T})$, where κ is a problem-dependent constant that can have exponential dependency on the number of attributes. In this paper, we propose an optimistic algorithm and show that the regret is bounded by $O(\sqrt{dT} + \kappa)$, significantly improving the performance over existing methods. Further, we propose a convex relaxation of the optimization step, allowing for tractable decision-making while retaining the favorable regret guarantee.

Smoothness-Adaptive Dynamic Pricing with Nonparametric Demand Learning – HANSHENG JIANG, UNIVERSITY OF TORONTO – Zeqi Ye

We study the dynamic pricing problem where the demand function is nonparametric and Hölder smooth, and we focus on adaptivity to the unknown Hölder smoothness parameter β . Traditional algorithm relies on the knowledge of β to achieve a minimax optimal regret of $\tilde{O}(T^{\frac{\beta+1}{2\beta+1}})$. We firstly prove no policy can adaptively achieve the optimal regret without knowledge of β . Motivated by the impossibility result, we propose a self-similarity condition to enable adaptivity, which does not compromise the problem's inherent complexity. Furthermore, we develop a smoothness-adaptive algorithm and prove that the algorithm achieves this optimal regret bound without the knowledge β .

A Benders Decomposition Algorithm for the Overland Search and Rescue Problem Considering Search Area Selection – SAEID ABBASIPARIZI, PH.D. STUDENT AT DEPARTMENT OF OPERATIONS AND DECISION SYSTEMS, UNIVERSITÉ LAVAL, QUÉBEC, CANADA – Michael Morin, Irène Abi Zeid

Search and rescue incidents occur frequently in Canada, some requiring the deployment of major search operations. In such situations, search and rescue decision support systems are vital. In this regard, an overland search and rescue model considering search area selection is proposed to maximize the probability of finding the search object, constrained by the search effort available. A solution to our problem includes the optimal set of search areas (rectangles where to position search patterns), the path between them, and the search pattern within each rectangle. We developed a Benders decomposition to solve the proposed model and compare its performance, for different instance sizes, to that of a MILP model in CPLEX. The obtained results confirm the efficiency of our Benders decomposition.

3e - WB82 - 514C - **Data-driven Optimization Under Uncertainty: Theory and Applications** Chair: KARMELE SHEHADEH

Robust Design of Critical Supply Chains – HAO HAO, CARNEGIE MELLON UNIVERSITY – Peter Zhang, Ningji Wei

Critical supply chains such as semiconductor and pharmaceutical supply chains require long-term resiliency in addition to efficiency. The challenge of balancing resiliency and efficiency with limited resources in industrial-scale supply chains requires a principled approach. To this end, we propose an end-to-end methodology for robust designs of critical supply chains. Among other things, our framework captures the trade-off between cost and robustness under worst-case supply and demand disruptions; as well as the multiple time scales of resiliency building and contingency planning. We formulate a two-stage supply chain design model under uncertainty, the model optimizes for the risk-mitigating first-stage network design and resource allocation that maximize the total profit under uncertainty. We deploy a recently popular approximate solution method — affinely adjustable robust optimization (AARO) formulations — to our problem. With extensive numerical experiments, we show that this method outperforms other solution methods, such as sample average approximation, by orders of magnitude in this type of robust network design problems. We also provide theoretical support for the (near) optimality and tractability of this method. Most interestingly, we collect large and real-world datasets for both types of supply chains and conduct case studies. Our case studies show an important insight: a small amount of risk-mitigating resources meticulously placed at strategic locations can lead to a significant gain in supply chain resilience. Our methodology allows the systematic identification of such leverages at an industrial scale under highly uncertain supply and demand environments. This insight implies that with the right data and analytical tools, there is an opportunity to strengthen critical supply chains with relatively affordable investments.

Integer Programming Approaches for Distributionally Robust Chance Constraints with Adjustable Risks – YILING ZHANG, UNIVERSITY OF MINNESOTA –

We study distributionally robust chance-constrained programs (DRCCPs) with right-hand side uncertainty. The DRCCPs treat the risk tolerances associated with the distributionally robust chance constraints (DRCCs) as decision variables to trade off between the system cost and risk of violations by penalizing the risk tolerances in the objective function. We consider two types of Wasserstein ambiguity sets: one with finite support and one with a continuum of realizations. By exploring the hidden discrete structures, we develop mixed integer programming reformulations under the two types of ambiguity sets to determine the optimal risk tolerance for the chance constraint. Valid inequalities are derived to strengthen the formulations. We test instances with transportation problems of diverse sizes and a demand response management problem.

On the Trade-off Between Distributional Belief and Ambiguity: Conservatism, Bias, and Asymptotic Properties – KARMEI SHEHADEH, LEHIGH UNIVERSITY – Man Yiu Tsang

We propose a new data-driven trade-off (TRO) approach for modeling uncertainty that serves as a middle ground between the optimistic approach, which adopts a distributional belief, and the pessimistic distributionally robust optimization approach, which hedges against distributional ambiguity. We

equip the TRO model with a TRO ambiguity set characterized by a size parameter controlling the level of optimism and a shape parameter representing distributional ambiguity. We first show that constructing the TRO ambiguity set using a general star-shaped shape parameter with the empirical distribution as its star center is necessary and sufficient to guarantee the hierarchical structure of the sequence of TRO ambiguity sets. Then, we analyze the properties of the TRO model, including quantifying conservatism, quantifying bias, and establishing asymptotic properties. Specifically, we show that the TRO model could generate a spectrum of decisions, ranging from optimistic to conservative decisions. Additionally, we show that it could produce an unbiased estimator of the true optimal value. Furthermore, we establish the almost-sure convergence of the optimal value and the set of optimal solutions of the TRO model to their true counterparts. We exemplify our theoretical results using an inventory control problem and a portfolio optimization problem.

Residuals-Based Contextual Distributionally Robust Optimization with Decision-Dependent Uncertainty – XIAN YU, THE OHIO STATE UNIVERSITY – Zoey Zhu, Guzin Bayraksan

We consider a residuals-based distributionally robust optimization model, where the underlying uncertainty depends on both covariate information and our decisions. We adopt regression models to learn the latent decision dependency and construct a nominal distribution (and thereby ambiguity sets) around the learned model using empirical residuals from the regressions. Ambiguity sets can be formed via the Wasserstein distance, a sample robust approach, or with the same support as the nominal empirical distribution (e.g., ϕ -divergences), where both the nominal distribution and the radii of the ambiguity sets could be decision- and covariate-dependent. We provide conditions under which desired statistical properties, such as asymptotic optimality, rates of convergence, and finite sample guarantees, are satisfied. Via cross-validation, we devise data-driven approaches to find the best radii for different ambiguity sets, which can be decision-(in)dependent and covariate-(in)dependent. Through numerical experiments, we illustrate the effectiveness of our approach and the benefits of integrating decision dependency into a residuals-based DRO framework.

3e - WC942 - 514C - **Data-driven optimization**
2 Chair: MOHAMMAD REZA BELBASI

Differentially Private Linear Optimization for Multi-Party Resource Sharing – NURSEN AYDIN, ASSOCIATE PROFESSOR – Utku Karaca, Sinan Yildirim, Ilker Birbil

This study examines a resource-sharing problem involving multiple parties that agree to use a set of capacities together. We start with modeling the whole problem as a mathematical program, where all parties are required to exchange information to obtain the optimal objective function value. This information bears private data from each party in terms of coefficients used in the mathematical program. Moreover, the parties also consider the individual optimal solutions as private. In this setting, the concern for the parties is the privacy of their data and their optimal allocations. We propose a two-step approach to meet the privacy requirements of the parties. In the first step, we obtain a reformulated model that is amenable to a decom-

position scheme. Although this scheme eliminates almost all data exchanges, it does not provide a formal privacy guarantee. In the second step, we provide this guarantee with a locally differentially private algorithm, which does not need a trusted aggregator, at the expense of deviating slightly from the optimality. We provide bounds on this deviation and discuss the consequences of these theoretical results. We also propose a novel modification to increase the efficiency of the algorithm in terms of reducing the theoretical optimality gap. The study ends with a numerical experiment on a planning problem that demonstrates an application of the proposed approach. As we work with a general linear optimization model, our analysis and discussion can be used in different application areas including production planning, logistics, and revenue management.

Metric Learning to Accelerate Convergence of Operator Splitting Methods – JAMES KOTARY, UNIVERSITY OF VIRGINIA –

Recent work has shown a variety of ways in which machine learning can be used to accelerate the solution of constrained optimization problems. Increasing demand for real-time decision-making capabilities in applications such as artificial intelligence and optimal control has led to a variety of approaches, based on distinct strategies. This work proposes a novel approach to learning optimization, in which the underlying metric space of a proximal operator splitting algorithm is learned so as to maximize its convergence rate. While prior works in optimization theory have derived optimal metrics for limited classes of problems, the results do not extend to many practical problem forms including general Quadratic Programming (QP). This paper shows how differentiable optimization can enable the end-to-end learning of proximal metrics, enhancing the convergence of proximal algorithms for QP problems beyond what is possible based on known theory. Additionally, the results illustrate a strong connection between the learned proximal metrics and active constraints at the optima, leading to an interpretation in which the learning of proximal metrics can be viewed as a form of active set learning.

It's All in the Mix: Wasserstein Machine Learning with Mixed Features – MOHAMMAD REZA BELBASI, IMPERIAL COLLEGE LONDON –

Problem definition: The recent advent of data-driven and end-to-end decision-making across different areas of operations management has led to an ever closer integration of prediction models from machine learning and optimization models from operations research. A key challenge in this context is the presence of estimation errors in the prediction models, which tend to be amplified by the subsequent optimization model – a phenomenon that is often referred to as the Optimizer's Curse or the Error-Maximization Effect of Optimization. **Methodology/results:** A contemporary approach to combat such estimation errors is offered by distributionally robust problem formulations that consider all data-generating distributions close to the empirical distribution derived from historical samples, where 'closeness' is determined by the Wasserstein distance. While those techniques show significant promise in problems where all input features are continuous, they scale exponentially when binary and/or categorical features are present. In this talk, we demonstrate that such mixed-feature problems can indeed be solved in polynomial time. We present a practically efficient algorithm to solve mixed-feature problems, and we compare our method against alternative techniques both

theoretically and empirically on standard benchmark instances. **Managerial implications:** Data-driven operations management problems often involve prediction models with discrete features. We develop and analyze a methodology that faithfully accounts for the presence of discrete features, and we demonstrate that our approach can significantly outperform existing methods that are agnostic to the presence of discrete features, both theoretically and across standard benchmark instances.

3e - ThB194 - 514C - Learning and Calibration in Data-Driven Optimization Chair: TIANYU WANG

Service Oriented Considerate Routing: Data, Predictions and Robust Decisions – MELVYN SIM, NUS BUSINESS SCHOOL – Yue Zhao, Zhixing Luo, Stanley Lima, Caihua Chen

In this research, we focus on improving service oriented routing by addressing the nuanced challenge of punctuality through the consideration of couriers' ability to ensure on-time deliveries. We utilize a comprehensive real-world dataset from a cold chain logistics firm for analysis. Our empirical investigation indicates that relying solely on travel distance is inadequate for accurate delivery time prediction. We highlight critical elements, including couriers' fixed effects and workload, as key covariates to improve prediction performance. Distinguishing our work from existing literature, we integrate couriers' workload and location familiarity into our Service Oriented Routing (SOR) model to enhance predictions of delivery times. We introduce the Courier Assigned Location Mismatch (CALM) metric as a less intrusive approach to incorporating couriers' location familiarity into their delivery efficiency. We propose the novel Service Oriented Considerate Routing (SOCR) model; by minimizing the CALM metric, couriers are assigned routes within familiar territories to the extent possible within the total routing distance constraint. The considerate routing strategies could potentially reduce the stress couriers face when delivering in unfamiliar areas. Additionally, we develop the connection of the SOCR model with a robust satisficing approach. This strategy guarantees timely deliveries by effectively mitigating the effects of predictive inaccuracies and potential model misspecifications. To solve the SOCR model, we apply Benders decomposition for an exact solution and Tabu Search for a heuristic approach, demonstrating their effectiveness and superior out-of-sample performance. Notably, our heuristic solutions significantly outperform exact solutions of classical vehicle routing problems with deadlines, resulting in substantial improvements in timely delivery performance.

Directional Gradients and Surrogate Losses in the Predict-then-Optimize Framework – VISHAL GUPTA, USC MARSHALL SCHOOL OF BUSINESS – Michael Huang

We propose a novel family of decision-aware surrogate losses based on directional gradients for the predict-then-optimize framework. These losses directly approximate the downstream decision loss and can be optimized using off-the-shelf gradient-based methods. Importantly, unlike existing surrogate losses, the approximation error of our PG losses vanishes as the number of samples grows. This implies that optimizing our surrogate loss yields a best-in-class policy asymptotically, even in misspecified settings. This is the first such result in mis-

specified settings, and we provide numerical evidence confirming our PG losses substantially outperform existing proposals when the underlying model is misspecified, and the noise is not centrally symmetric. Insofar as misspecification is commonplace in practice – especially when we might prefer a simpler, more interpretable model – PG losses offer a novel, theoretically justified method for computationally tractable decision-aware learning.

Decoupling Learning and Decision-Making: Breaking the $O(\sqrt{T})$ Barrier in Online Resource Allocation with First-Order Methods – CHUNLIN SUN, STANFORD UNIVERSITY – Wenzhi Gao, Chenyu Xue, Dongdong Ge, Yinyu Ye

Online linear programming plays an important role in both revenue management and resource allocation, and recent research has focused on developing efficient first-order online learning algorithms. Despite the empirical success of first-order methods, they typically achieve a regret no better than $O(\sqrt{T})$, which is suboptimal compared to the $O(\log T)$ bound guaranteed by the state-of-the-art linear programming (LP)-based online algorithms. This paper establishes several important facts about online linear programming, which unveils the challenge for first-order-method-based online algorithms to achieve beyond $O(\sqrt{T})$ regret. To address the challenge, we introduce a new algorithmic framework that decouples learning from decision-making. More importantly, for the first time, we show that first-order methods can attain regret $O(T^{1/3})$ by this new framework. Lastly, we conduct numerical experiments to validate our theoretical findings.

Optimizer’s Information Criterion: Dissecting and Removing Bias in Data-Driven Optimization – TIANYU WANG, COLUMBIA UNIVERSITY – Henry Lam, Garud Iyengar

We develop a general bias correction approach, building on what we call Optimizer’s Information Criterion (OIC), that directly approximates the first-order bias and does not require solving any additional optimization problems. Our OIC helps evaluate the objective performance in data-driven optimization, which crucially involves not only model fitting but also its interplay with the downstream optimization. As such OIC can be used for decision selection instead of just model selection. We apply our approach to a range of data-driven optimization formulations comprising empirical and parametric models, their regularized counterparts, and furthermore contextual optimization. Finally, we provide numerical validation on the superior performance of our approach under synthetic and real-world datasets.

3e - FA234 - 514C - **Data-driven decision making in high-stakes domains: notes on fairness and interpretability** Chair: SINA AGHAEI ↩

Is There an Accuracy-Simplicity Tradeoff in High-Stakes Decision Domains? – LESIA SEMENOVA, DUKE UNIVERSITY –

Finding optimal, sparse, accurate models of various forms (such as linear models with integer coefficients, rule lists, and decision trees) is generally NP-hard. Often, we do not know whether the search for a simpler model will be worthwhile, and thus we do not undertake the effort to find one. This talk addresses an important practical question: for which types of datasets would

we expect interpretable models to perform as well as black-box models? I will present a mechanism of the data generation process, coupled with choices usually made by the analyst during the learning process, that leads to the existence of simpler-yet-accurate models. This mechanism indicates that such models exist in practice more often than one might expect.?

Learning Interpretable and Fair Housing Allocation Policy for Individuals Experiencing Homelessness in Los Angeles – SINA AGHAEI, UNIVERSITY OF SOUTHERN CALIFORNIA –

We consider the problem of learning an optimal, highly interpretable, and fair policy for allocating scarce housing resources to people experiencing homelessness in Los Angeles. Using administrative data from the Homeless Management Information System, we propose a way to create simple and actionable policies in the form of prescriptive trees that can satisfy arbitrary fairness constraints or other domain specific requirements that stakeholders care about. In the case of enforcing fairness in outcomes, our policies improve outcomes by up to 2.3% and enhance fairness by up to 36.57%. When enforcing fairness in allocation, we achieve up to a 2.3% better outcome and increase fairness by up to 100%, i.e., eliminating all the discrimination. This paper is based on an ongoing collaboration with the Los Angeles Homeless Services Authority (LAHSA).

The Measure and Mismeasure of Fairness – JOHANN GAEBLER, HARVARD – Sam Corbett-Davies, Hamed Nilforoshan, Ravi Shroff, Sharad Goel

The field of fair machine learning aims to ensure that decisions guided by algorithms are equitable. Over the last decade, several formal, mathematical definitions of fairness have gained prominence. Here we first assemble and categorize these definitions into two broad families: (1) those that constrain the effects of decisions on disparities; and (2) those that constrain the effects of legally protected characteristics, like race and gender, on decisions. We then show, analytically and empirically, that both families of definitions typically result in strongly Pareto dominated decision policies. For example, in the case of college admissions, adhering to popular formal conceptions of fairness would simultaneously result in lower student-body diversity and a less academically prepared class, relative to what one could achieve by explicitly tailoring admissions policies to achieve desired outcomes. In this sense, requiring that these fairness definitions hold can, perversely, harm the very groups they were designed to protect. In contrast to axiomatic notions of fairness, we argue that the equitable design of algorithms requires grappling with their context-specific consequences, akin to the equitable design of policy. We conclude by listing several open challenges in fair machine learning and offering strategies to ensure algorithms are better aligned with policy goals.

f-FERM: Robust and Fair Empirical Risk Minimization via f-divergences – SINA BAHARLOUEI, UNIVERSITY OF SOUTHERN CALIFORNIA –

Training and deploying machine learning models that meet fairness criteria for protected groups are fundamental in modern artificial intelligence. While numerous constraints and regularization terms have been proposed in the literature to promote fairness in machine learning tasks, most of these approaches are not amenable to stochastic optimization due to the complex and nonlinear structure of constraints and regularizers. Here, the term “stochastic” refers to the ability of the algorithm to

work with small mini-batches of data. Motivated by the limitation of existing literature, we present a unified stochastic optimization framework for fair empirical risk minimization based on f -divergence measures (f -FERM). The proposed stochastic algorithm enjoys theoretical convergence guarantees. In addition, our experiments demonstrate the superiority of fairness-accuracy tradeoffs offered by f -FERM for almost all batch sizes (ranging from full-batch to batch size of one). Moreover, we show that our framework can be extended to the case where there is a distribution shift from training to the test data. Our extension is based on a distributionally robust optimization reformulation of f -FERM objective under ℓ_p norms as uncertainty sets. Again, in this distributionally robust setting, f -FERM not only enjoys theoretical convergence guarantees but also outperforms other baselines in the literature in the tasks involving distribution shifts.

3e - FA259 - 512B - New models and applications of data-driven optimization Chair: HODA BDKHORI 

Robust Portfolio Optimization Meets Arbitrage Pricing Theory – MATEUS WAGA, PUC RIO – Davi Valladão, Alexandre Street

Robust portfolio optimization models are crucial in mitigating the impact of significant forecasting errors in expected asset returns. However, despite their significance, existing approaches often overlook a fundamental characteristic of financial markets: the absence of arbitrage opportunities. This paper presents a novel portfolio optimization model that integrates the classical mean-variance approach, the Fama and French Factor Model, and the Arbitrage Pricing Theory within a robust optimization framework. Our model utilizes return statistics to shape the uncertainty set boundaries but further enhances its representation by explicitly incorporating the no-arbitrage condition. As a result, we obtain a non-convex uncertainty set and face a trilevel optimization problem. To effectively address these challenges, we propose a cutting-plane algorithm. Our numerical experiments using different datasets and transaction cost levels demonstrate the consistent outperformance of our proposed model over benchmark models in terms of cumulative returns and risk-adjusted metrics.

Data-Driven Robust Sequential Search – CAGIN URU, DUKE UNIVERSITY – David Brown

Models of sequential search arise in a variety of problems, including organizations hiring new employees, firms developing novel technologies, and individuals seeking housing or other investment opportunities. The classical papers in the sequential search literature rely on the assumption that the decision maker (DM) fully knows (or has some prior belief about) the underlying distribution of alternative values. In many applications, however, the DM explores alternatives with limited (if any) prior information on the alternative value distribution. As a motivating example, we can consider a recent graduate with no experience searching for a job with no prior knowledge of the wage distribution in the market or a seller searching for the best price among several buyers without any access to the distribution of buyers' valuation or any historical data for guidance. In these settings, it is natural to assume that the alternative value distribution is not known in advance. Since the previously prescribed decision rules in the sequential search

literature no longer apply, it is necessary to develop new decision rules to facilitate search under limited information on the value distribution. Motivated by these considerations, we study a generalization of the classical Pandora's problem in which the alternative value distribution is unknown prior to the search. In each period, the DM can either explore an alternative at a cost to continue the search or select and acquire any previously explored alternative to end the search. The goal is to find a stopping rule that maximizes the worst-case ratio of the expected reward compared with an oracle with full knowledge of the alternative value distribution. We consider several variations of this problem. For each variation, we design simple policies that perform competitively with the oracle policy. Furthermore, we develop upper bounds on the relative performance of any feasible policy with respect to the oracle benchmark. The resulting upper bounds show that our policies provide nearly optimal performance.

Missing data in data-driven optimization – HODA BDKHORI, GEORGE MASON UNIVERSITY –

Missing data is a common issue for many practical data-driven stochastic programming problems. The state-of-the-art approaches first estimate the missing data values and then separately solve the corresponding stochastic programming. Accurate estimation of missing values is typically inaccessible as it requires enormous data and sophisticated statistical methods. Therefore, this talk proposes an integrated approach, a distributionally robust optimization (DRO) framework, that simultaneously tackles the missing data problem and data-driven stochastic optimization by hedging against the uncertainties of the missing values. We construct several classes of ambiguity sets for our DRO model utilizing incomplete data sets, maximum likelihood estimation method, and different metrics. We prove the statistical consistency and finite sample guarantees of the corresponding models and provide tractable reformulations of our model for different scenarios. We perform computational studies on the multi-item inventory control problem and portfolio optimization using synthetic and real-world data. We validate that our method outperforms the traditional estimate-then-optimized approaches.

Pivotal Sampling for Online Bayesian Matching – TRISTAN POLLNER, STANFORD UNIVERSITY – Mark Braverman, Mahsa Derakhshan, Amin Saberi, David Wajc

We study the polynomial-time approximability of the optimal online stochastic bipartite matching algorithm, initiated by Papadimitriou et al. (EC'21). Here, nodes on one side of the graph are given upfront, while at each timestep, an online node and its edge weights are drawn from a time-dependent distribution. The optimal algorithm is PSPACE-hard to approximate within some universal constant.

We provide a state-of-the-art 0.678-approximate algorithm. Building on our algorithms and the recent black-box reduction of Banihashem et al. (SODA'24), we provide polytime (pricing-based) truthful mechanisms which 0.678-approximate the social welfare of the optimal online allocation for bipartite matching markets. Our online allocation algorithm relies on the classic pivotal sampling algorithm (Srinivasan FOCS'01, Gandhi et al. J.ACM'06), along with careful discarding to obtain strong negative correlations between offline nodes. Our analysis requires new tail expectation bounds on the weighted

sum of negatively correlated Bernoulli variables.

3e - FB141 - 514C - **Advances in data-driven decision making** Chair: VELIBOR MISIC

Learning Optimal Classification Trees Robust to Distribution Shifts – NATHAN JUSTIN, UNIVERSITY OF SOUTHERN CALIFORNIA – Sina Aghaei, Andres Gomez, Phebe Vayanos

We consider the problem of learning classification trees that are robust to distribution shifts between training and testing/deployment data. This problem arises frequently in high stakes settings such as public health and social work where data is often collected using self-reported surveys which are highly sensitive to e.g., the framing of the questions, the time when and place where the survey is conducted, and the level of comfort the interviewee has in sharing information with the interviewer. We propose a method for learning optimal robust classification trees based on mixed-integer robust optimization technology. In particular, we demonstrate that the problem of learning an optimal robust tree can be cast as a single-stage mixed-integer robust optimization problem with a highly nonlinear and discontinuous objective. We reformulate this problem equivalently as a two-stage linear robust optimization problem for which we devise a tailored solution procedure based on constraint generation. We evaluate the performance of our approach on numerous publicly available datasets, and compare the performance to a regularized, non-robust optimal tree. We show an increase of up to 12.48% in worst-case accuracy and of up to 4.85% in average-case accuracy across several datasets and distribution shifts from using our robust solution in comparison to the non-robust one.

Randomized Nyström Preconditioned Interior Point-Proximal Method of Multipliers (IP-PMM) – YA CHI CHU, STANFORD UNIVERSITY – Luiz Rafael Santos, Madeleine Udell

We present a new algorithm for convex separable quadratic programming (QP) called NysPCG-IP-PMM, a regularized interior-point solver that uses low-rank structure to accelerate solution of the Newton system. The algorithm combines the interior point proximal method of multipliers (IP-PMM) with the randomized Nyström preconditioned conjugate gradient (NysPCG) method as the inner linear system solver. Our algorithm is matrix-free: it accesses the input matrices solely through matrix-vector products, as opposed to methods involving matrix factorization. It works particularly well for separable QP instances with dense constraint matrices. We establish convergence of NysPCG-IP-PMM. Numerical experiments demonstrate its superior performance in terms of wall clock time compared to previous matrix-free IPM-based approaches.

On the Equivalence and Performance of Distributionally Robust Optimization and Robust Satisficing Models in OM Applications – LONG HE, GEORGE WASHINGTON UNIVERSITY – Zhiyuan Wang, Lun Ran, Minglong Zhou

Distributionally robust optimization (DRO) has become ubiquitous to address uncertainties inherent in many operations management (OM) problems. Recently, an alternative goal-driven framework, robust satisficing (RS), is proposed. Robust satisficing aims to attain a prescribed target, such as avoid-

ing overshooting the cost budget as much as possible under uncertainty. The goal-driven modeling philosophy naturally fits many OM problems, yet there is a lack of direct comparisons between DRO and RS in OM applications. In this paper, we uncover connections between DRO and RS. Suppose both models are based on the Wasserstein metric and consider a risk-aware convex objective function affected by uncertain parameters. We demonstrate that they share the same solution family. We establish the correspondence between the radius parameter in DRO and the target parameter in RS such that the optimal solutions to the two models are the same. Inspired by the globalized distributionally robust counterpart (GDRC), we extend the analysis to GDRC and the globalized robust satisficing (GRS). We reveal that GDRC and GRS have the same solution families as DRO and RS, respectively, indicating that globalized models do not necessarily enhance model performance. More importantly, we establish novel results on the equivalence of DRO, GDRC, RS, and GRS models under previously stated conditions. To illustrate our findings, we conduct numerical studies on adaptive network lot-sizing, facility location, and portfolio management problems. Although the equivalence result holds, the performance of the DRO and RS models can vary depending on how the model parameters are selected. The experiment results show that different operations management problems exhibit different model preferences. Additionally, the experiments provide insights for practitioners, such as the use of cross-validation can reflect the actual model preference when data is sufficient.

Randomized policy optimization for high-dimensional optimal stopping – VELIBOR MISIC, UCLA –

Optimal stopping is the problem of determining when to stop a stochastic system in order to maximize reward, which is of practical importance in domains such as finance, operations management and healthcare.

The majority of existing approximate dynamic programming methods for high-dimensional optimal stopping practice produce deterministic linear policies – policies that deterministically stop based on the sign of a weighted sum of basis functions – but are not guaranteed to find the optimal policy within this policy class given a fixed basis function architecture. The most popular of such methods is least squares Monte Carlo (LSM), which recurses backwards in time from the last period to the first, using least squares regression to predict the current continuation value at each period. Such an approach is motivated by the fact that least squares regression recovers conditional expectations exactly when one carries out the regression over an unrestricted function class and with an infinite sample of trajectories, leading to the optimal policy. However, when one restricts to a finite sample and a function class defined as the span of a finite set of basis functions, one is not guaranteed to obtain the best performing policy within that function class. This observation mirrors recent developments in contextual stochastic optimization, where it has been observed that predictive models chosen to optimize predictive performance (e.g., squared error) often do not yield good prescriptive performance when they are used within a downstream optimization problem (the “predict-then-optimize” paradigm).

In this paper, we propose a new methodology for optimal stopping based on randomized linear policies, which choose to

stop with a probability that is determined by a weighted sum of basis functions. We motivate these policies by establishing that under mild conditions, given a fixed basis function architecture, optimizing over randomized linear policies is equivalent to optimizing over deterministic linear policies. We additionally motivate these policies by developing simple pathological instances in which LSM can be made to perform arbitrarily poorly relative to the optimal policy, while an exact optimization over deterministic/randomized policies either exactly recovers the optimal policy or otherwise perform arbitrarily closely to optimal.

We formulate the problem of learning randomized linear policies from data as a smooth non-convex sample average approximation (SAA) problem. We theoretically prove the almost sure convergence of our randomized policy SAA problem and establish bounds on the out-of-sample performance of randomized policies obtained from our SAA problem based on Rademacher complexity. While the SAA problem is in general NP-Hard, we develop a novel alternating maximization approach for solving it that is structurally distinct from the backward recursion technique that is typical of classical approaches such as LSM.

Through numerical experiments on a benchmark family of option pricing problem instances, we show that our approach can substantially outperform state-of-the-art methods.

4a - TB943 - 512G - **Networks and Telecom 1**

Chair: CESAR A URIBE

Budgeted spanning trees and polyhedra – CHARLES NOURRY, LAMSADE UNIVERSITÉ PARIS DAUPHINE – Ali Ridha Mahjoub, Hassan Aissi

The minimum budgeted spanning tree problem is a problem situated at the intersection of the minimum spanning tree and the multi-knapsack problems. The minimum budgeted spanning tree problem is central in telecommunication and network design areas, where the knapsack constraints may represent business constraints. In order to effectively solve the problem, we investigated a polyhedral analysis of the budgeted spanning tree problem. We provide several exact extended formulations for special cases of the problem. In the special case where exactly one knapsack constraint is involved, our approach builds upon a combination of matroid intersection and an application of the disjunctive programming. This extended formulation allows us to gain some insights on the structure of the problem. Moreover, by using the projection technique, the formulation could be exploited to generate valid inequalities that can be integrated in a Branch and Cut procedure.

On Graphs with Finite-Time Consensus and Their Use in Gradient Tracking – CESAR A URIBE, RICE UNIVERSITY – Edward Duc Hien Nguyen, Xin Jiang, Bicheng Ying

This paper studies sequences of graphs satisfying the finite-time consensus property (i.e., iterating through such a finite sequence is equivalent to performing global or exact averaging) and their use in Gradient Tracking. We provide an explicit weight matrix representation of the studied sequences and prove their finite-time consensus property. Moreover, we incorporate the studied finite-time consensus topologies into Gradient Tracking and present a new algorithmic scheme called Gradient Tracking for Finite-Time Consensus Topologies (GT-

FT). We analyze the new scheme for nonconvex problems with stochastic gradient estimates. Our analysis shows that the convergence rate of GT-FT does not depend on the heterogeneity of the agents' functions or the connectivity of any individual graph in the topology sequence. Furthermore, owing to the sparsity of the graphs, GT-FT requires lower communication costs than Gradient Tracking using the static counterpart of the topology sequence.

4a - WB944 - 512G - **Networks and Telecom 2**

Chair: SACHIN JAYASWAL

Online Convex Optimization for On-Board Routing in High-Throughput Satellites –

OLIVIER BÉLANGER, POLYTECHNIQUE MONTRÉAL – Jean Luc Lupien, Olfa Ben Yahia, Stéphane Martel, Antoine Lesage Landry, Gunes Karabulut Kurt

In recent years, there has been a critical rise in the utilization of low Earth orbit (LEO) satellite internet connectivity services. In certain regions, the growing demand outpaces the available data rates, thereby compromising future quality of service. While deploying more satellites is an acceptable short-term solution, a more durable alternative involves the design of higher-performance satellites with higher transmission capabilities. These next-generation payloads will require a very high capacity only achievable by inter-connecting multiple onboard processors to scale throughput significantly, in comparison to the traditional single-processor satellites. However, there is a lack of current research on the internal packet routing of high-throughput satellites between these onboard processors. To tackle this challenge, after modelling the internal satellite architecture, we design a real-time optimal flow allocation and priority queue scheduling method using online convex optimization-based model predictive control. We formulate the problem as a multi-commodity flow instance, where incoming data streams with distinct priorities are the commodities, the uplink beam is the source, and both the downlink beam and discarded packets are sinks. We use an online interior-point method to iteratively solve the routing and scheduling optimization problem, accounting for time-varying equality constraints as information about incoming flow is observed. Then, we feed back its decision in the satellite model. Our approach minimizes packet loss cost and facilitates real-time rerouting in response to incoming information and exogenous uncertainty. Our online convex optimization method provides low computational overhead and is implementation-ready. Our method is tested in a next-generation high-throughput satellite model and is compared to a reference batch optimization model and traditional methods such as cost-proportional allocation.

Application placement simulations over HEAVEN platform – ANTOINE BERNARD, POLYTECHNIQUE MONTRÉAL – Antoine Legrain

In recent years, deploying distributed applications efficiently over cloud platforms has been extensively studied. Various algorithms were developed to solve this optimization problem, such as virtual network embedding algorithms.

The HEAVEN platform is an embedded wireless datacenter solution that provides resilient and flexible internet coverage. This solution is an efficient way to provide emergency connectivity through to its mesh network. The next logical step after establishing connectivity is to provide processing

power to the users. Exploiting drones' processing power can be a method for offloading computations, providing essential offline emergency services. With limited connectivity between devices, link allocation becomes a key component of the problem.

This talk presents a comprehensive analysis of the capabilities such platform could offer, by presenting a simulation based on individual device capability and connectivity. The simulation framework incorporates models for device capability, connectivity, and application requirements, enabling the evaluation of efficient application deployment strategies over HEAVEN. Optimizing application placement in real-time and in a distributed way would expand the network resiliency to its applications.

Additionally, this talk presents preliminary results on implementing various placement algorithm to the simulation framework, following the work presented in a previous conference, aiming to understand and solve this distributed virtual application placement problem under heavy network and resource constraints.

Fiber-to-the-Home Passive Optical Access Network Design: A New Formulation and Valid Inequalities – SACHIN JAYASWAL, INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD – Yogesh Kumar Agarwal

We study the problem of the optimal design of fiber-to-the-home optical access networks. Given a network of nodes and edges with a given demand for optical fibers at a subset of these nodes, the problem entails finding the optimal placement of splitters, which allows multiple demand points to share a common fiber between the central office and a splitter, such that the cost of fiber cables is reduced without incurring too much of splitter cost. Additionally, it needs to decide on the optimal selection of a cable type of appropriate capacity on each edge of the network to carry the required traffic. For this, we provide a mixed-integer programming (MIP) formulation of the problem and propose several valid inequalities (VIs), with or without a pre-specified template, to strengthen the formulation. Through our computational experiments, we demonstrate the efficacy of our proposed VIs, which help improve the lower bound of the problem from 79% to 86.9% of the MIP optimal cost, on average. For the special cases of the problem studied in the literature, we show that our formulation produces much tighter lower bounds. On top of that, our proposed VIs are comparatively much more effective in tightening the bounds. Specifically, our proposed formulation with our VIs consistently outperforms that available in the literature, being as much as 500 times faster in some instances.

4a - WC212 - 514A - **Telecommunication Network Optimization** Chair: DIMITRI PAPADIMITRIOU

Robust optimization for the Segment Routing Traffic Engineering Problem – BERNARD FORTZ, HEC LIÈGE MANAGEMENT SCHOOL OF THE UNIVERSITY OF LIÈGE – Hugo Callebaut, Jérôme De Boeck

Segment routing, a modern protocol enhancing traffic engineering, introduces flexibility by enabling traffic to take detours.

This talk tackles the challenge of optimizing routing in the face of uncertain traffic distribution. Unlike traditional methods relying on a single traffic matrix, our approach considers an infinite set of matrices defined by linear constraints. Our goal is to optimize routing under the worst-case scenario within this set. Through innovative formulations, our results showcase a substantial speed improvement in optimization compared to traditional methods, offering an efficient alternative to exploring all extreme points in the matrix set.

Joint Optimization of inter- and intra-frequency events in Multi-Carrier Scenarios with UE Mobility in Radio Networks – PEGAH ALIZADEH, ERICSSON RESEARCH CENTER PARIS – Anastasios Giovanidis, Koutsoukis Vasileios, Ingemar Johansson S

This work presents a novel approach for jointly optimizing inter- and intra-frequency events in multi-carrier scenarios with user equipment (UE) mobility for radio networks. We propose leveraging sequential learning techniques to learn the optimal group of continuous global handovers (HOs). Our method includes performance evaluation for UE traffic mix with various Quality-of-Experience (QoE) requirements and different load/mobility patterns. Furthermore, we select an appropriate global multi-objective function to optimize performance metrics such as the number of HOs, ping-pong effect, inter-carrier load, etc for a given input traffic mix.

Stochastic Lagrangian-based method and application to multi-constrained network design – DIMITRI PAPADIMITRIOU, ULB –

The Augmented Lagrangian Method (ALM) is one of the most common approaches for solving linear and nonlinear constrained problems. However, for nonconvex objectives, the handling of nonlinear inequality constraints remains challenging. In this paper, we propose a stochastic ALM with Backtracking Line Search that performs on a subset (mini-batch) of randomly selected points for the solving of nonconvex problems. The considered class of problems include both nonlinear equality and inequality constraints. Together with the formal proof of the convergence properties (in expectation) of the proposed algorithm and its computational complexity, the performance of the proposed algorithm are then numerically compared against both exact and inexact state-of-the-art ALM methods. Further, we apply the proposed stochastic ALM method to solve a multi-constrained network design problem. We perform extensive numerical executions on a set of instances extracted from SNDlib to study its behavior and performance as well as potential improvement of this method. Analysis and comparison of the results against those obtained by extending to nonlinear constraints methods developed for the approximation of separable nonconvex optimization programs are then provided.

4b - MB190 - 511E - **Network Design for Transportation Planning II** Chair: TEODOR GABRIEL CRAINIC

The Intermodal Railroad Blocking and Railcar Fleet-Management Planning Problem – JULIE KIENZLE, UNIVERSITY OF MONTREAL – Serge Bisailon, Teodor Gabriel Crainic, Emma Frejinger
Railroads constitute an important component of the North

American intermodal transportation network as they move containers long distances inland. They face difficult large-scale decision-making problems, and in this work, we introduce such a problem: the tactical Intermodal Blocking and Railcar Fleet-Management problem. Taking train schedule and related capacity as an input, we aim to build a cost-effective plan that satisfies demand expressed as numbers of containers of different types to move. It corresponds to a service network design problem with fleet management decisions: the design decisions are associated with the selection of extra train services, and the selection of so-called railcar blocks. In addition, we manage a heterogeneous fleet of railcars that need to be matched to the demand characteristics. We formulate the problem as a mixed-integer linear program, and it cannot be solved within a ten-hour time limit with a general-purpose solver. We propose an exact solution approach based on iterative fixing of variables which allows us to solve real large-scale instances from our industrial partner, the Canadian National Railway company, that is one of the largest railroads in North America. The results show that solving this problem can lead to important gains in terms of cost-reduction and increased demand satisfaction.

Multi-layer network design – TEODOR GABRIEL CRAINIC, CIRRELT AND UNIVERSITÉ DU QUÉBEC À MONTRÉAL –

Multilayered networks are used to model systems with multiple interacting components, each being represented by a network layer. More than two layers may be involved. The “multi-layer network” term is generally associated with complex multilayered problem settings, frequently encountered in planning transportation and telecommunication systems, where an arc in a given layer is defined with respect to a set of arcs in another layer. Multi-layer network design simultaneously designs all layer-specific networks to satisfy at minimum overall cost a given set of origin-destination demands. Extending the already-difficult network design problem class, multi-layer network design presents additional modelling and algorithmic challenges arising from the connectivity relations and requirements characterizing the design, flow, and attribute decision variables associated with the arcs of interconnected layers. We recall and enhance the basic definitions and formulations, and introduce new ones for richer multi-layer networks, with more than two layers and connectivity relations involving several layers simultaneously. We illustrate these concepts through multi-layer service network design models for tactical planning of consolidation-based freight transportation carriers.

The Integrated Multi-Tier Hierarchical Hub Location and Scheduled Service Network Design Problem – AMIRMOHAMMAD FATHOLLAHI-FARD, UNIVERSITY OF QUEBEC IN MONTREAL –

During this seminar, we embark on an in-depth discussion of the freight transportation system, emphasizing a multi-tier hierarchical hub network integrating air, rail, and highway transportation modes. This structure encompasses primary hubs, secondary hubs, and service nodes interconnected to meet time-sensitive demands. Our main focus lies in developing an optimization model for the integrated hub location and scheduled service network design problem, catering to strategic and tactical planning levels. Additionally, we introduce a novel decomposition-based metaheuristic algorithm to effectively address the complexities of this integrated system. Through this

method, we aim to break down the original problem into manageable subproblems, thereby enhancing efficiency and solution quality. For each subproblem, we do exploration and exploitation phases to find near-optimal solutions. Our results empower decision-makers to optimize hub network design and allocate capacity configurations, thereby enhancing service efficiency and ensuring cost-effectiveness.

Profit maximizing network design problem with location decisions considering incompatible commodities and binary demand selection – EHSAN MIRZAEI, UNIVERSITÉ DE MONTRÉAL ET TARBIAT MODARES UNIVERSITY (IRAN) – Teodor Gabriel Crainic, Walter Rei, Ehsan Nikbakhsh

This study focuses on a profit maximization variant of the multi-commodity network design problem with location decisions and incompatible commodities. A novelty of the problem is the restriction of commodity flows for safety and security reasons, as commodities have a predetermined list of other commodities that cannot be shipped together. Therefore, decisions regarding serving commodities should be made simultaneously with design and flow decisions as demand selection is affected by them. The problem, therefore, is to design a network by selecting both nodes and arcs that maximizes the total profit including the revenue obtained from selecting demands, design cost (for the selection of nodes and arcs), and the flow costs to transport the commodities. In this research, we assume that commodities are transferred through transshipment locations, and direct shipment between the origin and destination of a commodity is not allowed. The shipment between transshipment nodes is allowed to use several arcs. Finally, an extensive set of computational experiments is conducted on a set of instances from the literature.

4b - MC202 - 516E - **Dynamic Facility and Network Design** Chair: SANJAY DOMINIK JENA

Solving a stochastic multi-horizon facility location problem with capacity expansion – PETER SCHÜTZ, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY – Sarka Stadlerova

We study the problem of locating production facilities and selecting production technology for hydrogen in Norway. Hydrogen demand is expected to increase over the next years but is also the main source of uncertainty. Further, hydrogen production costs highly depend on electricity prices, which are also uncertain. We present a stochastic multi-stage multi-horizon formulation, which enables the modelling of strategic as well as operational uncertainty. The strategic uncertainty is related to increasing but uncertain future demand, while the operational uncertainty is related to uncertain future electricity costs directly affecting the operational costs. We consider two production technologies, which differ in production flexibility and costs, and allow for their combination at one location. Therefore, capacity expansion is modelled as the opening of a new facility. The objective is to minimize the expected sum of investment, production and distribution costs while satisfying customer demand. We use Lagrangian relaxation to solve our problem. The lower bound is calculated by solving the linear relaxation of the Lagrangian subproblem. We implement a heuristic based on a restricted MIP approach to obtain a fea-

sible solution. The results show that our solution method can find high-quality solutions for large problem instances with up to 150 operational scenarios.

A slope scaling heuristic for the multiperiod strategic planning of carbon capture and sequestration value chains – GABRIEL HOMSI, CANMETENERGY – étienne Ayotte Sauv , Sanjay Dominik Jena

Around the world, various decarbonization strategies are being evaluated to support countries meeting their net-zero emissions objectives. One of these decarbonization strategies is carbon capture and sequestration (CCS). In CCS, CO₂ is captured at emitter sites (e.g. industrial plants) and transported to geological sequestration sites (e.g. saline aquifers, depleted oil fields), where it is injected underground for long-term storage (10,000+ years). Recent studies indicate that without CCS, countries may not be able to meet their net-zero carbon emissions objectives.

The deployment of CCS involves billions of dollars of investment and requires planning ahead for decades. For this reason, in this presentation, we focus on the multiperiod strategic planning of a pipeline-based CCS value chain. This problem combines characteristics of two classical problems in Operations Research: facility location and network design. Facility location decisions are related to the activation and operation of CO₂ capture units at emitter sites, and the opening and operation of geological reservoirs. Network design characteristics consist of the activation and operation of the pipeline network.

To account for multiple potential sources of uncertainty, this problem may have to be solved thousands of times. Therefore, reaching high-quality solutions in a few minutes is important. Computational experiments show that commercial solvers struggle to attain these requirements. Adding further problem attributes such as multiple CO₂ transportation modes may exacerbate this issue. To address this computational challenge, a novel slope scaling heuristic was previously introduced by Homs , Ayotte-Sauv , and Jena. This heuristic is based on existing work on single-period CCS problems and network design problems. It approximates the cost of design variables, has long-term memory search strategies, generates upper bounds with dynamic programming, and has a final improving phase where a restricted model is solved iteratively.

In this presentation, we provide updated results for the slope scaling heuristic along with additional performance insights. These new results show that the heuristic generates better solutions than CPLEX for most (58%) experiments (average relative improvement of 10%), at a fraction (10%) of the CPU time. For the same fraction of CPU time, results also show the satisfactory performance of the heuristic when it underperforms CPLEX: the vast majority of SS solutions (95%) have a relative cost that is at most 1% larger than CPLEX.

Cumulative Customer Demand in Facility Location and Network Design – SANJAY DOMINIK JENA, UNIVERSIT  DU QU BEC A MONTR AL – Warley Almeida, Mohammad Daneshvar, Margarida Carvalho, Walter Rei

Dynamic Facility Location and Network Design are a popular class of combinatorial optimization problems that provide the infrastructure necessary to satisfy customer demand. They

have been applied to a large variety of planning contexts, where customer demand may have different interpretations, such as consumer goods, manufacturing components, transportation services, or medical relief items. Existing literature commonly assumes that customer demand quantities are defined independently for each time-period. In many planning contexts, however, unmet demand carries over to future time periods. Demand that is unmet for some time periods may therefore affect decisions of subsequent time periods.

In this talk, we discuss the implications of cumulative customer demand in two representative planning contexts for multi-period network design and facility location: humanitarian supply chains, where unmet demand for relief aid may result in a spread of disease, and therefore increases future demand; and temporary facility location, such as pop-up stores, where customer demand for necessary items gradually builds up until it is satisfied by a nearby facility. We discuss two principal types of cumulative demand propagation and their respective modeling in mathematical programming formulations. We show how a failure of modeling such behavior may result in severe economic underperformance, how the corresponding models can be modeled and solved efficiently, and how different problem characteristics impact the computational complexity.

4b - TA192 - 516E - Network Design for Transportation Planning III Chair: WALTER REI [↔](#)

Design for integrated Freight and Passenger Transportation: Some recent advances – TOM VAN WOENSEL, EINDHOVEN UNIVERSITY OF TECHNOLOGY –

Freight-on-Transit introduces the innovative approach of blending freight with passenger services in city transport systems. It encourages freight companies to adopt greener transport options through government policies. Companies can opt for direct (via trucks) or combined (using trucks and trains) methods depending on cost considerations in this mixed transport model. The decision involves selecting road transport or a mix of road and rail options for moving goods. This study presents several groundbreaking developments we are exploring, including the simultaneous planning and scheduling for freight and passenger rail services. We evaluate the impact of policies like fuel taxes and subsidies for trains on making freight transport more environmentally friendly. The issue is framed as a Stackelberg game, with a leader-follower dynamic between the regulatory body and the freight company. The authority imposes a fuel tax for road usage and offers subsidies for the combined use of freight and passenger rail services. In response, the freight company, as the follower, organizes shipments in the most economically beneficial way. "Direct" shipments are done entirely by road, whereas "indirect" ones incorporate rail services into the logistics chain.

A Construction Matheuristic for Two-Tier Synchronized City Logistics – JULIA LANGE, UNIVERSITY OF KAISERSLAUTERN LANDAU – Teodor Gabriel Crainic, Timo Gschwind, Walter Rei

Sustainable city logistics planning focuses on multi-tier and multi-modal transportation with efficient consolidation and vehicle types suited for each tier. Integrating logistics ser-

vices of different providers and shared transportation of multi-directional demand flows are key strategies to reduce congestion. This work introduces a detailed mathematical description of a day-before planning problem in two-tier multi-modal city logistics with on-time synchronization, where no storage exists at handover locations. While this enables the use of existing resources like supermarket parking lots in the distribution process, the requirement of delivery vehicles to meet for synchronized activities is a challenge. The planning approach is based on two-tier scheduled service network design, where transportation services with routes, departure time windows, and capacities are given, and waiting time policies exist for customer and handover locations. Demands involve inbound, outbound, and inner-city commodity flows. The goal is to select services, including a schedule for each of them, and allocate the demands such that both operating costs and waiting times are minimized. We present a two-step construction matheuristic, where the search space is reduced by fixing a subset of binary selection variables for services based on solutions obtained for a relaxation of the model. Since the adherence of capacity restrictions of handover locations and the precise determination of vehicle waiting times constitute the main drivers of model complexity, these aspects are simplified or neglected thereby. In a broad computational study, the simple construction matheuristic validates the accurate representation of structural and temporal requirements and demonstrates the applicability of the model to obtain near-optimal solutions for medium-size instances. Further, we explore computational effects of varying instance parameters and discuss dependencies within the solution's structure. Interesting methodological and managerial insights are drawn and conclusions on future research directions are provided.

A multi – commodity location – network design problem with vehicle selection in City Logistics – FRANCESCO CONTU, UNIVERSITY OF CAGLIARI –

This work aims to study and develop a decision-support system for planning the distribution of freight packed in containers arriving from the sea to cities built around a port. The problem is approached from the perspective of an urban mobility manager who aims to create a strategic plan for delivering freight from the port to customers located within the city, while managing existing transportation resources represented by firms located on the outskirts of the city. Containers cannot be opened at the port and cannot reach destinations within the city due to local regulations; they must be unpacked at intermodal facilities, and the freight must be delivered to customers by a set of city-freighters. A location-network design problem is examined to achieve this goal: a set of containers carrying pallets needs transportation from the port to facility locations, where pallets are unpacked from the containers, loaded onto vehicles allowed to transit on city streets, and dispatched to their final destinations. This problem involves selecting facilities and vehicles, assigning containers to facilities, determining paths for chosen vehicles, and managing pallet flows. An ALNS-based metaheuristic has been proposed; an initial solution is built and modified at each iteration of the algorithm: a destroy operator is selected to de-assign a certain number of containers from their corresponding satellites. These containers are then reassigned to any open satellite by a repair operator. Some proposed destroy operators remove containers after an explicit

opening or closing of a satellite. At each iteration of the algorithm, after the container assignments to satellites have been changed, a multi-depot vehicle routing problem is solved to determine the routes of the city-freighters that will deliver the pallets to customers. The proposed ALNS meta-heuristic algorithm is compared to a heuristic algorithm which is based on decomposition by satellite.

The Two-Echelon Multicommodity Location-Routing Problem with Stochastic and Correlated Demands – WALTER REI, CIRRELT AND UNIVERSITÉ DU QUÉBEC À MONTRÉAL –

David Escobar Vargas, Teodor Gabriel Crainic, Stein W. Wallace

In this presentation, we introduce the stochastic two-echelon multicommodity location routing problem with stochastic and correlated demands. We propose a two-stage stochastic programming formulation where the first stage involves decisions on the design of second-echelon facilities, and the second stage consists of recourse decisions addressing the distribution of observed demands. The objective is to minimize the combined costs of first-stage design decisions and the expected total routing costs incurred in the second stage. To solve the resulting stochastic optimization model, we develop a progressive hedging metaheuristic enhanced with algorithmic improvements to accelerate the solution space exploration. These enhancements include: 1) the use of population structures to derive diverse solutions for the scenario subproblems solved during the search process; 2) alternative strategies for defining reference solutions that efficiently guide the solution-consensus search; and 3) a reset procedure to mitigate the risk of the method becoming trapped in local optima. We validate the efficiency and effectiveness of our strategies through extensive computational experiments, assessing their ability to generate high-quality solutions under various problem settings and demand correlation settings.

4b - TA274 - 512G - Large Scale Optimisation in Transportation Chair: FRANCOIS SOUMIS ←

A Branch-and-Price Algorithm for Multiple Depot Bus scheduling Problem in Public Transit Systems – NADIA RASOULI, PHD CANDIDATE – Guy Desautniers, Mohammed Saddoune, Francois Soumis

Bus scheduling problem is one of the most important scheduling problems in public transit since it guarantees minimum cost running for the transportation companies. A bus schedule defines the number of trips made by a bus in a workday, from the time it leaves the depot until it returns to the depot. Our study focuses on a large-scale multiple-depot bus scheduling problem in the public bus transportation industry with homogeneous fleets of vehicles. Using a timetable of trips, this research targets finding feasible bus schedules that cover each trip exactly once. With this approach, we use a set partitioning model with side constraints, and a connection-based network. Our solution approach is based on column generation (CG) embedded in a branch and bound method. As part of CG, we use a shortest path subproblem with resource constraints, which is then solved with a labeling algorithm considering dominance rules. To speed up the solution process, we use some heuristics branching including variable fixing and inter-task fixing. Despite the complexity of our model compared to most of those in literature reviews, computational experiments show that our

approach is fast.

Considering regularity in the aircrew pairing problem – FRÉDÉRIC QUESNEL, UNIVERSITÉ DU QUÉBEC À MONTRÉAL – Mohamed Faouzi Benam-mour, Francois Soumis

The Crew Pairing Problem (CPP) seeks to find a set of feasible crew pairings (sequences of flight connections, and rests, forming one or several days of work for a crew) at minimum cost that cover all flights in a given period. However, some airlines are willing to sacrifice a portion of their savings to achieve regularity. This is because regular schedules are much easier to implement and manage, including booking the same hotels, rooms, and limousines for the crew. It also enhances employee satisfaction. Regularity is achieved by having pairing patterns that are repeated several times in the CPP solution. The CPP is typically solved using a branch-and-price algorithm. However, existing mathematical models for the CPP with regularity are incompatible with branch-and-price. In this study, we propose a two-step solution method to solve the CPP with regularity. The first step consists of solving an auxiliary optimization problem called aggregated daily CPP (ADCPP). ADCPP identifies a set of highly regular, yet efficient pairings. Solving the ADCPP requires a new branch-and-bound algorithm. In the second step, the pairings created by ADCPP are given as input to a traditional branch-and-price solver. Bonuses are given to those pairings to incentivize selecting them in the final CPP solution. Preliminary results on large-scale commercial instances of the CPP are presented.

Accelerated insight into the crew rostering problem with machine learning methods – PHILIPPE RACETTE, ÉCOLE POLYTECHNIQUE DE MONTRÉAL –

The crew rostering problem (CRP) is a complex crew scheduling task where pairings, or sequences of flights starting and ending at the same airport, are assigned to pilots. These assignments form a collection of schedules (called roster) spanning one month. The CRP typically seeks to optimize some criterion such as cost or crew satisfaction.

In this presentation, we describe a new line of research concerning the CRP. We address the possibility of generating very fast solutions to the problem with the help of machine learning (ML) techniques. Using a sequential assignment procedure whose arc weights are computed with ML weights, we generate solutions to infer about the quality of a CRP instance's configuration. We also use a partial re-optimization technique with windowing to improve the solutions. The rosters obtained in this way are high-quality and produced in 5% to 20% of the time normally taken by a state-of-the-art solver. In each case, the ML weights are learned with reinforcement learning and an evolutionary algorithm (CMA-ES).

We conclude with a discussion of the methodological framework introduced and outline future efforts, such as extensions of our method to other crew scheduling tasks, and its relationship with different ML models, such as GNNs.

Solving the Crew Split Problem: Innovative Team Formation Strategies for the Crew Pairing Problem – MOHAND RYAD CHELBANI, POLYTECHNIQUE MONTRÉAL – Frédéric Quesnel, Francois Soumis

Airline companies are constantly looking at optimizing their resources by improving their decision-making capabilities. In this context, crew scheduling significantly influences cost-saving decision-making processes. It is advantageous for airlines when their cabin crews operate in teams throughout entire pairings, which are sequences of flights, connections, and layovers, covering one or more workdays. The benefits of this approach are numerous: better cohesion between cabin crews, which increases effectiveness and boosts morale; enhanced logistical operations involved; reduced spread of delays throughout the pairings; and improved robustness of the schedules. However, maintaining the same team for consecutive flights isn't always feasible due to the different crew requirements of the various types of aircraft operated by companies in a single pairing. In these instances, using a core team to fulfill the majority of roles, supplemented by additional crew members for specific roles, can be a good strategy to use.

In our presentation, we will introduce the Crew Split Problem (CSP), aiming at determining the optimal composition of a core team and the complementary crew members following it for each flight. The CSP is a multi-objective optimization problem that, among others, seeks to minimize the cost associated with repositioning crew members (or deadhead). Our model also considers the possibility for higher-ranked crew members to assume roles below their usual rank; this concept is called downranking. The first promising results are shown in real-world instances.

4b - TB166 - 513D - Network Design for Transportation Planning I Chair: TEODOR GABRIEL CRAINIC

Benders Decomposition for Two-layer Network Design with Capacity Decisions – ALI ROUHANI, DÉPARTEMENT D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE, UNIVERSITÉ DE MONTRÉAL, MONTREAL, QC, CANADA – Teodor Gabriel Crainic, Walter Rei, Walid Klibi

Two-layer network design is a subclass of network design problems that simultaneously addresses two interwoven sets of design decisions, each pertaining to a different network layer. We focus on the Two-Layer Network Design Problem with Capacity Decisions (2LNDCD), where two network layers are interconnected through design, flow, and capacity requirements. The objective is to design a cost-efficient two-layer network that meets multicommodity demand and capacity restrictions while adhering to standard network design constraints and specific inter-layer coupling requirements. The 2LNDCD problem introduces additional modelling and algorithmic challenges due to its complex connectivity relations and requirements. In this presentation, we provide the fundamental definitions and formulation of 2LNDCD and we propose a Benders decomposition method to solve the resulting optimization model. Preliminary results are discussed to highlight the intricacies associated with the application of this solution approach in the given context.

From Tactical to Operational Planning: Implementation Policies for Consolidation-based Freight Transportation Systems – SAEDEH DEHGHANI, UQAM – Teodor Gabriel Crainic, Walter Rei

The objective of this research is to narrow the divide between theoretical frameworks regarding tactical planning within service network design models and their practical applications at the operational level, particularly in response to emerging information. While the concept of tactical planning is well-established within consolidation-based freight transportation systems, its adaptability in operational contexts remains insufficiently explored. Through the systematic definition of various scenarios and strategies, this study seeks to furnish operations managers with the expertise necessary to effectively tackle unforeseen operational hurdles, ultimately aiming to improve network efficiency and resilience. The findings of this research encompass refined decision-making procedures, enhanced operational efficacy, and a solid plan for setting tactical planning choices into action at the operational level.

Service Network Design with Uncertainty on Water Levels for Intermodal River Transport – BITA PAYAMI SHABESTARI, PHD STUDENT – Teodor Gabriel Crainic, Walter Rei, Ioana C Bilegan

Inland waterway transport presents a sustainable alternative to road transport, offering the potential to alleviate congestion and reduce costs. However, increasingly frequent and severe drought seasons, attributed to climate change, challenge the resilience of this mode of transport. Decreased water levels restrict navigation, impact vessel size, and reduce vessel capacity. This study introduces a comprehensive modeling framework for tactical planning in consolidation-based freight transportation on inland waterways. By examining the relationship between water levels and vessel load capacity, with a specific focus on vessel dimensions, the framework aims to evaluate possible impacts on the efficiency of inland waterway transport. We also introduce three Stochastic Scheduled Service Network Design with Resource and Revenue Management models to tackle water-level uncertainties, each addressing a distinct tactical plan adjustment strategy. These models aim to establish a tactical plan, given predicted water levels, that maximizes the expected carrier's revenue while accounting for future adjustments to the plan when information is revealed and predictions are reliably updated, to fulfill the demands of shippers and optimize the utilization of the carrier's resources. Through extensive experimentation using commercial software and a novel decision-based scenario clustering algorithm, we assess the quality of the solutions obtained with the different model variants and analyze the impact of the stochastic parameters on the results.

Tactical Planning in Multi-Stakeholder Freight Transportation Systems under Uncertainty – NADIA TAHMASBI, CIRRELT AND UNIVERSITÉ DU QUÉBEC À MONTRÉAL – Teodor Gabriel Crainic, Walter Rei, Janosch Ortmann

A two-sided digital platform for logistics service providers is defined as an integrated multi-stakeholder freight transportation system where the demands of many shippers (e.g., retailers, distributors, and manufacturers) are matched to the capacity offers from many carriers (e.g., freight carriers, terminal managers, and other logistics service providers) using a centralized intelligent decision-support entity. From a system design perspective, this platform has a Many-to-One-to-Many structure, which we refer to as an M1M system. The intelligent platform efficiently matches the demand and supply sides, ensuring timely satisfaction of demand while respecting oper-

ational restrictions and generating profit for the platform. We investigate the tactical planning required for M1M systems under uncertainty in demand volume and service capacity offers. A tactical plan involves selecting and matching requests and offers, and consolidating them in time and space to maximize the platform's profit.

4b - TB945 - 511F - **Transportation and logistics** 1 Chair: JEAN FRANÇOIS CÔTÉ

A Comprehensive Stochastic Programming Model for Transfer Synchronization in Transit Networks – ZAHRA ANSARILARI, UNIVERSITY OF TORONTO ALUMNA – Merve Bodur, Amer Shalaby

To realize the benefits of network connectivity in transfer-based transit networks, it is critical to minimize transfer disutility for passengers by synchronizing timetables of intersecting routes. Nonetheless, in real-world practice, transit agencies commonly build timetables without accounting adequately for stochastic operations and with slight or no consideration of synchronization among intersecting routes at transfer nodes. Given the significance of transfers for many transit users, it is imperative to develop seamlessly synchronized timetables that account for the stochastic nature of transit operations. In this research, we investigate the stochastic transfer synchronization problem, which seeks to synchronize the timetables of different routes in a transit network to reduce transfer waiting times, delay times, and unnecessary in-vehicle times. We present a sophisticated two-stage stochastic mixed-integer programming model that takes into account variability in passenger walking times between bus stops, bus running times, dwell times, and demand uncertainty. Our model incorporates new features related to dwell time determination by considering passenger arrival patterns at bus stops which have been neglected in the literature on transfer synchronization and timetabling. We solve a sample average approximation of our model using a problem-based scenario reduction approach, and the progressive hedging algorithm. As a proof of concept, our computational experiments on two single transfer nodes in the City of Toronto, with a mixture of low- and high-frequency routes, demonstrate the potential advantages of the proposed model. Our findings highlight the necessity and value of incorporating stochasticity in transfer-based timetabling models.

Flexible and modular energy systems modelling with JuMP; A case study from the Arctic – LARS HELLEMO, SINTEF – Sigrid Aunsmo, Saket Adhau, Brage Rugstad Knudsen

Radical cuts in emissions typically involves the integration of multiple energy carriers in the energy system, putting additional strain on energy system modelling tools and practitioners. The modelling framework JuMP developed for mathematical optimization in Julia offers great speed and flexibility in combining multiple modules/packages.

EnergyModelsX is a framework for energy systems modelling that is designed to be easy to extend, e.g. with more detailed technology descriptions for certain analyses. PiecewiseAffineApprox is a package to simplify finding good linear approximations from e.g. more detailed dynamical simulation models which may be used to define MILP formulations from non-linear models.

We illustrate how EnergyModelsX and PiecewiseAffineApprox can be used together on a use case from the Arctic. The first use case in the ZEESA project concerns reducing emissions at Isfjord Radio at Svalbard, a remote location in the Arctic. The energy system there includes existing renewable energy from solar, and potentially generation from wind turbines. The model includes use of surplus heat from existing diesel generators or potentially hydrogen technologies such as electrolyzers and fuel cells. To improve the quality of the results, the standard EnergyModelsX model will be extended with more detailed description of e.g. electrolyser efficiency operating at different loads and more detailed description of available heat at given operating temperatures.

Electric Vehicle Routing Problem with Robots, Parcel Lockers, and Time Windows – NIMA MORADI, INFORMATION SYSTEMS ENGINEERING, CONCORDIA UNIVERSITY – Fereshteh Mafakheri, Chun Wang

Integrating electric vehicle (EV) and robot-based parcel delivery alongside pick-up points, known as parcel lockers (PL), could benefit from the EV's efficiency, the robot's maneuverability, and the PL's flexibility. Parcels are loaded onto EVs, equipped with robots at the central depot. These vehicles then distribute parcels directly to customers (EV-based home-delivery service) or PLs for customer collection (self-collection service). Additionally, EVs may make stops at designated points, typically parking sites, to deploy robots that deliver packages to nearby customers (robot-based home-delivery service). Afterward, the robots return to the same EV to retrieve another package or recharge their batteries, employing the "dispatch-wait-collect" tactic. Due to its theoretical complexities, this multi-modal delivery system has yet to be addressed in the literature. The introduced parcel delivery approach is called the Electric Vehicle Routing Problem with Robots, Parcel Lockers, and Time Windows (EVRP-RPLTW). It minimizes the EV route and usage costs, the route cost of robots, and the utilization cost of PLs while satisfying the EVs and robots' cargo load and battery capacities, the PL's accommodation capacity and coverage range, and customers' (hard) time window constraints. A novel mixed-integer linear programming (MIP) model is proposed for EVRP-RPLTW to study and analyze the optimal solution of the problem under various scenarios. Several valid inequalities are proposed to strengthen the MIP model, and the sensitivity of the problem's parameters is analyzed. The computational results show the superiority of the introduced approach over the EVRP-TW, and some managerial insights are also presented.

Service level requirements in real-life-sized bike sharing systems – JEAN FRANÇOIS CÔTÉ, CIRRELT, UNIVERSITÉ LAVAL QUÉBEC, CANADA

Bicycle sharing systems (BSS) are vital to urban transportation networks, presenting complex optimization challenges. One such challenge is managing service levels for large-scale BSS, where the aim is to strategically redistribute bicycles across stations to anticipate fluctuating demand for origin-destination trips. While numerous studies tackle this issue, none, to our knowledge, have effectively handled real-life-sized BSS comprising thousands of stations, tens of thousands of bicycles, and hundreds of thousands of daily trips. We propose a two-step approach. First, we determine next morning service re-

quirements by formulating and solving stochastic programs to compute the bicycle quantities at each station. Second, we design vehicle routes to satisfy as many of these bicycle quantities as possible. Computational experiments utilize data from major BSS in Boston, Montréal, New York, and Washington D.C. Our results demonstrate efficient problem-solving with solutions closely approaching optimality. Additionally, we offer managerial insights into bicycle and station usage within BSS, facilitating better operational strategies.

4b - TB946 - 516E - **Transportation and logistics 2** Chair: SHEN PENG

Neural Network Estimators for Optimal Tour Lengths of Traveling Salesperson Problem Instances with Arbitrary Node Distributions – TAHA VAROL, HEC MONTRÉAL – Okan örsan özener, Erinc Albey

Complex routing problems are critical in achieving operational efficiency. These problems are typically addressed sequentially using cluster-first, route-second frameworks. Unfortunately, this method can lead to suboptimal results due to a myopic clustering phase. To overcome this limitation, we propose a strategy that incorporates information about the optimal tour lengths of potential clusters early in the process. This adjustment aims to make the framework less myopic. We introduce rapid and highly accurate estimators of Traveling Salesperson Problem (TSP) tour lengths based on neural networks (NNs). Our study is distinct in that we train our models using newly designed instances that closely replicate real-world logistics networks and morphologies. These instances introduce significant computational challenges, which we address using an unsupervised learning method to derive tight lower bounds and partial TSP solutions, subsequently used as predictors. By combining theoretical knowledge in the routing domain with neural networks, our approach achieves predictions with an average deviation of less than 0.7% from optimality. Extensive benchmarking shows that our models exhibit up to 100 times lower prediction errors than existing machine learning (ML) approaches on out-of-distribution test instances. Moreover, we integrate our models into a metaheuristic framework, creating an enumeration-like solution framework, capable of efficiently solving routing problems involving up to 100,000 cities. Compared to the current state-of-the-art solvers, our method significantly improves both solution time and quality, demonstrating the effectiveness of our features, models, and methodology.

Online Stochastic Optimization for Real-Time Transfer Synchronization in Public Transportation Networks – LAURA KOLCHEVA, POLYTECHNIQUE MONTRÉAL –

Transfer speed and protection are critical factors that influence passengers' willingness to use public transportation. Due to unpredictable traffic patterns, fixed transfer schedules may not always align. This work proposes two online stochastic optimization algorithms for the transfer synchronization problem using the hold, skip-stop, and speedup tactics. First, we design an offline arc-flow model using time-expanded graphs to enumerate all possible tactics. The model minimizes total passenger travel time by reducing transfer times. Then we implement two online stochastic optimization algorithms based

on the offline model in a discrete-event dynamic environment. Decisions are made based on predictions of the future state of the bus network using sampling to generate scenarios from historical and real-time data. The performance of the algorithms is compared using a real dataset from the public transit system of Laval, Canada. The results show significant improvements in both the number of successful transfers and total passenger travel time across 29 bus lines. This supports the practicality of using online stochastic optimization algorithms to solve the transfer synchronization problem in real-world transportation systems.

Distributionally robust facility location problem under facility disruption, uncertain demand and cost – SHEN PENG, XIDIAN UNIVERSITY – Kewei Jie

This paper focuses on facility location problems with facility disruption, uncertain demand, and transportation costs without capacity constraints. Firstly, we analyzed the sources of uncertainty information in the facility location problems and the adaptability and tractability of the distributionally robust model under different uncertain information. Secondly, we incorporate the probability of disaster occurrence and the effects of disruption propagation into the construction of an ambiguity set for the disruption scenario. Furthermore, we introduce information regarding support, mean, and upper-bound distribution into the construction of ambiguity sets for uncertain demand and transportation cost scenarios. In terms of tractability, for all the mentioned scenarios, we reformulate the distributional robust model as a mixed-integer linear program. Finally, we validate the effectiveness and applicability of our model through case studies involving emergency facility placement in earthquake scenarios and the positioning of epidemic detection facilities.

4b - WA947 - 516E - **Transportation and logistics 3** Chair: DAVID SCHINDL

A two-stage stochastic programming model with recourse for a Production Routing Problem with uncertain availability of vehicles – MICHEL GENDREAU, CIRRELT AND MAGI, POLYTECHNIQUE MONTRÉAL – Alline Zanette, Jorge E. Mendoza, Walter Rei

Supply chain planning has become a major concern of many companies. It is now well known that optimizing one activity of the supply chain often prevents the achievement of better solutions in others. Therefore, one should consider integrated problems, such as the Production Routing Problem (PRP), in which one performs the joint and simultaneous optimization of production, inventory, distribution, and routing decisions. Most PRP models consider only deterministic data, although uncertainty is a major issue in supply-chain management. Past studies have considered uncertainty in demand, quantity of products returned and quantity of components in each product. Here, we consider a PRP over a finite horizon, with a single capacitated production facility, a single product, a fleet of homogeneous vehicles, and uncertainty in the availability of vehicles. This is a problem setting seldom considered in previous studies, but commonly found in industrial environments.

Our PRP is defined on a complete graph. The plant is

located at node 0 and the other nodes represent customers. A single product is produced by the plant with production time-dependent capacity along a finite discrete-time horizon. When production takes place, a fixed setup cost and a variable production cost per unit are incurred. The products can be stored by the plant and by the customers up to an inventory limit, incurring an inventory holding cost per unit.

The distribution is performed using a limited fleet of homogeneous vehicles. For each period, a set of routes starting and ending at the plant is defined. However, the number of available vehicles is an independently distributed random parameter, with respect to periods, and its value depends on the general random event. To obtain a manageable model, we sample scenarios from the random space to empirically define the possible realizations of the random parameter. For a given period, if the minimum number of vehicles required to serve all the customers is not available, some routes cannot be performed, and a recourse policy must be adopted. The recourse policy that we study involves redirecting the customers who cannot be served due to the lack of vehicles; in fact, these clients are outsourced to a third party, which will individually deliver the products to each customer. The goal of the problem is to simultaneously minimize production, inventory, routing, and recourse costs, while respecting inventory limits, and production and vehicle capacities.

We propose a classical stochastic programming two-stage formulation with recourse: production, inventory and routing plans are decided in the first stage, while the recourse decisions belong to the second stage, after the realization of the random event. The objective function minimizes the sum of production, transportation, inventory holding, and expected recourse costs.

To solve efficiently the proposed model, one must exploit the inherent structure of the problem at hand. An obvious approach is to apply Benders decomposition by separating the two stages of the problem. Computational experiments are now under way on instances derived from existing benchmark instances described. Detailed computational results will be reported at the conference.

Generalized Dial-A-Ride Problem on Road Networks with Time-dependent Travel Times – BAHMAN BORNAY, CIRRELT AND MAGI, POLYTECHNIQUE MONTRÉAL – Michel Gendreau, Bernard Gendron

Classical Dial-a-Ride problem (DARP) is typically formulated on a complete customer-based graph, where vertices represent customer locations and arcs denote the shortest path between them. Further, a unique pair of pickup and dropoff (PD) nodes is associated with each request. We argue that these models might not be responsive to the contemporary transportation needs, hence leading to suboptimal solutions. First, in reality, an actual road network connects customer locations, where the speed on links changes over time, resulting in multiple shortest paths within various intervals throughout the day. Moreover, in modern logistics, flexibility is no longer a feature but a necessity. To be competitive in today's market, carrier companies and on-demand transit providers may need to offer clients a selection of potential pickup and dropoff locations, each with its own time windows, based on availability and constraints. Despite this, the rich vehicle routing problem body of knowl-

edge does not entail any article which explicitly investigates the combination of the aforementioned aspects. Hence, we formally introduce a new class of DARP with added layers of complexity described earlier, termed Generalized Dial-A-Ride Problem (GDARP). This model assumes a set of potential pickup and dropoff locations for each transit request, studied on real road networks with time-dependent travel times (TD-GDARP-RN). We present a Mixed-Integer Programming (MIP) model and formulate the scheduling problem for each vehicle as a Dynamic Programming (DP) problem. To solve the DP to optimality, we propose two algorithms: one employing forward propagation linearly, and the other utilizing a binary tree data structure in both recursive and iterative methods. Solving the DP for each route yields an optimal solution to the NP-Hard, Optimal Start Time Problem (OSTP). Additionally, we present an efficient DP-embedded hybrid adaptive large neighbourhood search (ALNS) metaheuristic that generates high-quality solutions.

A Stochastic Approach for Mobile Clinic Deployment Planning – ROSEMARIE SANTA, GEORGIA INSTITUTE OF TECHNOLOGY – Teodor Gabriel Crainic, Marie ève Rancourt, Marilene Cherkesly

A Stochastic Prize Collection Methodology for mobile Clinic Deployment Planning Mobile clinic deployments are often used to provide healthcare to underserved or disaster-affected populations. Practitioners aim to deploy mobile clinics to reach populations with the most critical healthcare needs. However, during humanitarian operations uncertainty arises in the travel time, usability of roads, and access. To address these challenges, we model mobile clinic deployment as a Two Stage Stochastic Prize Collection Problem, to maximize the benefit offered by mobile clinics while considering the uncertainty. We also propose and study the effect of four recourse policies, that defer in their routing decisions. Results and insights will be presented based on the cases of Indonesia, Iraq, Kenya, and Malawi.

A column generation approach for the routing of electricity technicians – DAVID SCHINDL, UNIVERSITY OF APPLIED SCIENCES WESTERN SWITZERLAND, HEG GENÈVE – Elise Bangerter, Pacheco Meritxell

The maintenance of an electricity distribution network involves numerous daily technical interventions. In this problem, we are given a set of interventions each with associated time windows, location, necessary skills and duration, as well as a set of teams of technicians with associated set of skills, each represented by a vehicle. We need to find feasible routes on the interventions for each vehicle, taking into account the time windows and skills, and ensure that each vehicle returns to its departure depot before the end of the day. The primary objective is to maximize the total duration of completed interventions and as a secondary objective, we aim to minimize the overall routing cost. This problem can be formulated as a capacitated vehicle routing problem with time windows. Due to the large number of vehicles and interventions, this results in a large-scale optimization problem, and its operational nature limits the time available for exact solving. Here, we propose a column generation approach where the subproblem decomposes into a subproblem per vehicle and each potential route of a vehicle is considered as a new column in the master problem. To generate these routes, we rely on dynamic programming. Real-world

instances from EDF (Electricité de France) of historical technicians' interventions will be used to evaluate the effectiveness of the proposed methods.

4b - WB147 - 516E - Large-scale optimization in action: methods and applications Chair: ALEXANDRE JACQUILLAT

A Better Match for Everyone, Reinforcement Learning for Ridesharing – SEBASTIEN MARTIN, NORTHWESTERN UNIVERSITY, KELLOGG SCHOOL OF MANAGEMENT –

I teamed up with Lyft, a popular ridesharing platform, to improve its online matching system for drivers and riders. To create a more efficient algorithm that can quickly adapt to changing conditions, we introduced a novel online reinforcement learning approach with a hierarchical linear approximation of the platform's state space adapted to a nonstationary environment. This overhaul of Lyft's core system resulted in millions of additional rides and more than \$30 million in extra revenue annually. Our approach is the first documented implementation of a ridesharing matching algorithm that can adapt to a nonstationary environment, and our experiments proved that this change benefits everyone involved - drivers, riders, and the platform. However, it was difficult for a large organization to trust an algorithm that can change itself.

Branch-and-price for prescriptive contagion analytics – ALEXANDRE JACQUILLAT, MIT SLOAN SCHOOL OF MANAGEMENT – Michael Li, Martin Rame, Kai Wang

Predictive contagion models are ubiquitous in epidemiology, social sciences, engineering, and management. This paper formulates a prescriptive contagion analytics model where a decision-maker allocates shared resources across multiple segments of a population, each governed by continuous-time dynamics. We define four real-world problems under this umbrella: vaccine distribution, vaccination centers deployment, content promotion, and congestion mitigation. These problems feature a large-scale mixed-integer non-convex optimization structure with constraints governed by ordinary differential equations, combining the challenges of discrete optimization, non-linear optimization, and continuous-time system dynamics. This paper develops a branch-and-price methodology for prescriptive contagion analytics based on: (i) a set partitioning reformulation; (ii) a column generation decomposition; (iii) a state-clustering algorithm for discrete-decision continuous-state dynamic programming; and (iv) a tri-partite branching scheme to circumvent non-linearities. Extensive experiments show that the algorithm scales to very large and otherwise-intractable instances, outperforming state-of-the-art benchmarks. Our methodology provides practical benefits in contagion systems; in particular, it can increase the effectiveness of a vaccination campaign by an estimated 12-70%, resulting in 7,000 to 12,000 extra saved lives over a three-month horizon mirroring the COVID-19 pandemic.

Heterogeneous Treatment Effects In Matching Marketplaces Under Interference: From Estimation to Decision – ARTHUR DELARUE, GEORGIA INSTITUTE OF TECHNOLOGY – Kleanthis Karakolios

Marketplace companies routinely use randomized experiments

to make operational decisions. Randomized controlled trials are often used to decide whether a specific intervention should be rolled out. Because rollouts can be global (affecting all users) or partial (targeting some users), it is important to measure the effect of the treatment intervention on specific subpopulations of users, also known as heterogeneous treatment effect analysis. However, marketplace experiments suffer from interference, where the treatment status of one unit can affect the outcome of another unit. We first establish that heterogeneous treatment effects are not necessarily uniquely defined in the presence of marketplace interference. We then distinguish two cases: when the heterogeneous treatment effect on a subset of users is uniquely defined, we show that a “shadow price estimator” based on linear programming duality can provably reduce bias from the standard estimator. When the heterogeneous treatment effect is not uniquely defined, we recast the problem of estimating the impact of treatment on each group of users as the problem of deciding whether or not to roll out the treatment to each group of users. We use the shadow price estimator to design a robust decision rule, which we show is more effective and less conservative than a similarly robust decision rule based on the standard estimator. We verify our theoretical results with numerical experiments on synthetic data.

Modeling and Optimizing Large-Scale Inventory Routing in East Africa: A Data-Driven Optimization Approach – SIMON THOMÄ, RWTH AACHEN UNIVERSITY – Sergio Camelo, Dan Iancu, Maximilian Schiffer

We investigate a large-scale inventory routing problem in the context of distributing clean cooking stoves in East Africa. Using a comprehensive dataset from an industry partner, we develop a stochastic model that accounts for significant supply and demand fluctuations, censored and incomplete data, and store substitution patterns and propose novel data-driven approaches to solve the resulting problem at scale. Our numerical results demonstrate that the proposed method yields significant improvements over the status quo, leading to significant cost savings and improved accessibility to clean cooking solutions.

4b - ThA948 - 511F - **Transportation and logistics 4** Chair: ALI KERMANI

Valid Inequalities for the Vehicle Routing Problem with Drones – FAIZ HAMID, INDIAN INSTITUTE OF TECHNOLOGY KANPUR – Yogesh Kumar Agarwal

We study the vehicle routing problem with drones (VRPD), a variant of the classical vehicle routing problem, in which both drones and trucks serve customers with the objective of minimizing transportation cost. Drones can reduce operating expenses because their per-unit travel cost is lower than that of trucks. However, due to limited carrying capacity and flying range, a fleet of drones must rely on trucks to serve long-distance or heavy-demand customers. Drones can be dispatched from and picked up by the trucks at the depot or any of the customer locations. VRPD is inspired by the increasing interest in commercial drone delivery by companies such as Amazon, Federal Express, DHL, and Walmart. Utilization of drones helps reduce energy usage and carbon emissions, repre-

sending a step toward fostering a sustainable and environmentally friendly supply chain.

We develop a mixed-integer linear programming model for VRPD. A number of valid inequalities are proposed to strengthen the MILP formulation. The effectiveness of the valid inequalities is verified using a computational study. The computational results demonstrate a significant reduction in the CPU time required to solve these problems to optimality.

A Matheuristic for the Stochastic Production Routing Problem with Adaptive Routing and Service Levels – ALI KERMANI, HEC MONTRÉAL – Jean François Cordeau, Raf Jans

In modern supply chains, improved coordination among different parties is sought-after as it can enhance overall efficiency and lower costs through integration. Therefore, the production routing problem (PRP) has emerged to tackle this integrated problem, combining production, inventory, and routing decisions into a unified framework to improve coordination throughout the supply chain. However, neglecting demand uncertainty can lead to profit losses and poor customer satisfaction.

To mitigate this risk, companies often adopt service level targets to meet a certain portion of demand while considering operational constraints and minimizing costs. In our study, we address the stochastic PRP with adaptive routing (i.e., second-stage) and service level constraints. We examine four distinct service levels, each tailored to address specific metrics aligned with different real-world requirements, such as managing stock-out occurrences and backorder or backlog ratios.

We propose a two-stage stochastic formulation where the setup decisions are made in the first-stage while the production, inventory, and routing decisions are made after demand realization, facilitating rapid responses to demand fluctuations. However, this approach brings computational challenges, which we overcome through an iterative matheuristic algorithm designed to deliver high-quality solutions within reasonable computation times.

This algorithm decomposes the problem into three subproblems. Initially, we tackle a two-level lot sizing problem, focusing on a single aggregate customer and direct shipments to rapidly generate setup decisions. Subsequently, we handle a restricted PRP with predefined setup plans and a vehicle with aggregate capacity, integrating service level constraints. Lastly, we focus on deriving feasible routing decisions, allowing for some flexibility in production and inventory choices while ensuring adherence to service level criteria. We then iterate over these phases to enhance the quality of solutions. By applying our algorithm to various benchmark scenarios, we compare results across different service levels and offer insights based on our findings.

4b - ThA949 - 516E - **Transportation and logistics 5** Chair: CAGLA DURSUNOGLU

An exact branch-and-price-and-cut algorithm for a practical dial-a-ride problem – MOHAMMAD KARIMI, DEPARTMENT OF MATHEMATICS AND INDUSTRIAL ENGINEERING, POLYTECHNIQUE

MONTREAL / CIRRELT – Guy Desautniers, Michel Gendreau

The Dial-a-Ride Problem (DARP) involves designing vehicle routes and schedules for customers with specified pickup and delivery requests between their origins and destinations. The objective is to minimize the cost of vehicle routes while adhering to constraints such as vehicle capacity, time windows, and maximum ride time. This paper proposes a branch-and-price-and-cut (B&P&C) algorithm to address a variant of DARP where both customers and the vehicle fleet are heterogeneous, alongside practical constraints like budget, breaks with different configurations, and maximum run duration. The algorithm employs a column generation approach at each node in the branching tree, decomposing the problem into a master problem and subproblems. Exact and heuristic versions of the labeling algorithm are developed to efficiently handle these subproblems. The effectiveness of the algorithm is assessed through a real-world case study involving 849 heterogeneous passengers and over 70 available vehicles, represented as contracts with varying cost structures. Notably, this instance marks the largest problem size addressed using an exact B&P&C algorithm to our knowledge. Results demonstrate that the proposed algorithm achieves highly competitive solution quality.

A Column-Generation Algorithm for the Intercity Electric Coach Scheduling Problem – MALTE BILLEN, THE UNIVERSITY OF EDINBURGH – Jörg Kalcsics, Sergio Garcia, Goncalo Dos Reis, Adrian Grütter

Governments globally are incentivising zero-emission vehicle sales and charging infrastructure development to cut transportation emissions. These incentives, along with higher market availability, lower fuel cost and extended battery range, make electric coaches viable for intercity public transportation. In collaboration with Ember Core Ltd., the first intercity operator in the United Kingdom to use an all-electric fleet, we have conducted research on maximising the utility of their electric coaches and charging infrastructure while ensuring that schedules fulfil the operational restrictions of intercity electric transportation. As a result, we introduce the Intercity Electric Coach Scheduling Problem along with a column generation algorithm that uses a variable fixing heuristic to find integer solutions. The algorithm is embedded in an iterative framework to generate schedules for a long planning horizon. It provides quality schedules that successfully meet Ember's real-world requirements. In the forthcoming presentation, we will discuss the challenges of sparse charging networks, regular maintenance scheduling and overnight timetables that make intercity electric coach scheduling a difficult problem to solve. However, we also explain how we can exploit these characteristics and provide a comprehensive analysis of the algorithm's solution quality and performance.

A unifying framework for selective routing problems – CAGLA DURSUNOGLU, BILKENT UNIVERSITY – Okan Arslan, Sebnem Manolya Demir, Bahar Yetis Kara, Gilbert Laporte

We present a unifying framework for Selective Routing Problems (SRPs) through a systematic analysis. The common goal in SRPs is to determine an optimal vehicle route to serve a subset of vertices while covering another subset. They arise in diverse fields such as logistics, public health, disaster response,

and urban development. To establish a unifying framework for different but related problems, we associate the notion of service with coverage and argue that routing is a tool of service. We classify SRPs according to their selectiveness degree and emphasize the breadth and depth of this problem in terms of its characteristics. This SRP framework helps us identify research gaps as well as potential future research areas. We present a generic mathematical model, use it to describe the connections among these problems and identify some identical problems presented under different names.

4b - ThB155 - 511F - **Models and Algorithms in Last-mile Delivery** Chair: OKAN ARSLAN ↔

EV delivery, charging and parking planning method under uncertain arrival times at logistic depots – HANG DONG, HITACHI, LTD. – Junko Hosoda, Akane Seto, Kazuya Uyama

In the logistics industry, the use of electric vehicles (EVs) in delivery operations is required to achieve a sustainable society. Since EVs have a shorter driving range than internal combustion engine vehicles (ICEVs), charging between deliveries is needed to extend the driving range of EVs. However, since the budget of the initial investment for electrifications is limited, the number of chargers installed at logistic depots may be less than the number of EVs. In addition, there may be fewer parking lots than EVs at logistics depots. For these reasons, it is difficult to charge all EVs simultaneously. Therefore, an EV fleet charging management solution is needed to replace ICEVs with as many EVs as possible. This research focuses on logistic depots, where vehicles are responsible for multiple delivery jobs each day, and a parking lot with a limited number of chargers is facilitated. When a delivery job is completed, the vehicles return to the depot and, if necessary, are recharged before the next delivery job. All delivery jobs are given in advance and must be assigned to vehicles. The first challenge of managing EV fleet at such a depot is that a plan that incorporates both delivery and charging is necessary. The second challenge is the power shortages caused by the delay in vehicle's arrival. This is because, if an EV arrives at the depot later than scheduled, its charging time may be shortened, which may lead to power shortages in subsequent delivery jobs. To address this issue, a robust EV fleet management method which accommodates uncertain arrival times is needed. The third challenge is the computation time limitation. We aim to solve this problem in a short time to apply this method to on-site operations. To overcome the challenges, we developed a robust EV fleet management method that integrates the determination of delivery job assignment, charging plan and parking plan for EVs, while considering the accommodation of robustness against uncertain delays. To solve this problem within a limited time, we propose a method that divides the problem into two steps: (1) generating diverse solutions and then, (2) evaluating their robustness under uncertainties. At step (1), we applied several parameterized constraints that may have an impact on power shortages, and help solutions be more resilient in terms of power, charging time, and interval time. Computational experiments show that a constraint on the lower bounds of possible charging time has the most impact on the result, and our method can replace 94% of ICEVs with EVs, an increase of 9 percentage points over a conventional method that does not consider uncertainties.

Optimizing Ultra-Fast Delivery Networks and

Service Guarantees Under Uncertainty – XIN WANG, HEC MONTRÉAL – Okan Arslan, Jean François Cordeau, Erick Delage

Ultra-fast delivery revolutionizes food and grocery services, with several companies advertising delivery times under 15 to 30 minutes. Motivated by the multi-billion-dollar industry that has emerged in recent years within the delivery business, we investigate the network design problem for ultra-fast delivery services. This involves decisions on micro-depot locations and customer allocations, considering various service guarantee levels. We develop robust probabilistic envelope constrained (PEC) programs to handle uncertainties in travel times and customer order arrivals, and jointly optimize the protection level to avoid both excessive risk and conservatism. To enhance the tractability of PEC models, we derive their equivalent semi-infinite linear programs and propose inner and outer approximations with finite linear constraints. We validate the accuracy of these approximations through extensive experiments using real-world data from Amazon and the Google API, along with a comparative study of different formulations. Varying service levels in ultra-fast delivery affect profitability and reliability, contingent on service level definitions and compliance probabilities of these guaranteed service levels. We find that a daily service level with multi-layer partial protection outperforms other policies investigated in this paper, yielding higher profitability and mild violations of service level guarantees, and it proves to be an effective strategy for profitable and reliable ultra-fast delivery without over-committing or under-delivering, regardless of ordering times or traffic conditions. Additionally, empirical evidence indicates that providing ultra-fast delivery in rural areas poses unique challenges compared to urban settings.

Technician Routing and Scheduling under Uncertainty – MILAD ELYASI, POSTDOCTORAL RESEARCH FELLOW – Amira Dems, Yossiri Adulyasak, Okan Arslan, Jean François Cordeau

This study introduces an advanced mathematical optimization approach for efficient scheduling and routing of technicians in field service operations, aiming to maximize profitability and customer satisfaction while addressing uncertainty. It categorizes customers into mandatory and optional visits and assigns technicians to vehicles based on skill levels and operational constraints. The model incorporates stochastic elements to manage unpredictability in travel and service times. The uncertainty in travel and service time is modeled using chance constraints. Constraints guarantee precise technician-to-vehicle assignments, feasible routes, adherence to time limits, and appropriate skill levels for each service. The problem is solved using a logic-based Benders decomposition approach that decomposes the original problem into a master problem and a series of subproblems.

A contextual framework for learning routing experiences in last-mile delivery – OKAN ARSLAN, HEC MONTRÉAL – Huai Jun (Norina) Sun

We present a contextual framework for solving the data-driven traveling salesman problem in last-mile delivery. The objective of the framework is to generate routes similar to historical high-quality ones, as classified by operational experts, by considering the unstructured and complex features of last-mile delivery operations. The framework involves learning a transition likelihood matrix between customer stops and using a classical TSP solver to generate routes similar to those high-

quality routes in a given dataset. We employ feature-based factorization of the transition likelihood matrix, which reduces the dimensions of the information to be learned. We develop a derivative-free algorithm by extending the coordinate search method and a complementary method for learning the transitions directly from the data. We test the efficiency of the methods using a case study based on the Amazon Last-Mile Routing Challenge and show that our methods are effective, interpretable, and flexible. A preliminary version of our methods achieved third place in the aforementioned challenge. We improve on this result by 23% in an out-of-sample dataset.

4b - ThB950 - 516E - **Transportation and logistics 6** Chair: ALEJANDRO TORIELLO

An integer optimization approach for the graph-free multi-agent pathfinding problem – SEYOUNG OH, SEOUL NATIONAL UNIVERSITY – Kyungsik Lee

The multi-agent pathfinding problem is the problem of finding a set of collision-free paths for multiple agents traversing a graph. In MAPF, the underlying graph is given, and agents move along adjacent vertex at each timestep. However, this assumption oversimplifies the actual movement of physical agents. Also, in real-world applications such as controlling AGVs or drones, each agent often has computational capabilities to plan and follow its own trajectory. This talk presents a novel integer optimization approach to plan collision-free paths for agents traversing the Euclidean space considering kinematic characteristics of agents. To avoid the computational burden of modelling the movements of agents in detail, we discretize a given space into areas and plan the occupancy of the areas by agents over time. Our approach reduces the computational burden in the planning phase and allows some autonomy to the agents in the execution phase. We present a condition on the discretization under which planned paths are conflict-free, with examples that satisfy the condition. The planning problem is represented as a pattern-based integer optimization model, where each agent's area occupancy over time corresponds to a pattern. We use a robotics simulator to evaluate our approach.

Optimizing Containerized Multimodal Transport toward Sustainability and resilience – ASEFEH HASSANI GOODARZI, DÉPARTEMENT DE GÉNIE DES SYSTÈMES, ÉCOLE DE TECHNOLOGIE SUPÉRIEURE – Armin Jabbarzadeh, Marc Paquet

Multimodal transport is a new transportation approach aimed at reducing the role of roadways in freight transportation, and moving toward eco-friendly solutions. This strategy has emerged as a promising method to address environmental concerns and challenges associated with global warming by utilizing multiple transportation modes for freight transshipment. Containerized intermodal transport is a specific type of this strategy, in which the loading unit of container is preserved throughout the journey. In this paper, we introduce a two-stage stochastic optimization model designed to enhance the sustainability and resilience of containerized transport, particularly in the face of unexpected disruptions. The proposed model integrates various environmental and social sustainability metrics, including greenhouse gas emissions, noise pollution, traffic con-

gestion, and accidents. To tackle the complexities inherent in the model, we develop an efficient solution methodology based on Lagrangian relaxation. Validation of the effectiveness of the proposed approach is conducted using real-world transportation system data. Through analysis of the case study, the practical application of the model is illustrated in generating valuable managerial insights for those involved in the freight transport industry. Our analysis reveals that allocating a mere 0.4% of total costs to preparedness and recovery initiatives can result in a substantial reduction of approximately 4.7% in overall costs.

Charging Station Location and Fleet Sizing for Shared Autonomous Electric Vehicles using Benders' Decomposition – MICHAEL LEVIN, UNIVERSITY OF MINNESOTA – David Rey

The emergence of shared autonomous electric vehicle fleets in transportation networks is expected to yield substantial societal and environmental benefits. In this paradigm, automated and electrified fleets are coordinated to serve passenger travel demand. The operations of shared autonomous electric vehicle systems requires the coordination of recharge trips and rebalancing trips in between passenger trips. Thus, the design of the charging infrastructure is key to the efficiency of the system. We address the charging station location problem in the context of shared autonomous electric vehicle systems. Building on a throughput-maximizing dispatch policy for shared autonomous electric vehicle operations, we propose a mixed-integer linear programming formulation to locate and size charging stations in the network. We exploit the structure of the problem to derive several classes of valid inequalities that aim to strengthen the formulation. We develop Benders' decomposition approaches to enhance the tractability of the solution methods. We conduct numerical experiments on three publicly available transportation networks to investigate the computational performance of four algorithmic configurations. We also conduct sensitivity analyses on problem data including demand, vehicle battery range, fleet cost and the maximum number of chargers available. Our findings show that the derived valid inequalities have a significant impact on reducing computational runtime. Furthermore, we observe that embedding on-the-fly valid inequalities in a Benders' decomposition algorithm can help in improving efficiency.

Optimizing the Trade-Off Between Batching and Waiting: Subadditive Dispatching – ALEJANDRO TORIELLO, GEORGIA INSTITUTE OF TECHNOLOGY – Ignacio Erazo

Motivated by applications in e-commerce logistics where orders or items arrive at different times and must be dispatched or processed in batches, we propose the subadditive dispatching problem (SAD), a strongly NP-hard problem defined by a set of orders with release times and a non-decreasing subadditive dispatch time function. A single uncapacitated vehicle must dispatch orders in batches to minimize the makespan, the time at which all orders have been dispatched. We propose a mixed-integer linear formulation for SAD; based on the linear relaxation's optimal solution, we construct a two-dispatch solution with a $3/2$ approximation ratio, and a solution with at most three dispatches that has a $4/3$ approximation ratio under an additional assumption. The guarantees are respectively best possible for solutions using at most two or three dispatches. Furthermore, we analyze FIFO solutions, discuss

cases when they are optimal, and provide a dynamic program to obtain them. Finally, we computationally test our methods on applications in same-day delivery and routing on restricted topologies.

4b - ThC951 - 516E - Transportation and logistics 7 Chair: ALI AKBAR SADAT ASL

Dive-and-cut-and-price heuristics for vehicle routing problems – STEFAN ROPKE, TECHNICAL UNIVERSITY OF DENMARK –

Branch-and-cut-and-price algorithms based on set-partitioning formulations are popular methods for optimally solving vehicle routing problems (VRPs). This talk explores vehicle routing heuristics based on the branch-and-cut-and-price approach. The exact method is turned into a heuristic by partially exploring the branch-and-bound tree and only solving the pricing problem heuristically. Results will be presented for the vehicle routing problem with time windows, the pickup and delivery problem with time windows, the dial-a-ride problem, the vehicle routing problem with drones, the electric vehicle routing problem with time windows, and the vehicle routing problem with stochastic demand. The handling of the master problem and the cutting and branching are generic for all considered VRP types, while the pricing sub-problem heuristics are tailored for the problem at hand.

Automatic model decomposition in Hexaly Optimizer – BIENVENU BAMBI, HEXALY –

Hexaly Optimizer, formerly known as LocalSolver, is a model-and-run solver that integrates heuristics and exact methods. A set-based modeling formalism was introduced to simplify the modeling of certain combinatorial problems like routing or packing problems. For instance, in a routing problem, list variables can be used to model the sequence of visits made by each truck. These decision variables are well suited for a heuristic search but are much more difficult to integrate in a mathematical programming approach to compute lower bounds. A direct reformulation in a Mixed Integer Linear Programming (MILP) model introduces a quadratic number of binary decisions with several big M constraints leading to poor scalability and bounds. Hexaly 13.0 automatically detects such structures in a user model and reformulates them in an extended MILP model to compute lower bounds. This model is solved efficiently using state of the art branch-cut-and-price technics from the literature. This talk will present the general approach and the algorithms used for the resolution and some benchmarks on the CVRP (capacitated vehicle routing problem) library.

An Efficient Real-time Multi-modal Single-leg Transportation System: Balancing Perspectives on User Needs, Operator Efficiency, and Sustainability – ALI AKBAR SADAT ASL, POLYTECHNIQUE MONTREAL –

Ridesharing has emerged as a promising solution to tackle urban mobility challenges, offering affordability and convenience while potentially alleviating congestion and environmental impact. This study investigates the use of transportation hubs, equipped with parking facilities, which serve as both the origin and destination of rideshare trips. The proposed system integrates personal vehicles with carpooling options and shut-

tle services as dynamic transit solutions to serve ride requests. Our study encompasses the needs and goals of users, operators, emissions reduction, and the overarching system efficiency. Employing a multifaceted approach, we solve the system optimization problem using column generation, ensuring that our investigation captures the diverse interests and priorities within the realm of urban mobility. Through this lens, we provide a holistic assessment of the proposed approach under various scenarios, illuminating its efficacy in addressing the complex challenges of large-scale urban mobility.

4b - FA193 - 516E - Network Design for Transportation Planning IV Chair: MARIA ELENA BRUNI

Stochastic Variable Cost and Size Bin Packing with Capacity Loss and Supplier Selection problem – GUIDO PERBOLI, POLITECNICO DI TORINO – Roberto Baretto, Teodor Gabriel Crainic, Walter Rei

In this work, we study the problem of a shipper planning its operations for delivering its orders using the transportation and storage services of contractor suppliers. The shipper needs to contract with the suppliers ahead of time because they do not provide their service in the short term. Moreover, capacity prices on the spot market are high and the shipper cannot run a cost-effective business by buying capacity just in time to meet the demand of orders. Uncertainty characterizes the decisions to make for computing ahead of time the shipper's plan of operations. At planning time, customer orders are still unknown, as well as the capacity loss of suppliers and the actual prices on the spot market for the spare capacity that the shipper may need to buy. Thus, the shipper wants to select a set of suppliers to contract with through medium-term agreements to secure sufficient capacity to meet the demand. Each agreement sets a fixed fee for engaging the supplier, the price to pay for the booked capacity, as well as the compensation the shipper pays to the supplier for the capacity booked, but not used, and the compensation the supplier pays to the shipper for the capacity finally not supplied. The shipper wants to diversify the risk of supply shortage by selecting more than one supplier and minimizing the expected costs of future operations. The time horizon of the shipper's problem is two-stage by definition, the shipper selects the suppliers in the first stage and ensures that all orders are shipped in the second one by packing all the orders into the supplied capacity and that one the shipper buys on the spot market (if needed). The two-stage capacity planning problem of the shipper is a variant of the stochastic variable cost and size bin packing problem that minimizes the expected total cost of the shipper's operations. For this problem, we provide a two-stage stochastic integer programming formulation and a solution framework based on state-of-the-art stochastic optimization techniques.

Continuous-Time Service Network Design with Stochastic Travel Times – GIACOMO LANZA, UNIVERSITY OF PISA – Teodor Gabriel Crainic, Maria Grazia Scutella, Mauro Passacantando
Service network design is inherently a stochastic problem involving decision-making in an environment with incomplete information. The aim is to tactically design a transportation plan to efficiently and profitably move origin-destination demand

through a network of transportation services in a consolidation-based freight transportation context (e.g., LTL-carriers, railroads, maritime liner navigation). The applicability of the resulting plan should consistently maintain the required high-quality service standards, crucial for customer satisfaction, despite the multiple sources of uncertainty that may impact its performance. Our focus lies on travel times, i.e., the time required by a service to travel from one point to another, the fluctuation of which can have ripple effects on the feasibility and profitability of the plan, hindering efficient consolidation and jeopardizing adherence to delivery deadlines, thus negatively impacting the carrier's reputation and revenues. We will describe a two-stage stochastic programming model prescribing the selection of services in the first stage when only probability distributions for travel times are available, and estimating costs (related to delays in operations, violating due dates in customer contracts, outsourcing) of such decisions in the second stage, where service operations are defined given travel time observations. The aim of the model is minimizing operational costs. The proposed formulation is defined on a compact network that models time continuously and, thus, mitigates the computational difficulties associated to traditional time-space networks. Preliminary results regarding the complexity of the problem and the efficiency of the formulation, together with hints on policies to hedge time fluctuations will be presented.

Door-to-Door Multimodal Long-Distance Freight Transportation Planning – MAHDI MOSTAJABDAVEH, POLYTECHNIQUE MONTRÉAL – Teodor Gabriel Crainic, Michel Gendreau, Walter Rei
Corridor-based logistics service providers facilitate freight exchanges between two specific regions (e.g., Southwest China and Canada) by orchestrating local operations in each region and the long-haul corridor connecting them. These logistics service providers work with multiple independent entities, including shippers (such as producers, wholesalers, and retailers) who need to transport items from one region to another, and carriers who offer services such as long-haul transportation, local transportation within each region, sorting and consolidating at terminals, and storage at warehouses.

In this study, we address the integrated, time-dependent operational planning of such a system, seeking to synchronize activities across the different parts of the system over a planning horizon. The problem focuses on three main components: route planning for local vehicles in one region, consolidating shipments on multimodal long-haul scheduled services for transfer along the corridor, and route planning for local vehicles in the other region.

To efficiently solve this complex problem, we present the Pattern-Guided Search (PGS) heuristic. This heuristic maintains a diverse, high-quality set of elite solutions and guides the search using patterns extracted from this elite set. PGS utilizes a two-layer local search, pattern-guided perturbation, and diversification mechanisms to achieve high performance. PGS significantly outperforms general solvers (such as Gurobi) on medium and large-sized problem instances. We conducted extensive experiments and derived practical insights for decision-makers in the operational planning of corridor-based logistics systems.

Scheduled Service Network Design with Packing Considerations – MARIA ELENA BRUNI, UNI-

UNIVERSITY OF CALABRIA – Teodor Gabriel Crainic, Sara Khodaparasti, Guido Perboli, Walter Rei

Modern transportation and Supply Chain systems require a holistic view of the sustainability of their operations, considering different aspects (operations, economies of scale, social and environmental impacts) while finding a trade-off between maximization of the revenues and high quality of service for customers. One of the principal modeling frameworks in this context turned out to be the Service Network Design. Some aspects of it, however, may limit its applicability. One of those is the absence of explicit packing aspects related to the services to be operated on the network. In this talk, we discuss new variants of service network design, namely the Service Network Design with Bin Packing Considerations, that incorporate the families of packing problems already in use in capacity planning. We will present a unified problem setting addressing simultaneously decisions on the selection of scheduled services (on a general time-space network), the selection of the type and number of capacity units to load and move the origin-destination demands, the assignment of demand flows to the loading units, and the construction of the demand itineraries within the selected service network.

4b - FB164 - 516E - Humanitarian Logistics

Chair: VITTORIO NICOLETTA

A Mutual Catastrophe Insurance Framework for Horizontal Collaboration in Prepositioning Strategic Reserves – HANI ZBIB, UNIVERSITY OF QUEBEC IN MONTREAL – Burcu Balcik, Marie ève Rancourt, Gilbert Laporte

Horizontal collaboration in humanitarian logistics can be highly beneficial through pooling and sharing resources, but it can be costly and difficult to create and operate. Moreover, communities may lack funds to establish it and often depend on donors for the initial investment. Integrating the logistical and financial functions within a catastrophic risk management framework addresses these challenges. The role of such a framework is to determine how to integrate these functions together, how to pool risk jointly, and what are the most effective cost and benefit sharing mechanisms.

In this talk, we present how the catastrophe insurance literature, which is rich in sharing mechanisms, can be integrated in a framework for horizontal collaboration for the prepositioning of strategic reserves. The framework consists of a risk-averse umbrella insurer offering multi-year insurance contracts to a portfolio of risk-averse policyholders wanting to collaborate and pool their risks against natural disasters. It encompasses four components, two operational and two financial. The operational functions consist of 1) planning the prepositioning network in preparedness for incoming insurance claims, and 2) insurance contracts design (setting coverage deductibles and limits of policyholders, and providing insurance coverage to the claims in the emergency response phase). The financial functions consist of 3) ensuring the insurer's solvency by efficiently managing its capital, and 4) allocating yearly premiums among policyholders.

We model the framework as a very large-scale non-linear multi-stage stochastic program, and present an exact method to solve it based on a tailor-designed Benders decomposition algorithm. We study the case of Caribbean countries establish-

ing a horizontal collaboration for hurricane preparedness and draw managerial insights from the case study.

Mitigating fire risk towards critical and residential structures near a high ignition area using Critical Node Detection – VITTORIO NICOLETTA, HEC MONTRÉAL – Valérie Bélanger, Denys Yemshanov, Raphaël D. Chavardès, Anne Cotton Gagnon, Jonathan Boucher

Wildfires pose a significant threat to both critical infrastructure and residential areas, especially in regions with a high proportion of wildland-urban interface (WUI) and prolonged fire seasons due to climate change. This study explores the application of Critical Node Detection (CND) to inform strategic landscape management plans aimed at limiting fire spread and intensity while considering the restrictive costs associated with hard-to-access areas. Our project focuses on a landscape with intensive military training, where ignitions are frequent. Our objective was to identify optimal fuel treatment locations to restrict wildfires from escaping the base and impacting neighboring communities. We integrated CND with fire-growth modeling and structural loss rate modeling to comprehensively assess wildland fire risk. Our preliminary results suggest which particular strategies to adopt according to the desired objective of the mitigation and the particular fire profiles, comparing their costs and relative efficacy. Our approach encompasses the assessment of fire hazards, impacts, and mitigation strategies, offering valuable insights for proactive wildfire management in comparable settings. This interdisciplinary framework serves as a robust tool for safeguarding communities and bridging the realms of fire science, land management, and military operations, thereby enhancing overall wildfire risk management effectiveness.

The Value of Demand Prediction for Improved Food Security – ALBORZ HASSANZADEH, HEC MONTRÉAL – Valérie Bélanger, Marie ève Rancourt, Feyza Sahinyazan

Hunger and famine pose great risks to global health and are the second in the UN's seventeen sustainable development goals to be achieved by 2030; however, they prove to be quite challenging to eradicate or alleviate. To mitigate their devastating impact, each year, aid agencies deliver tons of food commodities to populations in need. However, the delivery of food commodities is often expensive, and because of the complex intertwining factors shaping food security, it is very difficult to definitively predict future outcomes and demand for food aid. Without a timely identification of vulnerable populations, food aid often fails to arrive in the right place at the right time. We develop a stochastic optimization framework to assess the value of information: our analytical results quantify the advantages of incorporating food insecurity predictions in decision-making. Such predictions facilitate informed prepositioning decisions. Since acquiring relevant data might require heavy investments, we also analyze the delicate balance between allocating the limited available budget to the prepositioning of food commodities and investing in the acquisition of accurate data.

A bilevel optimization approach for shelter network design and evacuation planning problem: An application to flood preparedness in Haiti – MAEDEH SHARBAF, HEC MONTRÉAL – Marie ève Rancourt, Valérie Bélanger

We present a decision-support tool for flood preparedness developed through a collaboration with the World Bank in Haiti. The shelter network design and evacuation planning problem is formulated as a bi-level optimization model in which the leader is the decision-making authority, and the followers are the evacuees whose reactions to upper-level decisions is modeled. The model considers the inclusion of evacuees' behavior and the temporal evolution of flood disaster over time. The model is tested using socio-demographic, regional characteristics, attributes of hazard, and GIS data.

4b - FB172 - 511F - Demand-based and stochastic urban transportation planning Chair: LÉA RICARD 

Competing on Emissions Charges – GIAN-MARCO ANDREANA, UNIVERSITY OF BERGAMO – Nicole Adler, Gerben De Jong

We develop a game-theoretic framework in order to estimate market equilibria in the presence of environmental policies at both supra-national and global levels. In a two-stage game, regulators maximize social welfare over their jurisdiction by setting emission charges, whilst airlines compete through frequencies, fares, and fleet choice. Airline decisions include the extent to which costs of the environmental charges are absorbed or passed on to consumers and to what extent the fleet is upgraded and utilized across the network. The equilibria outcomes suggest the presence of multiple potential distortions in aviation markets that could undermine the effectiveness of environmental policies. We apply the model to North American, Western European, and Trans-Atlantic markets and find that the most efficient mechanism is the one of a single regulator able to tax discriminate between regions. Furthermore, we show that an increase in welfare would come at the expense of the environment because the effectiveness of the policy is limited by regulators interacting in a self-interested manner.

Hub Transportation Problem with Chance Constrained Due Dates – ÖYKÜ NAZ ATTILA, POLYTECHNIQUE MONTREAL – Antoine Legrain, Quentin Cappart

Traditional public transportation systems provide efficient mass transportation solutions between designated stations. Despite its advantages, public transportation systems suffer from being restricted to serve a fixed area. This reduces its accessibility considerably, particularly in remote areas. Hub transportation systems overcome this problem by offering shuttle services to serve passengers who face such restrictions. These systems are highly time sensitive, as shuttles are required to return to their station by a due time to ensure that all passengers are served on time. From a decision making perspective, it is essential that the routing decisions made for shuttles are reliable against changes in travel times. To achieve this, we propose a chance constrained programming approach within a column generation framework that guarantees that the final solution remains feasible (with respect to meeting the due time) for at least a given probability of travel time scenarios. We introduce a sampling-based approach to reduce the size of the problem, and incorporate procedures that maintain chance constraint feasibility under this framework. Computational experiments reveal that our approach yields routing decisions that remain immune against travel time uncertainty.

Collaborative Optimization of Rolling Stock Allocation and Timetable Coordination in a Multi-Modal Rail Network: MILP Formulation and Decomposition-Based Algorithm – JI-ATENG YIN, BEIJING JIAOTONG UNIVERSITY – Lixing Yang, Simin Chai

The transportation system is shifting towards a shared mobility ecosystem, emphasizing resource aggregation and coordination among different transport modes to enhance service quality and encourage passengers to shift from private cars to public transit. This study focuses on the coordination between metro and rail networks in a city, to alleviate congestion at transfer hubs and to decrease the travel time of passengers. Given the “tide-like” transfer passengers between rail and metro hubs, we investigate the potential of reserving a number of rolling stocks, strategically allocated to the metro trains with the highest demand to avoid over-congestion. We present a mixed integer linear programming (MILP) model to formulate this problem, with decision variables including rolling stock allocation and coordinated schedules for both metro and rail networks, by considering the time-varying demand of passengers. The objectives aim to minimize the passenger travel/transfer time and operational costs for managers. Due to the computational complexity in solving real-world instances, we analyze the mathematical properties and propose an exact decomposition-based solution algorithm. Our algorithm is similar to a constraint generation method, which reformulates the original problem into a relaxed master problem (corresponding to rolling stock allocation and timetabling) and a series of independent subproblems (corresponding to passenger flows). We provide optimality conditions and prove that the subproblems can be solved by an analytical procedure. According to these properties, we further propose a set of feasibility cuts and we prove that the new cuts provide tighter bounds in comparison with traditional Benders cuts. We test our integrated approach and solution algorithms on real-world instances from the Beijing railway network. The results show that our solution approach can obtain near optimal solutions, while the commercial solvers cannot even return feasible solutions in real-world instances. Compared to the current plan used in practice, our approach by combining rolling stock allocation and timetable coordination can reduce the passenger transfer waiting time by nearly 40%.

A Reinforcement Learning Approach for Dynamic Rebalancing in Bike-Sharing Systems – JIAQI LIANG, POLYTECHNIQUE MONTREAL – Sanjay Dominik Jena, Defeng Liu, Andrea Lodi

Bike-Sharing Systems provide eco-friendly urban mobility, contributing to the alleviation of traffic congestion and to healthier lifestyles. Efficiently operating such systems and maintaining high customer satisfaction is challenging due to the stochastic nature of trip demand, leading to full or empty stations. Devising effective rebalancing strategies using vehicles to redistribute bikes among stations is therefore of uttermost importance for operators. As a promising alternative to classical mathematical optimization, reinforcement learning is gaining ground to solve sequential decision-making problems. This paper introduces a spatio-temporal reinforcement learning algorithm for the dynamic rebalancing problem with multiple vehicles. We first formulate the problem as a Multi-agent Markov Decision

Process in a continuous time framework. This allows for independent and cooperative vehicle rebalancing, eliminating the impractical restriction of time-discretized models where vehicle departures are synchronized. A comprehensive simulator under the first-arrive-first-serve rule is then developed to facilitate the learning process by computing immediate rewards under diverse demand scenarios. To estimate the value function and learn the rebalancing policy, various Deep Q-Network configurations are tested, minimizing the lost demand. Experiments are carried out on various datasets generated from historical data, affected by both temporal and weather factors. The proposed algorithms outperform benchmarks, including a multi-period Mixed-Integer Programming model, in terms of lost demand. Once trained, it yields immediate decisions, making it suitable for real-time applications. Our work offers practical insights for operators and enriches the integration of reinforcement learning into dynamic rebalancing problems, paving the way for more intelligent and robust urban mobility solutions.

4c - TB952 - 512E - **Scheduling and manufacturing 1** Chair: MAURÍCIO GONÇALVES 

Growing a stochastic subtree for mixed integer linear programming under uncertainty – ZOE FORNIER, ENPC – Vincent Leclère, Bernardo Freitas Paulo Da Costa, Merve Bodur

We consider mixed-integer stochastic linear programs, which, under some Markovian assumptions, can be solved using Dynamic Programming principles. Were the problem convex, we could use the Stochastic Dual Dynamic Programming (SDDP) algorithm, shown to be efficient in solving large-scale problems. Unfortunately, relaxing all integrality constraints performs poorly. Moreover, if the problem is too large e.g., with more than a few time steps, the extensive formulation becomes intractable. In order to strike a balance between a complete relaxation and the exact problem, we propose to relax partially the integrality constraints. To take advantage of the structure of the problem, if we relax the integrality constraints in a node, then we also relax them in all its descendants: this whole section of the initial scenario tree can be approximated using SDDP. Then, we solve the extensive formulation of a stochastic sub-tree T , where we keep the integrality constraints in the nodes in T , while the leaves of T leverage the SDDP algorithm, to solve the problem. Naturally, the problem formulation gets more complex as the size of T gets larger. Further, the choice of nodes in T might be decisive to get a good lower-bound fast, but more importantly better first-stage integer decisions. Therefore, we propose different heuristics to iteratively grow the sub-tree T , inspired by branching methods in integer programming. We then compare the obtained strategies tested on different simple examples inspired by production problems with either shared resource constraints, minimum uptime/downtime constraints or counters constraints.

Stochastic optimization of bus schedules – LÉA RICARD, EPFL – Andrea Lodi, Guy Desaulniers, Louis Martin Rousseau

The vehicle scheduling problem (VSP) is one of the sub-problems of public transport planning. It aims to minimize operational costs while assigning exactly one bus per timetabled trip and respecting the capacity of each depot. Public trans-

port planning is subject to various endogenous and exogenous causes of uncertainty, notably affecting travel time and energy consumption. Despite the uncertainties involved, the VSP and its variants are usually solved deterministically to address trackability issues. However, considering deterministic travel time in the VSP can compromise schedule adherence, whereas considering deterministic energy consumption in the electric VSP (E-VSP) may lead to solutions with sub-optimal true costs (including recourse costs and the cost of ownership of battery electric buses). This presentation proposes a methodological framework aimed at integrating uncertainties, specifically travel time and energy consumption uncertainties, into the VSP. Three distinct stochastic, data-driven mathematical models and branch-and-price algorithms are introduced to address two variations of the problem: the multi-depot VSP (MDVSP) and the electric VSP (E-VSP). The objective is to find bus schedules that offer a good trade-off between operational costs and service reliability, as well as between operational costs and battery degradation in electric buses.

Integrated lot sizing and blending problems under demand uncertainty – MAURÍCIO GONÇALVES, UNESP IBILCE EXCHANGE PHD RESEARCHER AT HEC MONTRÉAL – Raf Jans, Silvio Araujo

This study addresses a production planning problem considering end products that are obtained through the blending of components. We approach the problem using mathematical programming formulations assuming integrated lot sizing decisions for components purchase and end products production over a finite planning horizon. The objective is to minimize setup, production, purchase, inventory, and lost sales costs, while incorporating constraints related to inventory balance, machine capacity, and quality standards. We extend the deterministic approaches explored in the literature by incorporating demand uncertainty through a two-stage stochastic programming formulation. The proposed formulation is designed to identify stochastic solutions based on the static uncertainty strategy derived from stochastic lot sizing problems. Our solution methodology encompasses an approximation procedure using finite samples to estimate bounds and identify stochastic candidate solutions for the problem. Enhancements are explored for the optimization routines, including adjustable strategies by fixing the value of some first-stage variables based on multiple candidate solutions, Benders decomposition techniques employed with a commercial solver, and specialized algorithms. Finally, we conduct a numerical evaluation of our proposed approaches through computational experiments using test instances based on randomly generated data. This research contributes insights into using stochastic lot sizing strategies for addressing demand uncertainty in integrated production planning and blending problems.

4c - TC953 - 513D - **Scheduling and manufacturing 2** Chair: KENZA ALIOUI 

Scheduling Cybersecurity Mitigations to Delay Attacker Projects – ASHLEY PEPER, UNIVERSITY OF WISCONSIN MADISON – Jim Luedtke, Laura Albert

We study the problem of deploying cybersecurity mitigations

to delay the completion times of a set of attacker projects. Existing literature studies how to select a portfolio of mitigations subject to a budget constraint to optimally delay such attack projects, but lacks consideration of the time it takes to implement the mitigations. We assume attackers are working on their projects at the same time the defender is working to complete mitigations to delay those attacks. Our proposed model captures the importance of scheduling mitigations, since to be useful they must be completed before an attacker exploits a vulnerability. This paper introduces an integer programming model of this problem that combines a defender's mitigation scheduling problem with the attackers' project completion problems. We consider alternative solution methods and examine the benefit of the proposed model.

Scheduling ISMP 2024 – MATHIEU GERVAIS DUBÉ, POLYTECHNIQUE MONTRÉAL – Antoine Legrain, Alain Hertz, Sébastien Le Digabel

Researchers around the globe participate to the International Symposium on Mathematical Programming to set forth their latest results in mathematics, algorithms, computation, and modeling. As of March 2024, the 2024 edition held in Montréal, has around 1000 participants registered and includes around 600 talks. Building a reasonable schedule with these figures cannot be done by hand. One has to wonder: should the assignment of its talks using mathematical programming be pipelined?

The assignment of the ISMP's presentations is modeled using an enumeration of promising aggregations of presentations which are formed by pre-processing common keywords. A bonus scheme is used to induce an order on the assignation of the aggregations for a given session. The suggested streams from the participant's submissions are mined to select possible streams for given keywords. The modeled problem is solved using the Gurobi Optimizer.

Inherent constraints of conferences are considered such as unallowed parallel presentations for the same speaker, session chair assignation, participant's time slot unavailabilities, and session's filling requirements.

Finally, a re-optimization process is designed to ensure that participant constraints communicated after the publication of the first schedule affect the least amount of assignments.

Open shop scheduling problem with reentrance, delays, and no-wait constraints – KENZA ALIOUI, AOTI ESG UQAM – Karim Amrouche, Mourad Boudhar

In this study, we consider a new scheduling problem which focuses on a two-machine reentrant open shop with no-wait and exact delays constraints, where each job should go back to the first machine, i.e. reentrance, to execute a second task after a fixed time lag that separates the two operations of the first machine. On the other hand, a job should start its execution on the second machine at its completion time on the first machine, and vice versa, this scenario enforces a no-wait constraint between the two machines. Our objective is to achieve the lowest possible total completion time, i.e. minimizing the makespan. We initially prove that this scheduling problem is NP-hard in the strong sense, indicating that finding an optimal solution is computationally demanding. However, we identify some specific cases that can be solved in a polynomial time. For the

general problem, we introduce three lower bounds to estimate the minimum makespan and propose two primary resolution methods: a new heuristic approach (H1) and a genetic algorithm (GA). Additionally, two different hybridizations are proposed between H1 and GA. Computational experiments have been conducted in order to assess the performance of the proposed algorithms.

4c - ThA954 - 512G - **Scheduling and manufacturing 3** Chair: PAUL PATIENCE

Fair Assortment Optimization for an Online Fashion Retailer in the Presence of Size Substitution Effect – ALIREZA ALAVI, PHD STUDENT – Javad Nasiry, Mehmet Gumus

Size stockouts in online fashion retailing occur when an item is not available in one of the sizes. In the size stockout cases, customers may opt to purchase a neighboring size (larger or smaller) rather than switching to a different style or leaving the store empty-handed. This customer choice pattern is known as the size substitution effect, and it has recently been revealed that it profoundly impacts customers' decisions. Size stockouts may also trigger fairness concerns among customers. In recent years, there has been substantial coverage of discrimination against plus-size customers as they cannot find the items they like in suitable sizes. Such concerns are increasingly important for retailers as "social inclusivity" gains traction among customers. Therefore, our goal in this research is to develop assortment optimization algorithms for an online fashion retailer in a static setting that, while considering the size substitution effect, are also fair toward different size groups. We model this situation as a zero-sum game between the retailer and an adversary (nature) and then develop an exact algorithm and an efficient approximation algorithm to find the Nash equilibrium of this game. After calibrating our setup with the coefficients coming from a real-world dataset, we ran a series of simulations to evaluate the effectiveness of our approximation algorithm. Among other things, our results show that it is possible to drastically reduce unfairness with a minuscule drop in the retailer's revenue using our algorithm.

A Nested Logic-Based Benders Decomposition for the Multi-Shop Car Resequencing Problem with A Painted Body Storage – XINYI GUO, TSINGHUA UNIVERSITY – Jean François Côté, Zhang Canrong, Miao Lixin

The cost of producing diverse cars depends on the sequence in which they are arranged in the body shop, paint shop, and assembly shop. Before entering the downstream assembly shop, the upstream car sequence shared by the body shop and paint shop is readjusted by a physical buffer named the painted body storage (PBS), which consists of several first-in-first-out lanes. The car resequencing problem (CRSP) addressed in this paper requires determining the upstream and downstream sequences and the car-to-lane assignment to minimize the total costs of the three shops. We propose a nested logic-based Benders decomposition (NLBBD) method with three levels, where the upstream sequence is determined in the first level by assigning each car a body and a color. By determining the configuration and the downstream sequence of each car, the cars are rearranged in the downstream sequence in the second level. A

feasible assignment of cars to lanes is sought in the third level to respect this sequence change. We provide two shortest-path problem reformulations for the first level, which are applicable to general sequencing problems that aim to minimize the number of changeovers. Based on the first reformulation, the classical logic-based Benders decomposition iterative procedure is transformed into a k shortest simple path problem. We verify the corresponding stopping criteria and the existence of the global optimal solution. A lower bound, three valid inequalities, and a math-heuristic method are also proposed to enhance our NLBBD. In our CRSP, the number of cars to be resequenced is not limited by the PBS size. Computational results show that our NLBBD can handle real-world cases and instances of up to 120 cars in one hour, about ten times more than previous studies. A sensitivity analysis from three perspectives is performed to provide some managerial insights.

Optimizing a cellular solid compression problem via the NOMAD blackbox optimizer – PAUL PATIENCE, POLYTECHNIQUE MONTRÉAL – Charles Audet, Bruno Blais

Cellular solids are porous materials used in sandwich panels, heat exchangers, tissue engineering scaffolds and catalysts. Advances in additive manufacturing have permitted the production of triply periodic minimal surface (TPMS)-like cellular solids, which have properties uniquely suited to these applications and have thus been the focus of recent research. Alas, the field of cellular solid design trails behind that of mathematical optimization — much of the literature relies on variants of grid search. We are the first to apply the MADS blackbox optimization algorithm, as implemented by the NOMAD optimizer, to cellular solid design. We have developed a cellular solid generator suitable for use as input to NOMAD, and a blackbox mimicking a problem in the literature consisting of the compression of a block composed of a cellular solid, with the goal of maximizing the cellular solid's relative Young's modulus. We will present the cellular solid generator as well as the results we obtained from applying NOMAD to the compression problem and a comparison with those from the literature. It is our hope that this will pave the way for researchers in cellular solid design to employ more advanced methods of mathematical optimization.

4c - FA955 - 512E - **Scheduling and manufacturing** 4 Chair: FARIN RASTGAR AMINI ←

Daily production scheduling method for mixed-model assembly line and different workload models – TAZU NOMOTO, HITACHI, LTD. –

We developed a daily production schedule method for mixed-model assembly, in which multiple models with different workload are produced on a single line. In mixed-model assembly, it is necessary to determine the production ratio and production order considering two objectives, on-time shipping and production efficiency. In order to consider the objectives, divide the production planning problem into two procedures: (1) selecting the daily mix-flow ratio from the production ratio pattern using mixed integer programming, and (2) classifying production lots based on the workload and determining the production order so that the production of lots of the same classification

is leveled. In this presentation, we show the contents of the method and the evaluation results.

Flexible job shop scheduling problem with sequencing flexibility: constructive heuristics and local searches considering position-based learning effect – DEBORA PRETTI RONCONI, UNIVERSITY OF SÃO PAULO PRODUCTION ENGINEERING DEPARTMENT – Kennedy A.G. Araujo, Ernesto Birgin

This paper addresses the flexible job shop scheduling problem with sequencing flexibility and position-based learning effect. Unlike the classical case where a total order is imposed, in this variant of the flexible job shop scheduling problem, the precedence constraints of the operations comprising a job are given by an arbitrary directed acyclic graph. For this problem, a mixed integer programming model is first introduced, and constructive heuristics are introduced to provide an initial solution for the model's solvers. Next, local search strategies are examined. We show that the classical strategy of relocating only the operations that are part of the critical path can lose better quality neighbors, contrary to what happens in the case where there is no learning effect. Consequently, we analyzed an alternative type of neighborhood reduction that eliminates only neighbors that are no better than the current solution. Furthermore, we propose a neighborhood cut, and experimental results show that this greatly shrinks the neighborhood, bringing efficiency with minimal loss of effectiveness. Numerous numerical experiments are conducted. The proposed methods, as well as the instances and solutions found, are freely available.

Solving an integrated lot sizing and preventive maintenance planning for a single machine – HARCÈNAGE DANSOU, ESG UQAM – Matthieu Gruson, François Lamothe, Julien Legavre

In manufacturing industries, production and maintenance operations are closely linked but are generally planned and optimized separately. Planning preventive maintenance operations makes it possible to determine the appropriate periods for carrying out maintenance upstream of production. Then, production planning consists of determining the quantity and timing of production of several items over a time horizon. This study proposes new mathematical formulations as mixed-integer linear programs for a problem integrating preventive maintenance planning into the Capacitated Lot Sizing Problem (CLSP). One of these formulations (M1) is based on the redefinition of the variables linked to preventive maintenance planning and another is a Dantzig-Wolfe decomposition formulation (M2). We demonstrate that the formulation M1 gives a relaxation equivalent to the column generation based on M2. Subsequently, we use a relax-and-fix heuristic combined with fix-and-optimize to solve this problem. Relax-and-fix is used to build an initial solution and fix-and-optimize allows us to improve the built solution. Computational experiments evaluate the performance of the heuristics and the models solved with Gurobi. The results show that the solutions obtained by the heuristic are close to those obtained by the models in terms of gap and in a reduced time.

Data Mining-Driven Shift Enumeration for Accelerating the Solution of Large-Scale Personnel Scheduling Problems – FARIN RASTGAR

AMINI, POLYTECHNIQUE MONTREAL – Daniel Aloise, Claudio Contardo, Guy Desautniers

This study addresses large-scale personnel scheduling problems in the service industry by combining mathematical programming with data mining techniques to enhance efficiency. The studied problem aims at efficiently scheduling skilled employees over a one-week planning horizon, minimizing costs while meeting diverse job demands. In service industries, shift planning is intricately tied to customer presence, leading to a multitude of potential shifts and to a difficult optimization problem that cannot be easily solved using a commercial mixed-integer programming solver. Nevertheless, these problems are categorized as recurrent problems, where distinct instances share common characteristics and solution structures that differ only in a few parameters over time. Consequently, we propose to use a data mining technique, namely, the k-nearest neighbors algorithm, to expedite the solution process while upholding solution quality. In particular, we suggest using schedules of past solutions to reduce the problem size. Specifically, for an upcoming instance, we identify similar historical instances and streamline the enumeration of shifts to align with the comparable historical instances' schedules. This approach allows us to effectively address the problem using a commercial solver within a reasonable timeframe while preserving solution quality. Significantly, our methodology offers decision-makers the flexibility to determine the extent to which they wish to scale down the problem. Our experiments, conducted on a total of 80 instances with up to 12 jobs and 190 employees, yield an average removal of 85.5% of decision variables. This resulted in a noteworthy average speedup factor of 15.5, with a marginal average cost increase of 1.2%.

4d - MB119 - 516E - **Optimization models for electric vehicle charging problems 1** Chair: MIGUEL ANJOS

Strengths of the Formulations of the Fuel Refuelling Location Problem – NAGISA SUGISHITA, UNIVERSITÉ DE MONTRÉAL – Margarida Carvalho, Ribal Atallah

The Fuel Refueling Location Problem (FRLP) consists of finding the optimal siting of charging stations for electric vehicles (EVs), specifically considering the limited travel range of EVs. This problem becomes particularly relevant when the goal is to facilitate long-distance travel using EVs. A variant of the FRLP, known as the Deviation Fuel Refueling Location Problem (DFRLP), accounts for the possibility that drivers may deviate from their preferred routes to ensure sufficient charging for completing their trips. In this presentation, we study formulations of the FRLP and DFRLP used in the literature, examining their strengths. Furthermore, we explore approaches to tighten the formulations.

A Constrained Traffic Assignment Problem Model for Congestion Analysis of Electric Vehicle Fast Chargers – ISMAIL SEVIM, UNIVERSITE DE MONTREAL – Margarida Carvalho, Ribal Atallah

In this study, we model the congestion in the electric vehicle (EV) fast chargers as a traffic assignment problem (TAP) with recharging constraints. The model aims to find the most congested stations and to further analyze them to increase the service level. We also work on a gradient projection algorithm

implementation for the problem. As the case study, we analyze the Province of Quebec, Canada in a collaboration with Hydro-Quebec, the administration responsible for establishing and maintaining EV charging stations. Preliminary computational results revealed that the proposed model can estimate the congestion in the stations accurately.

Optimal Electric Vehicle Charging with Dynamic Pricing, Customer Preferences and Power Peak Reduction – GAËL GUILLOT, INRIA – Luce Brotcorne, Miguel Anjos

We consider a provider of electric vehicle charging that operates a network of charging stations and uses time-varying pricing to maximize profit and reduce the impact on the electric grid. We propose a bilevel model with a single leader and multiple disjoint followers. The customers (followers) makes decisions independently from each the other. The provider (leader) sets the prices for each station at each time slot, and ensures there is enough energy to charge. The charging choice of each customer is represented by a combination of a preference list of (station, time) pairs and a reserve price. The proposed model takes thus into accounts for the heterogeneity of customers with respect to price sensitivity and charging preferences. We define a single level reformulation based on the reformulation for the rank pricing problem. Computational results highlight the efficiency of the new reformulation and the impact of the model on the grid peak.

Optimizing Electric Vehicle Charger Locations for Ride-hailing Services through Discrete Simulation-based Optimization – TOMMASO SCHETTINI, HEC MONTRÉAL – Fausto Errico, Jorge E. Mendoza, Carolina Osorio

The problem of planning the installation or extension of EV charging infrastructure has attracted considerable research attention in the scientific community. One interesting thread in this literature relates to the design of charging networks for supporting a privately owned fleet of EVs. In this paper, we consider the problem of optimizing the charging and parking infrastructure for supporting an autonomous fleet of ride-hailing vehicles with the objective of maximizing its expected operational profit. The dynamic and stochastic nature of ride-hailing services makes it challenging to measure the expected operational profit of a particular charger configuration. To overcome this issue, we make use of a simulator to estimate the expected operational profit associated with each solution and formulate the problem as a discrete simulation-based optimization (DSO) problem. We implemented a novel DSO method based on the classical algorithm of Nested Partitions (NP). Our proposed approach uses the mathematical structure of the feasible space of the problem to construct auxiliary integer linear programming formulations to inform partitioning and sampling. We discuss a strategy for biasing the proposed sampling strategy, so as to improve the average quality of the sampled solutions, and boost the overall performance. To illustrate the effectiveness of our method, we consider a case study in the context of New York City and compare the proposed strategy to NP, and to the solutions obtained through a non-simulation-based heuristic.

4d - MC337 - 513C - **Optimal transition of energy systems/power grids 2** Chair: MICHEL DENAULT

Modeling of a Hydrogen Economy Using Integrated Assessment Approaches – OLIVIER BAHN, GERAD, HEC MONTRÉAL – Kamyar Amir-moeini, Sara Ghaboulia Zare, Najmeh Neshat

Within the needed global energy transition, hydrogen presents a potential avenue for achieving deep decarbonization. To assess the decarbonization potential of a hydrogen economy, one may use an integrated assessment approach, a comprehensive framework that incorporates insights across diverse sectors like energy, land use, and the broader economy, along with their associated greenhouse gas emissions and their impact on the global climate. Integrated assessment models (IAM) are mathematical tools to conduct an integrated assessment. This presentation proposes a survey of 12 distinct family of IAM from 50 different studies, to assess the potential of a hydrogen economy for different scenarios developed within the IPCC Sixth Assessment Report WGIII. This presentation also reports on our ongoing efforts at GERAD to model hydrogen production pathways in our own integrated assessment model, called ADMERGE.

Robust Pathways to H2 Adoption: Leveraging ETEM-YUL and ETEM-CAN Frameworks – SARA GHABOULIAN ZARE, UNIVERSITY OF MONTREAL – Frederic Babonneau, Olivier Bahn, Alain Haurie, Normand Mousseau, Martin Trepanier

To meet global climate objectives, it is essential to reduce reliance on fossil fuels. Hydrogen (H₂) has emerged as a crucial component in this shift, supported by the federal and Québec governments' ambitious goals for emissions reduction and carbon neutrality. This study investigates the integration of H₂ technologies within the Greater Montréal (GM) region, aiming to achieve significant greenhouse gas (GHG) emissions reductions and aid the transition to carbon neutrality by 2050. Focusing on hard-to-abate sectors, the research enhances the Energy-Technology-Environment Model (ETEM) for the GM region (ETEM-MTL) to assess hydrogen contribution to decarbonization efforts, particularly in mobility and industry. The work develops optimized H₂ pathways to examine sector synergies and evaluate GHG emissions across hydrogen supply chain.

4d - TA117 - 513D - **Advances in generation, transmission and distribution modelling** Chair: ALEXANDRE BESNER

RALPH - Solving the unit commitment problem under hydraulic and transient stability constraints – ALEXANDRE BESNER, HYDRO QUÉBEC – Alexandre Blondin Massé, Abderrahman Bani, Mouad Morabit

Hydro-Québec is a vertically integrated utility that produces, transmits and distributes most of the electricity in the province of Québec. The power grid it operates has a particular architecture created by important hydroelectric dams located far North and an extensive 735kV transmission grid that allow

the generation to reach the majority of the load that is located thousand of kilometers away in the southern region of Québec. The specificity of the grid has forced Hydro-Québec to develop important measures to monitor the stability of the grid. Those stability limits are derived from multiple stability phenomena like angular stability, frequency stability or voltage stability. Those phenomena are highly non-linear but have to be taken into account when doing generation planning and operation. Since generation planning and operation algorithms normally use mixed integer linear programming formulation, Hydro-Québec had to develop novel approach to model stability limits into its tools for planning and operating the generation. RALPH software is a tool specifically developed to solve the unit-commitment problem under hydraulic and transient stability constraints.

Tight Big-Ms for Optimal Transmission Switching – SALVADOR PINEDA, UNIVERSIDAD DE MÁLAGA –

This work addresses the Optimal Transmission Switching (OTS) problem in electricity networks, which aims to find an optimal power grid topology that minimizes system operation costs while satisfying physical and operational constraints. Existing methods typically convert the OTS problem into a Mixed-Integer Linear Program (MILP) using big-M constants. However, the computational performance of these approaches relies significantly on the tightness of these big-Ms. In this work, we propose an iterative tightening strategy to strengthen the big-Ms by efficiently solving a series of bounding problems that account for the economics of the OTS objective function through an upper-bound on the generating cost. We also discuss how the performance of the proposed tightening strategy is enhanced if reduced line capacities are considered. Using the 118-bus test system we demonstrate that the proposed methodology outperforms existing approaches, offering tighter bounds and significantly reducing the computational burden of the OTS problem.

Application of operational research in Hydro Quebec distribution system planning – SANJA BAJIC, HYDRO QUÉBEC – Mihajlo Curcic

Hydro Quebec distribution network faces many planning and operational challenges that can be addressed using operational research. The focus is to adequately prioritize the network reinforcements and to improve the system reliability, i.e., reduce the interruption frequency and duration. In Hydro Quebec research center, we try to address these objectives through the innovation project SCÉNARIO. We are studying the some of the common distribution system optimization problems, including: - Controlled recharge of electrical vehicles - Residential heating control through aggregator - Staggered client reconnection after an outage - Reliability-based positioning of feeder sectionalizers - Optimal micro-grid control - Distribution network capacity expansion One of the particularities that Hydro Quebec network has compared to other utilities is the winter Cold Load Pick-up (CLPU). Due to the large and growing prevalence of electric heating and a significant projected EV load, the CLPU has become the most important criterion for the distribution system planning. Therefore, all the above-mentioned optimization problems must be casted considering the CLPU and the associated customer load models.

Optimizing physical and financial assets us-

ing SDDP – YVES ALAIN MBEUTCHA, HYDRO QUÉBEC, CANADA – Jasson Pina, Nathalie Letendre, Abdelghani Hammadia

Hydro-Québec (HQ), one of the world's largest hydropower producers, generates electricity to supply the Québec market and sells its excess output on wholesale markets. At HQ a variety of tools and methods have been implemented to jointly analyze operational scheduling and contract management. Stochastic Dual Dynamic Programming (SDDP) is one of the algorithmic solutions used to determine the operating policies of our extensive energy infrastructure. These operational policies encompass among others the evaluation of load factor contracts, commonly known as flexible contracts. In these agreements, terms regarding price, energy, and maximum power are predefined, yet usage remains adaptable. For instance, contracts such as CHPE (Champlain Hudson Power Express) and NECEC (New England Clean Energy Connect) are integrated into the SDDP framework represented as state variable. These contracts impose restrictions on both annual energy volume and instant power output. This decision support system allows us to better evaluate our risk associated to inflow uncertainty.

4d - TA956 - 513C - **Energy and Environment**

1 Chair: HONGYU ZHANG



Differentially Private Distributed Optimization with Constraints – MINSEOK RYU, ARIZONA STATE UNIVERSITY – Kibaek Kim

In distributed optimization (DO), multiple agents collaborate to minimize a global objective function by iteratively exchanging intermediate local solutions. Despite not directly transmitting locally stored data, it is feasible to reconstruct the data from these exchanges. To tackle this issue, the integration of differential privacy (DP) has been crucial, ensuring the privacy of each individual's data within the dataset. Within a DPDO framework, noisy local solutions are repeatedly communicated to safeguard data privacy, which may compromise solution quality due to the trade-off between privacy and accuracy. In this talk, we will first demonstrate that the existing DP mechanism exacerbates the trade-off when constraints are present. Subsequently, we will introduce our DP mechanism that significantly improves this trade-off. The relationship between our proposed DP mechanism and existing one will be presented, as well as our privacy and convergence analyses. Furthermore, we will showcase the performance of our approach numerically, illustrating its efficacy in applications such as federated learning and distributed optimal power flow.

Uniform Convergence of Kernel Estimators – PAUL DOMMEL, TU CHEMNITZ –

Estimating the conditional expectation is a central task in a variety of different optimization topics, including stochastic optimization and optimal control. The kernel estimator associated with kernel ridge regression provides a suitable tool for this objective, converging to the target function with respect to the L^2 -norm. However, to relate the optimizer of the estimated objective with the optimizer of the true objective, uniform convergence of the estimator is crucial.

In this talk, we relate uniform convergence of kernel estimators to convergence in L^2 as well as smoothness conditions of

the corresponding target function. These conditions are given in terms of the decay rate of the associated basis coefficients. We provide an interpolation bound for the uniform norm as well as applications.

An Interval-based Stochastic Unit Commitment Model Under Demand Uncertainty – HOJIN JUNG, SEOUL NATIONAL UNIVERSITY – Jongheon lee, Kyungsik Lee

The unit commitment problem aims to find a minimum-cost on/off status and amount of generation for each generator while satisfying the electricity demand and operational requirements. In this talk, we address the stochastic unit commitment problem where the uncertain demand follows a certain probability distribution and the prediction error is period-wise independent. In this situation, the two-stage stochastic optimization models with a finite number of demand scenarios have been widely used, where the on/off status is decided in the first stage and the amount of generation is in the second stage. However, they often suffer from excessive computational burden as the number of scenarios increases. In addition, operations based on the solutions from the models may be ineffective where the demand is sequentially realized. To overcome these drawbacks, we propose an interval-based stochastic unit commitment model, where a range of the amount of generation, called “interval”, is determined in advance for the purpose of reliable operation. Furthermore, the proposed model can avoid the use of a large number of scenarios by using the upper and lower bounds on the expected second-stage cost. We show the effectiveness and efficiency of the proposed model through the computational experiments in the context of the actual operation.

Integrated energy system planning under short-term and long-term uncertainty: Modelling and algorithms – HONGYU ZHANG, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY – Ignacio Grossmann, Brage Rugstad Knudsen, Ken McKinnon, Rodrigo Garcia Nava, Asgeir Tomasgard

We propose the REORIENT (REnewable resOuRce Investment for the ENergy Transition) model for energy systems planning with the following novelties: (1) integrating capacity expansion, retrofit and abandonment planning, and (2) using multi-horizon stochastic mixed-integer linear programming with short-term and long-term uncertainty. We apply the model to the European energy system considering: (a) investment in new hydrogen infrastructures, (b) capacity expansion of the European power system, (c) retrofitting oil and gas infrastructures in the North Sea region for hydrogen production and distribution, and abandoning existing infrastructures, and (d) long-term and short-term uncertainty. We utilise the structure of multi-horizon stochastic programming and propose enhanced Benders decomposition methods to solve the model efficiently. We propose: (1) stabilising Adaptive Benders with the level set method and adaptively selecting the subproblems to solve per iteration for more accurate information, (2) a centred point stabilisation approach when the level method problem is hard to solve, and (3) dynamic level set management to improve the robustness of the algorithm by adjusting the level set per iteration.

We first conduct a sensitivity analysis on retrofitting costs of oil and gas infrastructures. We then compare the REORI-

ENT model with a conventional investment planning model regarding costs and investment decisions. Finally, four algorithms are implemented for solving LP instances with up to 1 billion variables and 4.5 billion constraints, and two algorithms are implemented for MILP instances with high degeneracy. The results show that: (1) when the retrofitting cost is below 20% of the cost of building new ones, retrofitting is economical for most of the existing pipelines, (2) compared with a traditional investment planning model, the REORIENT model yields 24% lower investment cost in the North Sea region, and (3) for a 1.00% convergence tolerance, the enhanced Benders is up to 6.8 times faster than the reference algorithm for MILP instances, and is up to 113.7 times faster than standard Benders and 2.14 times faster than unstabilised Adaptive Benders for LP instances. Also, for a 0.10% convergence tolerance, the enhanced Benders is up to 45.5 times faster than standard Benders for LP instances, and unstabilised Adaptive Benders cannot solve the largest instance to convergence tolerance due to severe oscillation. Finally, the dynamic level set management makes the algorithms more robust and is very helpful for solving large problems.

4d - TB292 - 513C - **Optimization models for electric vehicle charging problems 2** Chair: MIGUEL ANJOS

Improving Cloud Service Sustainability with Bilevel Optimization under Demand Uncertainty – NATHALIA WOLF, INRIA – Luce Brotcorne, Bernard Fortz

Cloud computing's rapid growth, driven by escalating global data traffic, presents challenges in energy efficiency due to the underutilization of servers. We propose a cloud sharing system through a mixed integer bilevel optimization model under demand uncertainty. This model features two distinct types of customers in the public cloud, namely, long-term consumers with monthly subscriptions and short-term consumers who want on-demand access. In this model, pricing strategies at the upper level dynamically adjust to incentivize short-term consumers to utilize virtual machines (VMs) during periods of low-energy prices. Additionally, these strategies incorporate rewards for long-term consumers, encouraging them to share their VMs in idle periods. We consider a Distributionally Robust Optimization (DRO) approach to address uncertainties in the consumer bids, enhancing decision-making resilience against variability in willingness to pay.

The Electric Bus Rostering and Charging Scheduling Problem with Uncertain Energy Consumption: a Two-Stage Stochastic Programming Approach – KYRA ANN D'IGNAZIO, HEC MONTRÉAL –

Electric vehicles are one of the promising technologies that are being used in the transportation sector to help reduce the negative impacts associated with traditional technologies. In order to integrate electric vehicles into public transportation, this study focuses on the examination of how energy uncertainty in electric battery consumption impacts operating costs associated with charging decisions. This research formulates the energy uncertainty challenge as a two-stage stochastic programming model by employing mixed-integer linear programming. The model optimizes bus route assignment and charging deci-

sions such as when to charge, while accounting for the potential deviations in energy requirements that may require a recourse action in certain scenarios. The results establish a correlation between the level of uncertainty in consumption and the difficulty to solve the problem (i.e., CPU time). This research contributes to the realm of electric vehicle fleet management by providing a comprehensive methodology to optimize bus route assignments and charging protocols considering the real-world variability in energy consumption.

Integrated Location, Sizing and Pricing for Electric Vehicle Charging Stations – MIGUEL ANJOS, UNIVERSITY OF EDINBURGH – Luce Brotcorne, Alemseged Weldeyesus, Ikram Bouras, Clémence Alasseur, Riadh Zorgati

We present a bilevel optimization model to support decision-making about locating, sizing and pricing of electric vehicle (EV) charging stations by taking into account the behaviour of EV users. This is done by adopting a preference-list or rank-based approach that characterizes users by a set of distinct ordered sets of predefined preferences or products. In the upper-level, the charging service provider is in charge of making decisions on the location, size and pricing of the charging stations to maximize its profit. In the lower level, EV users select their first available charging station from their preference list. We solve the bilevel optimization problem using a KKT-based single-level reformulation and provide extensive managerial insights on randomly generated set of problem instances.

4d - TC142 - 513C - **Optimal transition of energy systems/power grids 1** Chair: MICHEL DENAULT

Multi-value benefits of transmission expansion in Western Canada – MADELEINE SEATLE, UNIVERSITY OF VICTORIA – Madeleine Mcpherson

It has been proven that expanding transmission capacity facilitates the achievement of net-zero targets by improving variable renewable energy (VRE) utilization (IEA, 2023). However, transmission expansion plans have traditionally only been assessed on the metrics of operational cost savings and curtailment reduction; this completely disregards the reliability benefits of expanded transmission and the quantification of increased access to high-value renewable potential (ESIG, 2022). A multi-value benefit planning framework has been developed by ESIG (2022) to more holistically assess the value of transmission expansion. In the framework, the metrics used in assessing transmission benefits still include operational cost savings, but also now includes five other metrics: emission reduction, renewable expansion capital cost savings, risk mitigation under uncertain future conditions, resource adequacy analysis, and resilience benefits. Using a mixed-integer optimization model to assess transmission expansion and a linear optimization model to evaluate grid operations, we have applied this multi-value planning framework to transmission corridors that show significant opportunity for expansion under the Canadian Energy Regulations (CER): British Columbia and Alberta, and Saskatchewan and Manitoba. Results indicate that there are significant benefits of expanding transmission in terms of improving the resilience and resource adequacy of the electricity grid, which would have gone unquantified with traditional

transmission expansion assessments. This multi-benefit analysis can be used to quantify the value of interprovincial transmission and analyse the specific benefits gained through transmission expansion more fully.

References: ESIG. (2022). Multi-Value Transmission Planning for a Clean Energy Future: A Report of the Transmission Benefits Valuation Task Force. Energy Systems Integration Group. <https://www.esig.energy/multi-value-transmission-planning-report> IEA. (2023). Net Zero Roadmap: A Global Pathway to Keep the 1.5 °C Goal in Reach. IEA. <https://www.iea.org/reports/net-zero-roadmap-a-global-pathway-to-keep-the-15-0c-goal-in-reach>

Reinforcement learning and large system operations – FLORIAN MITJANA, HEC MONTRÉAL – Michel Denault

When operating large power systems that include controlled (hydro) and uncontrolled (wind) production, optimizing the use of the reservoirs is the central task. One of the main difficulties is handling the uncertainties associated with water and wind inflows, particularly in a northern region like Quebec where water storage is managed over months, even years.

We model a large part of the region's hydroelectric capacity, including five main reservoirs and their interdependence. The objective is to satisfy load while maximizing profit, with sales in the event of overproduction and purchases in the event of power shortages.

We apply a reinforcement learning (RL) approach, with a simple REINFORCE policy gradient method. Benchmark methods such as Stochastic Dynamic Programming (SDP) or Stochastic Linear Programming (SLP) are limited, as the computational effort increases exponentially with the number of reservoirs (SDP) or with the representativity of the inflow profiles (SLP).

Generation Planning of Renewables Using Stochastic Dual Dynamic Programming: A Case Study on Quebec – SOROOSH SABAEI, PH.D. STUDENT – Michel Denault, Pineau Pierre Olivier

Our research focuses on scheduling electricity generation of hydro and wind power utilizing a multi-stage stochastic optimization approach since the variability of inflows and wind is a major concern. Using Quebec's hydro and wind power networks as a case study, we investigate long-term reservoir management strategies to meet the demand while minimizing the total costs. We applied a state-of-the-art technique Stochastic Dual Dynamic Programming (SDDP) to find the optimal generation policy. We also compare the results with a deterministic policy that utilizes the expected values to find the optimal policy. The last part of our analysis is a numerical investigation of expanded generation capacity. In particular, we examine the trade-offs between the costs of new capacity construction and the reduction of load shedding and power purchasing costs.

4d - WA206 - 513D - **Equilibrium modeling in energy markets** Chair: RYAN ENT ↩

Equilibrium Modeling in Natural Gas Markets: A Theoretical Analysis and a Case Study for Brazil, A Method for Solving Gas Market and Water Market Equilibrium Problems

– STEVEN GABRIEL, UNIVERSITY OF MARYLAND –

We present an abstract view of natural gas markets to offer a theoretical discussion on the existence and prove uniqueness of equilibrium solutions. In the abstract model, each player solves a separate profit-maximization optimization problem, the optimality conditions of which can be concatenated together along with market-clearing conditions to give rise to a mixed complementarity problem. The equilibrium-based approach features a general framework that allows for the analysis of a convex combination of price-taking and price-making strategies for producers with a particular Cournot coefficient. As a new motivation for such an analysis, we apply the model to a gas market in transition as is the case in Brazil, which has undergone significant regulatory changes aimed at making it more open and competitive. By demonstrating a sensitivity analysis of market competition on an illustrative Brazilian gas network, we offer new insights to competitive natural gas market modeling. We also describe a new method to solve the resulting linear complementarity problem and demonstrate its use on both the gas market problem as well as second-order cone complementarity problems in water markets.

When market incompleteness is preferable to market power. Insights from power markets.

– IBRAHIM ABADA, GRENOBLE ECOLE DE MANAGEMENT – Andreas Ehrenmann

Since the liberalization of power markets in Europe, almost all investments have required government subsidies. This demonstrates the failure of short-term markets to send the right price signals to investors, but other market failures have been brought to light, among which is market incompleteness. The literature has consistently shown the benefits of risk-sharing instruments to complete a market and boost investments in the absence of market power. Because of entry barriers and legacy assets, however, some producers today might find themselves in a dominant position with the possibility of abusing the market. This paper investigates the impact of completing a market with financial contracts in the presence of market power. Starting from the benchmark case of a capacity-expansion problem in a perfectly competitive and complete market, we formulate a set of equilibrium stochastic models of risk-averse agents investing in a first stage and then operating and trading electricity in a short-term market with possibly two types of failure: incompleteness and market power. We provide existence results and conduct a thorough numerical application inspired by the French case. Contrary to what is reported in the literature, we highlight conditions under which social welfare is worse off when partially completing the market with contracts in the presence of a price-making incumbent. We also show that a strategic incumbent might reduce the volume of hedging contracts to induce physical scarcity. Finally, using the entropic risk measure, we provide a theoretical proof of this result in the case of a monopoly.

Connection queues and shared connection costs

– EDDIE ANDERSON, IMPERIAL COLLEGE – Beste Akbas

Around the world there are large numbers of generation projects waiting in queues for connection to the grid. The wait times are a significant factor in delaying progress on decarbonisation. Connection costs involve a fixed component shared between the generators who are finally connected, but many proposals in the queue will eventually drop out, as more infor-

mation on costs becomes available. We consider the ways in which entry and withdrawal costs chosen by the system operator can lead to the system optimal set of generation projects both joining the queue and not withdrawing prior to connection.

Analysis of ISO New England's Energy Imbalance Reserve Using Mixed Complementarity Problems – RYAN ENT, UNIVERSITY OF MASSACHUSETTS AMHERST – Golbon Zakeri, Jinye Zhao, Tongxin Zheng

In wholesale electricity markets, operating reserves are contracts for generation capacity sold by electricity providers to provide the power system with flexibility in the case of N-1-1 contingencies. In 2025, the independent system operator of New England (ISO-NE) will introduce a co-optimized day ahead energy and reserve market that contains four reserve products with a call option settlement based on the difference between the real time price of energy and a strike price set by ISO-NE. Part of this market is a new product, the energy imbalance reserve, which will make up the difference between energy cleared on the day ahead market and the system operator's real time load forecast. ISO-NE's goals for this market are to incentivize the generators in their fleet, mainly reliant on natural gas, to invest in advanced fuel and increase energy purchased on the day ahead to a level equal to the expected real time load. To analyze how this market may or may not achieve these goals, we model the simultaneous equilibrium of multiple risk-averse generators participating in day ahead and real time markets for energy with and without the presence of energy imbalance reserve.

4d - WA957 - 513C - **Energy and Environment**
2 Chair: SEYYEDREZA MADANI

Using structural properties to optimize operating strategies of district heating networks under nonlinear constraints – REHLICH LEA, TU DARMSTADT – Marc Pfetsch, Stefan Ulbrich

To reduce carbon emissions the transformation of the heating sector is of great importance. In particular, district heating networks in combination with renewable energy generation and waste heat play a major role. This includes lowering the operating temperatures and transforming the existing networks to decentralized structures. Incorporating this, new operating strategies are needed.

We consider a global optimization approach that aims at finding cost-optimal operating strategies. The resulting problem is based on nonlinear physical equations for describing the network state and binary variables to determine flow directions in the network pipes. The problem is solved with the solver SCIP. Due to the MINLP structure the solving process leads to high computational costs. Therefore, we consider methods that aim at reducing the number of variables and nonlinear constraints, making use of the underlying network structure.

The developed methods are evaluated with numerical results based on a real district heating network.

Obtaining the Convex Hull Formulation for Optimal Investment and Operation of Energy Storage Systems Including Reserves – MAAIKE ELGERSMA, DELFT UNIVERSITY OF TECHNOLOGY

– Germán Morales España, Karen Aardal, Mathijs De Weerd, Juha Kiviluoma, Niina Helistö

Large-scale energy system optimization modeling is becoming an increasingly important field, since we need to harness larger shares of variable renewable energy. The main challenge is to include many detailed models of energy assets, while maintaining computational tractability. For example, energy storage systems have become a promising option to increase power system flexibility, but the standard MILP models that describe optimal investment and operation of these storage units, possibly including the optional capacity to provide up/down reserves, do not scale well. A common practice is therefore to relax the integrality constraints, but this results in LP relaxations that allow simultaneous charging and discharging, while this is not feasible in practice. In this talk, we discuss how we can use MILP theory to improve the tightness of such formulations, and how this can benefit energy system optimization models. Our methodology uses the disjunctive nature of the storage optimization problems to derive the convex hull of its solutions for one time period. When included in multi-period large-scale energy system models, these tighter MILP models are expected to scale better, and their LP relaxations are expected to better prevent simultaneous charging and discharging. We demonstrate these two benefits with illustrative case studies of a unit commitment problem and a transmission expansion planning problem. The improved MILP formulations and LP relaxations can be used for many different types of energy storage systems, as well as transmission lines.

A Convex Model Predictive Control Optimization Model of Active Distribution Network Flexibility Regions – LUIS LOPEZ, PHD STUDENT – Kristen Schell

Distributed energy generation (DER), such as roof-top solar panels, are expected to become prolific in the net zero electricity distribution system of 2050. The intermittency of DERs poses challenges for operating the distribution system. At scale, however, the intermittency of DERs could be leveraged to provide operational flexibility to higher voltage levels of the power system, including the transmission system. This work develops a new optimization model to characterize the flexibility region afforded by DERs in the distribution level, and analyzes opportunities for use at the transmission level. IEEE distribution test systems are used as case studies of the optimization formulation. This research demonstrates the value of flexibility regions and how to operate active distribution networks to optimize that value.

Smart Grids, Smart Pricing: Employing Reinforcement Learning for Prosumer-Responsive Critical Peak Pricing – SEYYEDREZA MADANI, HEC MONTRÉAL – Pineau Pierre Olivier, Laurent Charlin

This paper evaluates the potential effectiveness of Critical Peak Pricing (CPP) as a demand response strategy for peak load shaving in electricity grids, with a specific focus on Quebec's electricity grid. The study employs a unique approach by integrating prosumer behavior into the analysis, accounting for the adoption of distributed energy resources such as photovoltaic panels, batteries, and electric vehicles. Through comprehensive simulations and the application of reinforcement learning algorithms, we analyze the effectiveness of CPP programs, both in mass and targeted offering scenarios. The results reveal that

while CPP is effective in incentivizing load shifting, its efficacy diminishes with increasing prosumer participation, leading to new peaks. To counteract this, we propose targeted, dynamic pricing strategies demonstrating significantly improved performance and extended viability. The study also highlights the influential role of batteries and electric vehicles in peak load reduction, suggesting a need for focused policy and incentive structures.

4d - WB958 - 513C - **Energy and Environment**
 3 Chair: JORGE WESTON

Robust Optimal Design of Renewable Energy Systems within an Energy Storage Process using Iron as Energy Carrier – ELISA CORBEAN, TECHNICAL UNIVERSITY OF DARMSTADT – Stefan Ulbrich

For the current goal of limiting and in the long run eliminating CO₂ emissions, the expansion of renewable energy systems is indispensable. However, even with sufficient green energy at hand, further challenges arise regarding its storage, transport and distribution. Metals such as iron are promising candidates to serve as carbon neutral energy carriers due to the high availability and advantageous transport and storage characteristics. Iron could be used as an alternative to hydrogen, in a storage process in which green hydrogen only represents an intermediate product and is further used as reducing agent in a thermochemical reduction reaction. This way, energy is stored in form of iron resulting from the reduction of iron oxides with hydrogen.

This innovative concept requires the expansion and construction of infrastructure concerning renewable energy systems for green hydrogen production on the one hand, and reduction plants for iron oxide reduction on the other hand. Finding the cost-optimal design of these energy systems and related process components represents an interesting task that can be tackled by optimization techniques. To this end, a mathematical model describing the underlying physical processes is developed. In order to account for the location-specific potential of renewable energy sources and the inherent fluctuations of weather parameters, available weather time-series data is used within a robust optimization approach to obtain robust optimal system designs. Numerical results will be presented to showcase the investigated approach using realistic data for various locations.

A Robust Data-Driven Approach to Characterize Operational Flexibility Zones of Active Distribution Networks – ALEJANDRO ANGULO, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

The substantial growth of distributed generation is driving an increasing need for active distribution networks to adapt to dynamic operating scenarios. This expansion poses important challenges to TSO-DSO coordination for the power grid's efficient operation, given the distribution network's energy management capacity at the connection point. This research presents a strategy based on robust optimization to determine the operational flexibility of active distribution networks. The methodology involves determining dynamic operation charts for both active and reactive power, considering the system's operational constraints and the characterization

of the equipment associated with variable energy resources. The different modes of operation of these devices, based on power inverters, are also included in the analysis. The impact of control strategies based on droop, defined by the IEEE 1547 standard and equivalent policies documented in the literature, is also studied. The uncertainty sets necessary for the proposed robust approach are constructed directly from historical samples based on polyhedral approximation. Extensive computational experiments are performed to validate the proposal, using modified versions of IEEE test cases and considering out-of-sample evaluation to highlight the benefits of the proposed method compared to those documented in the literature.

A Heuristic Adversarial Approach for the Two-Stage Robust AC OPF Problem – JORGE WESTON, FRIEDRICH ALEXANDER UNIVERSITÄT ERLANGEN NÜRNBERG – Kevin Martin Aigner, Frauke Liers, Martin Schmidt

In this work, we present a two-stage robust approach to tackle uncertainty of power loads in the alternating current optimal power flow (AC OPF) problem. Most of the work related to robust versions of the AC OPF problem makes use of convex relaxations or linear approximations to reduce the complexity of the non-convex problem. Moreover, many studies assume the existence of a feasible adaption of the network's state after the uncertainty realization. While the first approach might lead to operational points that are actually infeasible for the AC physics model, the second assumption is also known to be violated in many cases. Here, we model the AC OPF problem under uncertainty as a two-stage robust problem to (i) identify approximate robust AC-feasible operation points under (discrete or continuous) power load uncertainty and (ii) find a linear decision rule to model the adaption of the network's state once the uncertainty manifest itself. We solve the problem via a decomposition approach. Where, we discretize the uncertainty set and use spatial branching within an adversarial approach where we solve the subproblems to global optimality. Preliminary computational results for small test cases from the IEEE library show that the approach can detect robust infeasibility with a small number of iterations. However, solution times increase for computing approximate robust feasible points. From our results, with affine decision rules and box-type uncertainty sets, we notice that protection against the extreme points of the uncertainty set leads to a feasible solution that is robust for the whole uncertainty set.

4d - WC171 - 513C - **Optimization of Energy Communities** Chair: KRISTEN SCHELL

Optimal Participation of Energy Communities in Electricity Markets under Uncertainty. A Multi-Stage Stochastic Programming Approach – ALBERT SOLÀ VILALTA, POLYTECHNIC UNIVERSITY OF CATALONIA – Marlyn D. Cuadrado, Ignasi Mañé Bosch, Heredia F. Javier

An energy community is a legal figure, recently coined by the European Union, that creates a framework to encourage active participation of citizens and local entities in the energy transition to net-zero. In this work, we study the optimal participation of energy communities in day-ahead, reserve, and intraday electricity markets.

The motivation to do so is that there are time periods

where energy communities cannot meet their internal demand, and periods where they generate excess electricity. This is because the electricity they generate often comes from variable renewable resources like solar and wind. Electricity market participation is a natural way to ensure they meet their internal demand at all times, and, simultaneously, make the most of the excess electricity.

We propose a multi-stage stochastic programming model that captures variable renewable and electricity price uncertainty. The multi-stage aspect models the different times at which variable renewable generation is considered to be known and electricity prices from different markets are revealed. This results in a very large scenario tree with 34 stages, and hence a very large optimization problem. Scenario reduction techniques are applied to make the problem tractable. Case studies with real data are discussed, considering different energy community configurations, to analyse proposed regulatory frameworks in Europe. The added value of considering stochasticity in this problem is also analysed.

Hierarchical optimization of cooperative distributed energy resource aggregations in power distribution systems – LORELEY SEPPO, STUDENT – Miguel Anjos, Amira Dems, Hanane Dagdougui, Antoine Lesage Landry

The rise in energy transition incentives has increased the prevalence of distributed energy resources (DERs), posing both challenges and opportunities for the electric grid. We propose a cooperative approach to optimize individual load consumption providing on-demand flexibility to support the distribution system operator (DSO). We introduce a hierarchical optimization strategy using two distinct timescales. First, a day-ahead phase coordinates loads and sets hourly individual setpoints. The objective is to adjust the aggregated consumption to follow a regulation consumption level provided by the DSO. The regulation signal represents the optimal aggregated consumption, enabling the DSO to ensure efficient network operations. At the day-ahead level, we consider uncertainties inherent to the problems through chance constraints. This allows us to account for intrinsic unpredictability of DERs, customer baseloads, and DSO targets. Additionally, we integrate power flow constraints into the day-ahead level to prevent exceeding physical network limits. Individual setpoints evaluated for each user are then utilized in the second optimization step where real-time implementation is considered. Each load employs model predictive control to track their corresponding individual setpoint while subject to only local information and uncertainties. Lastly, we propose a cooperative game-theoretic mechanism based on marginal contributions to provide a scalable and fair allocation of the aggregation payoff over all participating users. The approach is numerically implemented to demonstrate its ability to provide flexibility to modern power systems.

Interdependent Network Optimization Modelling for Energy Network Design to Support Post Disaster Services – KRISTEN SCHELL, CARLETON UNIVERSITY – Yicheng Wang

Social infrastructure systems such as police, hospitals and banking services, rely heavily on civil infrastructure systems, like power and telecommunications, to provide their services. However, the interdependencies between social and civil infrastructure systems are rarely modeled, as they are in this

modeling effort. This work extends an interdependent network optimization model through the development of a capacity expansion model for electricity generation, to optimally provide energy to both civil and social infrastructure post-disaster. The model is based on a detailed representation of a community that faces repeated natural disasters, including hurricanes and flooding – the CLARC community. The results analyze the impacts of both infrastructure outage level and outage type on energy system design. Energy capacity location and technology type is highly dependent on the availability of the energy transmission network, which limits the effectiveness of large-scale technologies like nuclear power.

4d - ThA46 - 513C - **Optimization Methods for Renewable Power Systems** Chair: ANTOINE LESAGE LANDRY

Privacy-Preserving Decentralized Local Electricity Markets – LESIA MITRIDATI, TECHNICAL UNIVERSITY OF DENMARK –

Amidst the worldwide efforts to decarbonize power networks, Local Electricity Markets (LEMs) in distribution networks are gaining importance due to the increased adoption of renewable energy sources and growing active participation of consumers and prosumers. Considering that LEMs involve data exchange among independent entities, privacy and cybersecurity are some of the main practical challenges in LEM design. This presentation will introduce a novel secure market protocol using innovations from distributed optimization and Secure MultiParty Computation (SMPC). The considered LEM is formulated as an uncertainty-aware joint market for energy and reserves with affine balancing policies. To achieve scalability and enable the use of SMPC, market clearing is solved using the Consensus ADMM algorithm. Subsequently, the data exchange among participants via ADMM iterations is protected using the Shamir secret-sharing scheme to ensure privacy. The market protocol is further reinforced by a secure and verifiable settlement process that uses SMPC and ElGamal commitments to verify market quantities and by a secure recovery scheme for missing network measurements. Finally, the feasibility and performance of the proposed LEM are evaluated on a 15-bus test network.

Data Valuation in Data-driven Stochastic Optimization – JUAN M. MORALES, UNIVERSITY OF MALAGA – Robert Mieth, Adrian Esteban Perez

In this talk, we deal with data-driven stochastic optimization, which has become a powerful tool to address decision-making problems across various domains by leveraging data in an effective manner. Despite its consolidated position, a significant gap remains in terms of theory and methodologies to assess the sensitivity of decision performance to the fundamental input of stochastic optimization: the dataset.

Against this background, we propose a methodology based on Distributionally Robust Optimization (DRO) to perform this sensitivity analysis by formally defining the marginal value of the quality of a dataset. Central to our approach is the utilization of the Wasserstein metric, which serves as a natural measure to encode the quality of data and allows us to posit a Wasserstein DRO formulation for computing the marginal value of datasets sourced from multiple providers. These datasets may pertain to the same uncertain input pa-

parameter or to different ones.

Finally, we use paradigmatic problems from the field of power system operations to illustrate how our methodology can be exploited to guide decision-making and data valuation and acquisition.

Hierarchical Feedback-Based Optimization of Distributed Energy Resources in Distribution Grids – JOHN W. SIMPSON PORCO, UNIVERSITY OF TORONTO –

We describe a novel hierarchical optimization-based control architecture for coordination of Distributed Energy Resources (DERs) in Distribution Networks (DNs). At a high level, the proposed control system enables the DN in aggregate to quickly respond to power set-point requests from the transmission level while maintaining local DN constraints. More granularly, the proposal allows for multiple independently-managed areas within the DN to optimize their local resources while coordinating to support the TN, and while maintaining data privacy; the only required inter-area communication is between physically adjacent areas within the DN control hierarchy. Using tools from variational analysis and control theory, we provide rigorous closed-loop stability guarantees and tuning recommendations. The proposal is further validated via case simulation studies on multiple three-phase unbalanced grids, ranging from 123 to 8500 buses, demonstrating scalability to tens of thousands of DERs under control of dozens of independent stakeholders.

Online Interior-point Methods for Time-Varying Equality-constrained Optimization – ANTOINE LESAGE LANDRY, POLYTECHNIQUE MONTRÉAL, GERAD – Jean Luc Lupien, Iman Shames

An important challenge in the online convex optimization (OCO) setting is to incorporate generalized inequalities and time-varying constraints. The inclusion of constraints in OCO widens the applicability of such algorithms to dynamic but safety-critical settings such as the online optimal power flow (OPF) problem. In this work, we propose the first projection-free OCO algorithm admitting time-varying linear constraints and convex generalized inequalities: the online interior-point method for time-varying equality constraints (OIPM-TEC). We derive simultaneous sublinear dynamic regret and constraint violation bounds for OIPM-TEC under standard assumptions. For applications where a given tolerance around optima is accepted, we propose a new OCO performance metric – the epsilon-regret – and a more computationally efficient algorithm, the epsilon OIPM-TEC, that possesses sublinear bounds under this metric. Finally, we showcase the performance of these two algorithms on an online OPF problem and compare them to another OCO algorithm from the literature.

4d - ThA204 - 513D - Mobile and static battery systems operations and planning Chair: ANDY SUN 

Decision-Focused Learning for Energy Storage Control and Behavior Prediction – ALGHUMAY-JAN SAUD, COLUMBIA UNIVERSITY – Ming Yi, Bolun Xu

We introduce a novel, decision-focused, end-to-end methodology that directly generates storage charge and discharge pro-

files at sub-hourly resolutions, leveraging past pricing data and storage activity. This approach does not rely on future price predictions or intermediary price forecasting models. Our method incorporates the prior knowledge of the storage model and infers the hidden reward that incentivizes the energy storage decisions. This is achieved through a dual-layer framework, combining a predictive layer with an optimization layer. We have also developed a hybrid loss function for effective model training. The numerical experiments on synthetic and real-world energy storage data show that our approach achieves the best performance against existing benchmark methods, which shows the effectiveness of our method.

Arbitrage and Regulation with Electricity Storage – DIRK LAUINGER, MIT – Andy Sun

Electricity storage is used for intertemporal price arbitrage and for regulation, i.e., for balancing unforeseen fluctuations in electricity supply and demand. We present an optimization model that computes bids for both arbitrage and regulation, and ensures that storage operators can honor their market commitments at all times for all fluctuation signals in an uncertainty set that is inspired by applicable market regulations. We encode this requirement with an infinite number of functional constraints. We show that the robust constraints with functional uncertainties are equivalent to a finite number of deterministic constraints, which leads to an exact bilinear reformulation that can be solved with off-the-shelf software.

Optimal Facility Planning for Electric Vehicle Battery Recycling – MATTHEW BRUN, MIT OPERATIONS RESEARCH CENTER –

The electric vehicle (EV) battery supply chain will face challenges in sourcing scarce, expensive minerals required for manufacturing and in disposing of hazardous retired batteries. Integrating recycling technology into the supply chain has the potential to alleviate these issues; however, players in the battery market must design investment plans for recycling facilities. In this work, we propose a two-stage stochastic optimization model for computing minimum cost recycling capacity decisions, in which retired batteries are recycled and recovered materials are used to manufacture new batteries. The model is a separable concave minimization subject to linear constraints, a class for which we design a new finitely convergent global optimization algorithm based on piecewise linear approximation that solves up to 10x faster than comparable algorithms. We propose an equivalent reformulation of the model that reduces the total number of variables by introducing integrality constraints. The reformulation can also be solved by our global algorithm with drastically reduced solve times. We detail a cut grouping strategy for Benders' decomposition in the second stage which improves convergence relative to single-cut and multi-cut implementations. To produce a set of second-stage scenarios, we design an approach for generating time-series projections for new battery demand, retired battery supply, and material costs. Analysis of the optimal solutions shows that effective investment in recycling can reduce battery manufacturing costs by 22% and reduce environmental impacts by up to 7%.

Optimal Planning and Scheduling for a Large-Scale Electric Bus Fleet Transition Problem – ROBIN LEGAULT, MASSACHUSETTS INSTITUTE OF TECHNOLOGY – Andy Sun

In recent years, several metropolitan areas worldwide have announced ambitious electrification targets for their public bus systems. The problem of electrifying these large-scale public transit systems poses significant logistic challenges. It involves long-term investment decisions, including planning the acquisition of different types of battery electric buses (BEBs), the retirement of conventional buses, and the location of charging infrastructure. At the operational level, the activities of an evolving mixed fleet of conventional and electric buses must be precisely scheduled to respect the charging dynamics of BEBs and ensure a sufficient level of service throughout the transition horizon. In this work, we present a multi-period integer programming formulation of the bus fleet transition problem that includes the joint planning of yearly investment decisions and hourly operations. We show that the feasible solutions to the operational problem can be uniquely mapped to a set of cycles in an underlying graph. Taking advantage of this structure, we propose an equivalent cycle-based formulation of the problem and develop a range of policy restriction and column generation heuristics that rely on both formulations. Extensive experiments on large-scale instances based on real data from North American cities show that our approach can generate solutions with an average optimality gap of less than 2% in a few hours. In comparison, general-purpose MILP solvers generally cannot identify a feasible solution with the same computational budget.

4d - ThB33 - 513C - **Learning from Optimization in Energy** Chair: DAVI VALLADÃO 

On the Coupling Between Charging Scheduling of Electrified Bus Fleets and Grid Pricing – NATHAN CHO, CORNELL UNIVERSITY – Andrea Lodi, Anna Scaglione, Tong Wu

Electric buses are increasingly becoming an important topic in urban transportation systems. Many cities are beginning to incorporate them for their environmental and health benefits. Despite these advantages, electric bus systems face many technical challenges when compared to traditional gasoline powered buses. Various strategies have been proposed in the literature to leverage the flexibility of individually owned electric vehicles and fleets. A crucial component of the problem lies in the local electric grid load created by the demand from charging the vehicles. The collective demand from charging an entire fleet of vehicles results in a significant demand spike, insofar that charging the vehicles becomes a price-making activity, rather than a price-taking one. This shift poses further challenges in the vehicle scheduling in that we must optimize by reducing costs, but the choices also affect and determine the costs. In this study, we introduce methods to tackle this chicken-and-egg problem by iteratively solving the electric bus vehicle scheduling problem, and learning the costs of the routes that affect and are affected by the grid load.

Advancing Renewable Energy Simulation: Integrating Time Variant Hidden Markov Models with SDDP for Enhanced Accuracy and Real-World Application – ANDRE FELIPE DE CARVALHO RAMOS, PUC RIO –

Simulating energy generation from intermittent sources like solar and biomass is vital for efficient energy planning. Accurate

estimation and analysis of these sources are crucial for system design and decision-making. However, conventional approaches fail to capture the intricate dynamics of these time series, necessitating more advanced modeling techniques. In this study, we highlight the limitations of traditional methodologies and propose an exogenous variable-dependent Hidden Markov Model (HMM) that incorporates time-variant transition matrices. This novel approach enables more realistic simulations, as demonstrated in the context of biomass generation. Moreover, we present an integrated application of a Time Variant Hidden Markov Model in a SDDP framework in a real world scenario of a hydrothermal power system model, where the model involves an independent system operator (ISO) responsible for monthly planning the economic dispatch for power plants and water stored in reservoirs while meeting energy demand. The main contributions of this research include the formulation of the adapted HMM, best practices and implementation details and real-world application in renewable energy simulations.

Task-Based Prescriptive Trees for Two-Stage Linear Decision-Making Problems: Reformulations, Heuristic Strategies, and Applications – BRUNO FANZERES, PUC RIO – Rafaela Ribeiro

Most decision-making under uncertainty problems found in industry and studied by the scientific community can be framed as a two-stage stochastic program. In the past decades, the standard framework to address this class of mathematical programming problems follows a sequential two-step process, usually referred to as estimate-then-optimize, in which a predictive distribution of the uncertain parameters is firstly estimated, based on some machine/statistical learning (M/SL) method, and, then, a decision is prescribed by solving the two-stage stochastic program using the estimated distribution. In this context, most M/SL methods typically focus only on minimizing the prediction error of the uncertain parameters, not accounting for its impact on the downstream decision problem. However, practitioners argue that their main interest is to obtain near-optimal solutions from the available data with minimum decision error rather than a least-error prediction. Therefore, in this talk, we discuss the new framework referred to as task-based learning in which the M/SL training function also accounts for the downstream decision problem. As the M/SL method, we focus on decision trees, and study decision-making problems framed as a two-stage linear program. We present an exact Mixed-Integer Linear Program (MILP) formulation for the task-based learning method and construct two efficient recursive-partitioning Heuristic Strategies for the MILP. We conclude the talk by analyzing a set of numerical experiments illustrating the capability and effectiveness of the task-based prescriptive tree learning framework, benchmarking against the standard estimate-then-optimize framework, and discussing the computational capability of the constructed heuristic strategies vis-à-vis the MILP formulation.

Contextually Robust Optimization: End-to-End Learning in Power System Applications – DAVI VALLADÃO, LAMPS PUC RIO – Carlos Gamboa, Alexandre Street

We have developed a novel analytics framework that computes contextually robust data-driven decisions with recourse. In contrast to existing contextual prescriptive approaches, our framework completely disregards the predictive step. Considering contextual information reduces uncertainty, leading

to significant economic benefits, such as lower operating costs than those estimated by traditional robust optimization models. A simple parameterization allows robustness control, enabling optimal decisions ranging from contextual to unconditional robustness. We provide a single-level reformulation with the same complexity as the unconditional problem. Finally, we demonstrate the framework's performance in a economic dispatch problem with uncertain demand and renewable generation.

4d - ThB118 - 511E - **Optimization of Gas Networks** Chair: LARS SCHEWE 

Adjustable Robust Nonlinear Network Design under Demand Uncertainties – JULIA GRÜBEL, UNIVERSITY OF TECHNOLOGY NUREMBERG (UTN) – Johannes Thürauf, Martin Schmidt

We study network design problems for nonlinear and nonconvex flow models under demand uncertainties. To this end, we apply the concept of adjustable robust optimization to compute a network design that admits a feasible transport for all, possibly infinitely many, demand scenarios within a given uncertainty set. For solving the corresponding adjustable robust mixed-integer nonlinear optimization problem, we show that a given network design is robust feasible, i.e., it admits a feasible transport for all demand uncertainties, if and only if a finite number of worst-case demand scenarios can be routed through the network. We compute these worst-case scenarios by solving polynomially many nonlinear optimization problems. Embedding this result for robust feasibility in an adversarial approach leads to an exact algorithm that computes an optimal robust network design in a finite number of iterations. Since all of the results are valid for general potential-based flows, the approach can be applied to different utility networks such as gas, hydrogen, or water networks. We demonstrate the applicability of the method by computing robust gas networks that are protected from future demand fluctuations.

Mixture of Gases on Networks – PASCAL BÖRNER, TU DARMSTADT – Marc Pfetsch, Stefan Ulbrich

With the move towards a climate neutral usage, gas suppliers and transport companies have begun to mix a certain percentage of hydrogen into the gas networks. This poses new challenges for the modeling and optimization of gas transport. Therefore, we have developed a model for the mixture of gases on networks. The model is based on an equation of state for the mixture, the stationary isothermal Euler equations and coupling conditions for the flow and the mixture. The equation of state or pressure law which we developed is based on the change of the speed of sound in a mixture of gases. We prove that the gas flow is unique, even in the case of a mixture. This is not trivial since the mixture is changing the flow properties and it is not clear anymore if there exist different network flows with different mixing ratios. Further, we use the model to solve stationary gas flow problems to global optimality on large networks. Therefore, the model is implemented and solved with the help of the MINLP-solver SCIP. We examined different implementations of the model and their impact on computational performance.

A tailored interior point method for fast op-

timization on gas networks – ROWAN TURNER, UNIVERSITY OF EDINBURGH AND HERIOT WATT UNIVERSITY –

The usage of hydrogen gas as an energy carrier is one likely factor towards achieving a net zero future. There are many possible applications, but one important question is whether existing natural gas infrastructure can be re-used for hydrogen and the challenges this brings for the control of these future networks. In particular, we expect that a hydrogen network which uses hydrogen generated from excess renewable electricity would be more difficult to control as the patterns of injection and withdrawal would be much less regular than today. Additional challenges arise from new operating parameters required for hydrogen and not for natural gas – such as controlling for pressure fluctuations to prevent pipe-ageing. Motivated by a need for instationary optimization methods on networks at scale, we present a specialized interior point method for gas problems. Our test problem is a line-pack optimization problem using a discretization of the 1d isothermal Euler equations, as a step towards understanding the important questions above. By incorporating a bespoke preconditioned iterative solver to tackle the linearized systems at each iteration of the interior point method, which form the key computational bottleneck in such a method, we utilize the highly structured nature of the problem. This arises from both the network structure itself and the time discretization. The expectation is that the method will scale well with both network size and time windows, and be generalizable to broader instationary gas network optimization problems.

Optimization methods for the analysis of gas markets – LARS SCHEWE, UNIVERSITY OF EDINBURGH –

We discuss how optimization methods can be used to analyze the European Gas Market. Specifically, we discuss how we can analyze the effect of market power on booking decisions and how uncertainty makes it difficult to find market equilibria. We will close with an outlook of how a switch to hydrogen will make these questions more difficult and/or interesting to analyze.

4d - ThC261 - 513C - **Research at Hydro-Québec: Black-box Optimization** Chair: MIGUEL DIAGO 

Black box optimization with virtual sensors for turbine deformations during startup events – VINCENT MAI, HYDRO QUÉBEC – Arthur Favrel, Martin Gagnon, Quang Hung Pham

During startup, hydroelectric turbine blades undergo important deformations, which significantly reduce the device's Remaining Useful Life (RUL) through fatigue. Directly measuring these deformations can only be done for a dozen of trials during the commissioning of the turbines. Our objective during these trials is to find the optimal startup controller parameters which minimize the deformation peak-to-peak range while reaching the synchronous rotation speed in given time. After gathering some initial startups, we train a neural model to predict the deformation envelope from the turbine's startup trajectory in the 2-D space defined by the rotational speed and the wicket gates opening. This model, coupled with a turbine dynamics surrogate simulator, defines a black box. Given a set

of startup parameters, it outputs a cost based on the deformation amplitude, the model's uncertainty, and the time to reach the synchronous speed. The NOMAD optimizer is then used to find the optimal parameters. First, in an active learning phase, the cost considers the model's uncertainty so that the optimized parameters, when tested on the turbine, provide relevant data to improve the deformation model. Then, we only consider the deformation amplitude and time to synchronous speed to output the optimal startup parameters.

Risk-Based Optimization of Periodic Maintenance Policies for Electrical Transmission Equipment – MIGUEL DIAGO, HYDRO QUEBEC – Xavier Lebeuf, Jean Marc Meango, Alain Côté

Maintenance of electrical transmission equipment is key to ensure the reliability of power supply. Maintenance tasks may be triggered, for example, by equipment failures (reactive maintenance), observed anomalies expected to lead to failures (condition-based maintenance), and time (periodic preventive maintenance). Periodic maintenance tasks are typically scheduled at fixed time intervals to detect ongoing degradation mechanisms and apply corrective measures as needed. Usually, engineers choose the period between tasks using their knowledge of failure mechanisms. In the context of electric transmission networks, the period between tasks should be a compromise driven by factors that include equipment reliability, maintenance costs (inspections, repairs, replacements...), the value of lost load (VoLL) and other risks inherent to power transmission (environment, health and safety...). Herein, an asset behaviour model, an event stochastic simulator, a power-flow simulator, and a risk model with a VoLL estimator are combined to quantify the total cost of periodic maintenance strategies. Then, a blackbox optimization solver searches for periodic maintenance strategies that minimize costs within specified constraints. Given that the event simulator uses a Monte Carlo method to output network states where equipment fails following set statistical distributions, the blackbox is non-deterministic. However, timely and meaningful results are obtained through proper adjustment of the number of Monte Carlo cycles, the length of the timespan simulated, and other parameters. This opens the way to multifidelity optimization, where these parameters are automatically adjusted during optimization. Ultimately, engineers may use this approach to select optimal periodic maintenance schedules that minimize the global risk for the network operator. The procedure is implemented with NOMAD, an open-source blackbox optimizer.

Comparison of optimization methods for hyperparameter tuning of a deep learning forecasting model – ALEXIA MARCHAND, HYDRO QUÉBEC – Arnaud Zinflou

Hyperparameter optimization remains one of the main challenges in machine learning and often relies on manual tuning of hyperparameters, or the use of naive approaches such as grid search and random search. Unfortunately, manual hyperparameter tuning is a lengthy and inefficient process that heavily depends on the user's expertise. Grid search and random search, on the other hand, are techniques that are not very efficient and require a significant amount of computation.

With an electric short-term demand forecasting model based on a Temporal Fusion Transformer (TFT) architecture,

we studied the impact of using different optimization methods, including Bayesian Optimization and Hyperband, Latin hypercube, and Mesh Adaptive Direct Search (MADS), on the performance of the model. Results show that the MADS optimization method was the most efficient and that the choice of the objective function and constraints has a significant impact on the resulting optimized forecasting model.

4d - FB959 - 513C - Energy and Environment

4 Chair: HANDE BENSON



Efficient hydrogen infrastructure design using JuMP and convex approximations – TRULS FLATBERG, SINTEF – Frida Sæther

The transition to a sustainable energy future requires innovative solutions that reduce greenhouse gas emissions across various sectors. In this context, hydrogen has emerged as a promising alternative fuel, particularly for applications where electrification is challenging. One such sector is fisheries, which heavily relies on diesel-powered vessels and equipment.

Green hydrogen is produced by electrolysis and transported to end users by truck or ship. Currently, hydrogen is transported in compressed form with low energy density. Therefore, it is beneficial to reduce transport costs by implementing distributed production. However, this cost-saving approach needs to be balanced against the more efficient production and lower costs of larger electrolyzers. The electrolyzers are assumed grid-connected and may be exposed to time-varying electricity and power tariffs. Additionally, they produce surplus heat that can be sold for commercial value.

To integrate all these aspects into one model, we use the JuMP modelling framework in Julia. Our modular approach allows us to set up and solve a mixed-integer linear optimization problem for determining the location and dimensions of the hydrogen infrastructure. To address non-linearities arising from varying electrolyser efficiency and dynamics in the electricity market, we employ a package that identifies optimal piecewise linear convex approximations based on a set of data points. This approximation of production costs is then combined with separate modules for storage and transport modeling, resulting in an overall facility location model.

Initial results from a case study in the ZeroKyst project based on a region of Northern Norway will be used to illustrate the approach.

Applications of Optimization in Land Conservation – HANDE BENSON, DREXEL UNIVERSITY –

Given an allotment of land divided into parcels, government decision-makers, private developers, and conservation biologists can collaborate to select which parcels to protect, in order to accomplish sustainable ecological goals with various constraints. In this talk, we propose several mixed-integer optimization models that considers the presence of multiple species on these parcels, subject to predator-prey relationships and crowding effects. Recognizing that the relationships within and across species, their resources, and environment are quite complex, we also incorporate simulation-based estimates of long-term population viability into our models.

4e - MB219 - 512E - **Applications of Stochastic optimization in Healthcare** Chair: ANTOINE LEGRAIN

A Network-Flow Approach to Increasing Inter-Facility Transfers for Nova Scotia Health Long-Term Care – AMELIA LANE, DALHOUSIE UNIVERSITY – Borna Baradaran Noveiri, Lindsay Cormier, Joachim Mwebembezi, Neel Patel, Peter Vanberkel

Due to access and flow challenges within acute care, nearly 3,000 long-term care residents in Nova Scotia, Canada are on a waitlist to be transferred to their preferred facilities. Reasons for transferring include moving closer to families, cultural communities, and to reunite with a life partner in a different facility. We developed a new method to identify an optimal series of transfers which considers residents' care needs (i.e. required bed type, smoking, life partner), priority level, facility preferences, and the features of a vacant bed. We model this by generating a network of every resident on the waitlist and solving it as an elementary path problem, generating a chain of transfers starting with a vacancy. We use dynamic programming to solve this problem in practice. In this talk, the mathematical models and algorithms used to generate these chains, as well as an overview of their potential impact, will be discussed.

Quantifying the benefits of customized vaccination strategies: A network-based optimization approach – SU LI, TEXAS A&M UNIVERSITY

We study the problem of designing vaccine distribution strategies that maximally mitigate the negative impact of an infectious disease outbreak. This is achieved through a multiperiod optimization-based framework that embeds important subject-specific risk and contact information into the decision-making process. By analyzing the structure of the resulting optimization problem, we identify key structural properties which we use to construct a globally convergent solution scheme (suitable for smaller problem instances) and two, more scalable, heuristic schemes. We demonstrate the benefits of the considered framework through a case study on COVID-19 in Texas. Our results highlight the importance of considering risk and contact information as doing so substantially reduces the total expected number of fatalities over conventional compartmental-based approaches. These findings indicate that customization can have a significant benefit, particularly for community-scale planning.

Multi-objective scheduling of chemotherapy drug preparation – CAMILLE PINÇON, POLYTECHNIQUE MONTRÉAL – Antoine Legrain, Nadia Lahrichi

Most cancer patients receive chemotherapy. This type of treatment relies on dangerous and expensive drugs. The quality of drug preparation, which must be delivered in the right quantities and on time, is therefore a key point in a patient's treatment pathway. To address this, a multi-objective scheduling method has been developed for drug manufacturing in a chemotherapy clinic, to reduce patient waiting times and minimize preparation expenses. A lexicographic approach coupled with a linear programming model that approximates scheduling is used. The objective is to jointly minimize the maximum

delay of a task, the associated losses of raw materials, and overall lead times during production days. The model also enables the selection of medications to be produced in advance, aiming to increase the production buffer for the following day while also minimizing raw material losses. The resulting solution is then reconstructed to clarify the scheduling details and establish a schedule that can be used by pharmacists within clinical constraints. This reconstruction makes it possible to reduce delays, select drugs for early preparation, and improve productivity while reducing costs. This scheduling method contributes to a more efficient and cost-effective workflow in chemotherapy clinics, resulting in improved quality of patient care and resource utilization.

Real-time demand-driven inventory management in a hospital pharmacy – ALI JAFARI, POLYTECHNIQUE MONTREAL – Antoine Legrain, Nadia Lahrichi

Hospital pharmacies receive thousands of prescriptions per day, consequently, it is crucial that the hospital medication circuit be as optimal as possible. Nevertheless, hospital inventory management is highly complex for various reasons, including space limitations, demand uncertainty, and human resource limitations. Furthermore, the inventory policy must be adaptable to real-time fluctuations in patient demands, with a simultaneous emphasis on preventing backorders and stock-outs, given their potential to harm patients' well-being. Additionally, it is not operationally effective for care units (CUs) to repeatedly request new supplies every day to avert situations of insufficient stock. To address this problem, we present a real-time inventory control model based on the characteristics of medication demand while considering uncertainty in demand and space limitations, with the main objective of minimizing CU replenishment frequency. First, we categorize the medications into two groups; fast-moving and slow-moving medications, then, we utilize a continuous and dynamic review inventory control policy for these categories respectively. For optimizing the inventory control parameters we propose a stochastic optimization model for each category. To efficiently solve the proposed inventory policy in real-time, we employ a receding-horizon control (RHC) strategy, where the models are solved iteratively over a predetermined time horizon. To validate the effectiveness of our proposed approach, we use a real-world inventory management setting for a hospital in Montréal and conduct a comparative analysis between the proposed model and the existing state of the inventory policy to demonstrate the advantages of this research.

4e - MC283 - 514A - **Oncology and Clinical Decision Support** Chair: ADRIAN MANUEL REIMANN

Data-Driven Alternatives for Feedback Mechanisms in Acute Myeloid Leukaemia – CARL JULIUS MARTENSEN, OTTO VON GUERICKE UNIVERSITY MAGDEBURG – Sebastian Sager, Vijay Ivaturi, Niklas Korsbo

Modeling plays a crucial role in pharmaceutical and medical applications by facilitating the comprehension and analysis of intricate biochemical processes and the design of optimal treatments. Particularly in pharmacokinetic and pharmacodynamic modeling, which investigates the absorption, dis-

tribution, metabolism, and effects of drugs, developing interpretable, precise, and robust models is paramount. However, modeling is a resource-intensive process that demands significant expertise, time, data, and computational resources. Researchers often resort to pre-existing models with fixed coefficients that may not be tailored to their specific dataset due to its size or characteristics. This might lead to the creation of overly complex models, excessive parameterization, incorrect calibration, or even erroneous insights into underlying mechanisms. The advancement of automatic differentiation and machine learning has provided a data-driven alternatives to traditional modeling. Scientific machine learning, in particular, merges first order principles with adaptable algorithmic frameworks, offering semi-automated approaches to model discovery. In pharmacokinetic-pharmacodynamic (PKPD) systems, deep nonlinear mixed effect models enable the integration of universal approximators within fixed model architectures, allowing for the identification and extraction of previously unknown signals. This study delves into the utilization of deep nonlinear mixed effect models in conjunction with symbolic regression to explore data-driven alternatives to a conventional baseline model. Specifically, the objective is to substitute the feedback term in the classical Friberg model using a previously published dataset related to acute myeloid leukemia treatment with intermediate to high-dose cytarabine.

Inference for parameters of selection and copy number aberrations from DNA-sequencing data – KHANH DINH, COLUMBIA UNIVERSITY –

Over the last decade, bulk DNA-sequencing (DNA-seq) has allowed us to appreciate the sheer amount and diversity of the genomic changes associated with cancer development. More recently, advances in single-cell DNA-seq has enabled profiling of copy number aberrations (CNAs) at high resolution in thousands of cells, and has uncovered the level of intra-tumor chromosomal instability (CIN). These technologies promise quantitative measurements of tumor dynamics, and measurements of the rate of chromosomal aneuploidy, whole-genome duplications and replication errors in tumors. We have developed CINner, a state-of-the-art mathematical model and simulation algorithm for studying single-cell dynamics in a population of cells, incorporating clonal selection of somatic driver mutations and copy number aberrations (CNAs), as well as accumulation of neutral passenger mutations and CNAs. CINner follows population dynamics as input by the user, generates the clonal evolution forward in time, where clones are defined by their copy number and driver mutation profiles. The phylogeny of a sample is then computed backward in time. CINner is designed to be efficient for large cell populations while maintaining statistical accuracy in the sampled cells. The mathematical model and simulation method provide an efficient framework for modeling genomic changes during tumor growth, and are easily adaptable to accommodate newly uncovered genomic alteration mechanisms.

We present several applications of the simulator package. The first study uncovers the selection forces driving chromosomal CNAs, estimated from large bulk DNA-seq datasets. The inferred selection parameters can predict the prevalence of whole-genome duplication in each cancer type. Moreover, the selection parameters inferred from a pan-cancer dataset show a strong correlation with the chromosomal driver gene count and potency. Together, these observations prove the bi-

ological relevance of the uncovered selection parameters. The second part of this presentation delves into the development of an inference method for single-cell DNA-seq data, which holds great potentials for more accurate parameter estimation but also poses several distinct challenges over bulk data. Based on a novel implementation of random forests within the Approximate Bayesian Computation (ABC) framework, we developed an inference method to uncover both occurrence rates and selection parameters driving specific CNA mechanisms in CINner. The inference recovers the parameters of interest well, even for datasets with small sample sizes. In combination, the width and depth of these studies showcase CINner's applicability in analyzing current and upcoming DNA-seq data, toward the goal of reconstructing tumor history and predicting patient outcome.

Optimal Leukemia Treatment – ADRIAN MANUEL REIMANN, OTTO VON GUERICKE UNIVERSITY MAGDEBURG – Sebastian Sager

With an overall 5-year relative survival rate of 30% Acute Myeloid Leukemia (AML) is a serious disease with a complex dynamic acting on the whole organism. Due to this complexity the treatment choices made by clinicians are crucial for the prospects of recovery of every single patient. Finding and researching new and promising improvements in treatment strategies for this deadly disease using clinical trials is therefore a very slow and tedious process. We aim to accelerate this process by using modelling, simulation, and optimization methods to find new treatment regimens that both improve the chance of survival and reduce the stress on the patients during these treatments. In this talk we will discuss a pharmacokinetic and pharmacodynamic model of ordinary differential equations that allows to simulate both chemotherapy and immunostimulator intake at arbitrary points in time. Considering white blood cell measurements obtained during clinical trials we estimate sets of patient specific parameters for this model which serve us as digital twins allowing for detailed analysis of simulated drug response. With these digital twins we are able to reproduce clinically observed treatment responses strengthening the trust in our model. This validation allows us to expand the analysis upon yet unconsidered treatment combinations and gives insight into clinically hard to observe behaviours.

4f - MB960 - 512G - Finance and social sciences **1 Chair: JOÃO VITOR PAMPLONA**

Optimal hedging of the interest rate swap book – JÖRGEN BLOMVALL, LINKÖPING UNIVERSITY –

With an optimization model interest rate curves are measured with increased accuracy from Overnight Index Swaps. Principal Component Analysis identifies the significant risk factors in interest rate markets. With these a Stochastic Programming model is formulated to determine the optimal hedge of the Overnight Index Swap book, where significant improvements are found relative to traditional delta hedging.

Predicting Limit Order Book Prices: A Machine Learning Approach – FRANCIS HUOT CHANTAL, UNIVERSITÉ DE MONTRÉAL – Fabian Bastin

In high-frequency trading, the dynamics of a double auction market emerge, with "buy" orders and "sell" orders queued

and managed using a limit order book (LOB). Within this framework, accurate prediction of price movements becomes crucial for optimizing the performance of high-frequency algorithms. This talk explores machine learning models for forecasting the LOB mid-price, comparing Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM). First we will explore various methods for modeling and predicting LOB mid-prices. This forecasting involves diverse methodologies, including stochastic differential equations, which capture market dynamics such as mean-reversion, momentum, and volatility asymmetry. Second, we will discuss the practical implementation across various scenarios. The models will yield both regression and classification outputs, incorporating time-defined and event-defined intervals, the latter offering advantages in mitigating intraday seasonal effects. We will also explore the impact of excluding the first 30 minutes and the last 30 minutes of data on performance, as price volatility and market activity during these intervals tend to be greater than the rest of the day. Finally, we will evaluate performance using machine learning metrics such as the F1-score and R squared.

Numerical experiments will utilize a full year of high-frequency data from the S&P/TSX 60, comprising 60 major companies spanning ten distinct sectors traded on the Toronto Stock Exchange. Machine learning and deep learning methods present challenges in finance due to data volatility and underlying complexity. CNNs excel in capturing spatial patterns, originally designed for image data analysis, while LSTMs effectively handle complex time series, even with noisy input sequences.

Contextual Portfolio Optimization: Some Recent Results – ROY KWON, UNIVERSITY OF TORONTO – Andrew Butler

Smart Predict and Optimize’ is an emerging paradigm where the parameters and optimal solutions for an optimization problem are jointly determined. We present some recent results using this paradigm for mean-variance optimization. We show that the integrated approach can result in parameter estimations and optimal solutions that perform better out of sample than those obtained from the ‘predict or estimate first, then optimize’ approach. Computational challenges will be discussed which motivates the development of decomposition methods based on the alternating direction method of multipliers (ADMM) for the optimization layer in the neural network that represents the integrated (joint) approach.

Fair Mixed Effects Support Vector Machine – JOÃO VITOR PAMPLONA, TRIER UNIVERSITY – Jan Pablo Burgard

When using machine learning for automated prediction, it is important to account for fairness in the prediction. Fairness in machine learning aims to ensure that biases in the data and model inaccuracies do not lead to discriminatory decisions. E.g., predictions from fair machine learning models should not discriminate against sensitive variables such as sexual orientation and ethnicity. A fundamental assumption in machine learning is the independence of observations. However, this assumption often does not hold true for data describing social phenomena, where data points are often clustered based. Hence, if the machine learning models do not account for the cluster correlations, the results may be biased. Especially high is the bias in cases where the cluster assignment is correlated

to the variable of interest. We present a fair mixed effects support vector machine algorithm that can handle both problems simultaneously. With a reproducible simulation study we demonstrate the impact of clustered data on the quality of fair machine learning predictions.

4g - MB162 - 512B - Quantum Computing and Optimization I Chair: BRANDON AUGUSTINO

Symplectic Extra-gradient Type Method for Solving General Non-monotone Inclusion Problem – YI ZHANG, INSTITUTE OF COMPUTATIONAL MATHEMATICS AND SCIENTIFIC/ENGINEERING COMPUTING –

The min-max optimization has attracted substantial interest given its critical role in the training of Generative Adversarial Networks and solving Nash equilibrium. Because the first-order condition of min-max problem can be written as an inclusion problem, the extragradient method, a efficient method to solve inclusion problem, has gained wide attention. Recently, accelerated versions of the extragradient method inspired by Halpern’s iteration have been under intensive study. Notably, however, Halpern’s iteration can suffer from computational inefficiencies since its iterative rule is a convex combination of the starting and current feasible points. In this talk, we present an overview of the fundamental theory behind a brand new acceleration technique, referred to as symplectic acceleration. Leveraging this symplectic acceleration, we derive the symplectic extra-gradient method and its subsequent variations. We employ a Lyapunov function to demonstrate that the symplectic extra-gradient method and its variants exhibit an inverse of quadratic convergence rate in terms of the squared gradient norm. Furthermore, we refine the proof techniques for proving sharper convergence results of symplectic acceleration to establish a sharper convergence rate for the symplectic extra-gradient type method. For improved practical performance, we apply a line search technique into the symplectic extra-gradient framework. From a theoretical standpoint, we guarantee the convergence of the symplectic extra-gradient type method with line search. Empirically, numerical experiments show computational efficiency of the symplectic extra-gradient method when combined with line search. This adaptation exhibits faster convergence rates in practice compared to existing extra-gradient type methods.

Quantum Algorithms for Multi-objective Optimization – STEFAN WOERNER, IBM QUANTUM –

Quantum computing is a new computational paradigm that has the potential to disrupt certain disciplines, with (combinatorial) optimization frequently being mentioned as one of them. However, classical algorithms and heuristics can often achieve very good results quickly, leaving little room for further improvements.

Multi-objective optimization represents a class of optimization problems where the goal is to find the optimal trade-offs between multiple objective functions. Thus, it is usually not the goal to find a single solution, but (a good approximation of) the Pareto frontier, i.e., the collection of all Pareto optimal solution.

In this context, the sampling-based nature of many quantum optimization algorithms can be beneficial as it can quickly

produce a variety of good solutions to approximate the Pareto frontier.

We demonstrate how quantum optimization can be efficiently applied to certain multi-objective combinatorial optimization problems and present first promising results that show that multi-objective optimization could be a problem class particularly suitable for quantum computers.

Solving the semidefinite relaxation of QUBOs in matrix multiplication time, and faster with a quantum computer – GIACOMO NANNICINI, UNIVERSITY OF SOUTHERN CALIFORNIA – Brandon Augustino, Tamás Terlaky, Luis Zuluaga

Recent works on quantum algorithms for solving semidefinite optimization (SDO) problems have leveraged a quantum-mechanical interpretation of positive semidefinite matrices to develop methods that obtain quantum speedups with respect to the dimension n and number of constraints m . While their dependence on other parameters suggests no overall speedup over classical methodologies, some quantum SDO solvers provide speedups in the low-precision regime. We exploit this fact to our advantage, and present an iterative refinement scheme for the Hamiltonian Updates algorithm of Brandao et al. (Quantum 6, 625 (2022)) to exponentially improve the dependence of their algorithm on the precision ϵ , defined as the absolute gap between primal and dual solution. As a result, we obtain a classical algorithm to solve the semidefinite relaxation of Quadratic Unconstrained Binary Optimization problems (QUBOs) in matrix multiplication time. Provided access to a quantum read/classical write random access memory (QRAM), a quantum implementation of our algorithm exhibits $O(ns + n^{1.5} \text{polylog}(n, \|C\|_F, 1/\epsilon))$ running time, where C is the cost matrix, $\|C\|_F$ is its Frobenius norm, and s is its sparsity parameter (maximum number of nonzero elements per row).

Quantum Langevin Dynamics for Optimization – YUCHEN LU, TSINGHUA UNIVERSITY – Zherui Chen, Hao Wang, Yizhou Liu, Tongyang Li

We initiate the study of utilizing Quantum Langevin Dynamics (QLD) to solve optimization problems, particularly those non-convex objective functions that present substantial obstacles for traditional gradient descent algorithms. Specifically, we examine the dynamics of a system coupled with an infinite heat bath. This interaction induces both random quantum noise and a deterministic damping effect to the system, which nudge the system towards a steady state that hovers near the global minimum of objective functions. We theoretically prove the convergence of QLD in convex landscapes, demonstrating that the average energy of the system can approach zero in the low temperature limit with an exponential decay rate correlated with the evolution time. Numerically, we first show the energy dissipation capability of QLD by retracing its origins to spontaneous emission. Furthermore, we conduct detailed discussion of the impact of each parameter. Finally, based on the observations when comparing QLD with classical Fokker-Planck-Smoluchowski equation, we propose a time-dependent QLD by making temperature and $\hbar$ time-dependent parameters, which can be theoretically proven to converge better than the time-independent case and also outperforms a series of state-of-the-art quantum and classical optimization algorithms in many non-convex landscapes.

4g - MB209 - 512A - **A glimpse of computational optimization at Google I Chair: HOSSEIN BATENI** ←

Practical Performance Guarantees for Pipelined DNN Inference – MATTHEW FAHRBACH, GOOGLE RESEARCH – Aaron Archer, Kuikui Liu, Prakash Prabhu

We optimize pipeline parallelism for deep neural network (DNN) inference by partitioning model graphs into stages and minimizing the running time of the bottleneck stage, including communication. We give practical and effective algorithms for this NP-hard problem, but our emphasis is on tackling the practitioner's dilemma of deciding when a solution is good enough. To this end, we design novel mixed-integer programming (MIP) relaxations for proving lower bounds. Applying these methods to a diverse testbed of 369 production models, for 2, 4, 8, 16, 32, 64, we empirically show that these lower bounds are strong enough to be useful in practice. Our lower bounds are substantially stronger than standard combinatorial bounds. For example, evaluated via geometric means across our production testbed with pipeline stages, our MIP formulations raised the lower bound from 0.4598 to 0.9452, expressed as a fraction of the best partition found. In other words, our improved lower bounds closed the optimality gap by a factor of 9.855x.

Contextual Nearest Neighbor Search in Sub-linear Time – MOHAMMAD HAJIAGHAYI, UNIVERSITY OF MARYLAND –

Over the past couple of decades, locality-sensitive hash functions (LSH) appeared as the main technique for finding near neighbors with a vast list of applications in machine learning, data mining, and database management, among others. In this work, we introduce the concept of "contextual near-neighbor queries", where each query consists of a point accompanied by a context that specifies the dimensions that matter and those that should be ignored. We introduce an extension of the near-neighbor problem to accommodate such queries named the in-context near-neighbor problem.

Multi-Vector Retrieval: The New Frontier of Similarity Search – RAJESH JAYARAM, GOOGLE RESEARCH NYC – Vahab Mirrokni, Majid Hadian, Jason Lee, Laxman Dhulipala

Over the past decade, dense embedding models have become a fundamental component of modern information retrieval (IR) paradigms. These models produce a single d -dimensional embedding per data-point, allowing for fast retrieval via highly optimized maximum inner product search (MIPS) algorithms. Recently, beginning with the landmark ColBERT model, it has been shown that multi-vector models, which produces a set of embedding for a single data point, can achieve significantly higher quality results for IR tasks than single embedding models. Unfortunately, retrieval (nearest neighbor search) for these models is computationally expensive, as multi-vector similarity is a significantly more complex metric requiring new techniques. This dilemma is the primary bottleneck preventing the widespread use of multi-vector models, posing multi-vector retrieval as a central modern challenge in similarity search.

In this talk, we discuss the first scalable and direct search algorithms for multi-vector similarities. Our approach is based

on a randomized, data-dependent embedding from a set of vectors to a single vector, called "Fixed Dimensional Encodings" (FDE), which employ random space partitions and projections to approximate multi-vector similarity by single-vector dot products. This enables the usage of highly optimized MIPS solvers, backed by decades of prolific research, for multi-vector retrieval. The method is based on the theory of probabilistic tree embeddings underlying decades of theoretical nearest neighbor search literature. Empirically, it achieves latency on the same order of magnitude as single-vector retrieval, while nearly optimally capturing the more powerful expressivity of multi-vector similarity.

Attention-based Model Structure Optimization – HOSSEIN BATENI, GOOGLE RESEARCH – Kyr-
iakos Axiotis, Lin Chen, Matthew Fahrbach, Thomas Fu,
Vahab Mirrokni, Taisuke Yasuda

We propose a feature selection algorithm called Sequential Attention that achieves state-of-the-art empirical results for neural networks. This algorithm is based on an efficient one-pass implementation of greedy forward selection and uses attention weights at each step as a proxy for feature importance. We give theoretical insights into our algorithm for linear regression by showing that an adaptation to this setting is equivalent to the classical Orthogonal Matching Pursuit (OMP) algorithm, and thus inherits all of its provable guarantees. Our theoretical and empirical analyses offer new explanations towards the effectiveness of attention and its connections to overparameterization, which may be of independent interest.

4g - MC163 - 512E - **Quantum Computing and Optimization II** Chair: BRANDON AUGUSTINO [↔](#)

The quantum central path algorithm – BRANDON AUGUSTINO, MASSACHUSETTS INSTITUTE OF TECHNOLOGY –

We propose a novel quantum algorithm for linear optimization problems by quantum mechanical simulation of the central path. By combining our approach with iterative refinement techniques, we demonstrate an end-to-end asymptotic speedup over the best performing classical algorithms for a certain sparsity regime. We then discuss how to generalize our techniques to more complex classes of convex optimization problems, including linearly constrained convex quadratic optimization, and semidefinite optimization.

Quantum speedups for stochastic optimization – CHENYI ZHANG, STANFORD UNIVERSITY – Aaron Sidford

We consider the problem of minimizing a continuous function given quantum access to a stochastic gradient oracle. We provide two new methods for the special case of minimizing a Lipschitz convex function. Each method obtains a dimension versus accuracy trade-off which is provably unachievable classically and we prove that one method is asymptotically optimal in low-dimensional settings. Additionally, we provide quantum algorithms for computing a critical point of a smooth non-convex function at rates not known to be achievable classically. To obtain these results we build upon the quantum multivariate mean estimation result of Cornelissen et al. 2022 and provide a general quantum-variance reduction technique of independent interest.

Online learning of a panoply of quantum objects – ALICE ZHENG, VIRGINIA TECH – Akshay Bansal, Ian George, Soumik Ghosh, Jamie Sikora

In many quantum tasks, there is an unknown quantum object that one wishes to learn. In the online learning setting, one is tasked to adaptively refine a hypothesis to reproduce the measurement statistics of such an object. A common evaluation metric for such a strategy is its regret, or roughly the accumulated errors in the hypothesis statistics. We prove a sublinear regret bound for learning over general subsets of positive semidefinite matrices via the regularized-follow-the-leader algorithm and apply it to numerous settings. We establish a variety of matrix analysis results useful in quantum information theory, including a generalization of Pinsker's inequality for arbitrary positive semidefinite operators with possibly different trace.

4g - TC282 - 512E - **Quantum Computing and Optimization III** Chair: RODOLFO ALEXANDER
QUINTERO OSPINA [↔](#)

Binary control pulse optimization for quantum systems – SVEN LEYFFER, ARGONNE NATIONAL LABORATORY –

Quantum control aims to manipulate quantum systems toward specific quantum states or desired operations. Designing highly accurate and effective control steps is vitally important to various quantum applications, including energy minimization and circuit compilation. We consider discrete binary quantum control problems and propose optimization techniques with provable error bounds. We apply an alternating direction method of multipliers (ADMM) algorithm to solve the continuous relaxation of the penalized model, and round the resulting solution. We show how to amortize the computationally expensive factorization of the Hamiltonians by exploiting the discrete nature of the controls. Our algorithms can obtain high-quality control results, as demonstrated by numerical studies on diverse quantum control examples.

Hybrid Quantum Branch-and-Bound Method for Quadratic Unconstrained Binary Optimization – DAVID BERNAL NEIRA, PURDUE UNIVERSITY – Zedong Peng, Pedro Maciel Xavier

With the advancement of quantum hardware and theory, quantum algorithms have demonstrated their effectiveness in finding high-quality solutions for Quadratic Unconstrained Binary Optimization (QUBO) problems. Moreover, there exists a myriad of novel methods aimed at finding reasonable solutions to QUBO problems efficiently, named Ising solvers, given the equivalence of QUBO to the transverse field Ising problem in statistical physics. However, optimization problems of interest to process system engineering are more general than QUBOs, usually involving discrete non-binary and continuous variables and constraints. Although it is possible to map such problems onto QUBOs, the size of solvable QUBOs is inherently limited by the number of qubits in available quantum computers. Moreover, the global optimality of the solution cannot be guaranteed due to the thermal noise and the nature of the algorithms implementable for this problem class in quantum devices. Guaranteeing global optimality has been recently addressed using hybrid quantum-classical algorithms. This work presents a hybrid quantum branch-

and-bound method to address these challenges. This method utilizes the computational strengths of quantum hardware alongside the global search capabilities of the branch-and-bound method. The strategy of breaking the problem into smaller subproblems aligns well with the capabilities of quantum hardware. Contrary to previous work on characterizing the hardness of a branch-and-bound enhanced with quantum algorithms for exploring the nodes, this work provides a practical implementation of the method relying on a classical branch-and-bound enhanced with Ising solvers, potentially leveraging quantum mechanics, in an open-access repository (<https://github.com/SECQUOIA/QuantumBranchAndBound>). The Ising solver serves as a heuristic method, and we further explore the strategy of when and where to apply it within the branch and bound framework. Particularly, limitations in quantum annealers related to the embedding problem set such limitations. Additionally, a customized branching rule and an optimized embedding of QUBO problems are applied to accelerate the convergence of the algorithm, where even without enhancing the commercial solver Gurobi 11 with the solutions coming from Ising solvers, we can achieve a speedup of 20% with respect to the default branching rule in Gurobi. The performance of the proposed method is evaluated on hundreds of QUBO instances from QUBOLib.jl (<https://github.com/SECQUOIA/QUBOLib.jl>), using Gurobi as the branch-and-bound solver, several physics-inspired classical heuristics and the D-Wave quantum annealer as the quantum solver. Detailed computational results are presented, and the proposed method, Gurobi, and the hybrid B&B method are compared with classical heuristic approaches. Improvements in the solution time go up to 14%, and the reduction of the number of nodes explored to prove optimality reaches 16% using Ising solvers with respect to default Gurobi when solving QUBO problems to 0.01% of optimality, highlighting the value of this hybrid approach.

Breaking barriers in two-party quantum cryptography via stochastic semidefinite programming – AKSHAY BANSAL, VIRGINIA TECH – Jamie Sikora

In the last two decades, there has been much effort in finding secure protocols for two-party cryptographic tasks. It has since been discovered that even with quantum mechanics, many such protocols are limited in their security promises. In this work, we use stochastic selection, an idea from stochastic programming, to circumvent such limitations. For example, we find a way to switch between bit commitment, weak coin flipping, and oblivious transfer protocols to improve their security. We also use stochastic selection to turn trash into treasure yielding the first quantum protocol for Rabin oblivious transfer.

4g - WA961 - 512E - **High performance implementation and quantum computing 1** Chair: PRAKASH GAWAS

Steepest-edge simplex algorithms for quadratic programming – SHOJI SHIMIZU, NTT DATA MATHEMATICAL SYSTEMS INC. – Koichi Fujii, Julian Hall

In the field of optimization, one of the algorithms used to solve linear programming (LP) problems is the simplex method, which was proposed by G. B. Dantzig in 1947. The simplex

method was extended to be applicable to quadratic programming (QP) problems in the 1950s and 1960s, but since then, it does not seem to be as popular as the LP simplex method. The QP simplex method has the advantage that the tools of linear algebra used in the LP simplex method (such as LU factorization, updates, etc.) can be utilized without modification, allowing for efficient implementation that fully takes advantage of sparsity. In this talk, we will provide an overview of the QP simplex method and discuss extending the steepest-edge rule to the QP simplex method.

An imitation-based learning approach using Dagger for the Casual Employee Call Timing Problem – PRAKASH GAWAS, POLYTECHNIQUE MONTREAL –

Predictive models are increasingly important in enhancing decision-making processes. This study proposes an innovative approach utilizing Dagger, an imitation learning algorithm, to iteratively train a policy for addressing stochastic sequential decision problems. These problems can be challenging, especially when expert input is costly or unavailable. Our focus lies in crafting an effective expert within the Dagger framework, drawing from deterministic solutions derived from contextual scenarios generated at each decision point. Subsequently, a predictive model is developed to mimic the expert's behavior, aiding real-time decision-making. To illustrate the applicability of this methodology, we address a dynamic employee call-timing issue concerning the scheduling of casual personnel for on-call work shifts. The key decision involves determining the optimal time to contact the next employee in seniority order, allowing them to select a preferred shift. Uncertainty arises from the varying response times of employees. The goal is to strike a balance between minimizing schedule changes induced by early notifications or calls and avoiding unassigned shifts due to late notifications. Unlike traditional predict-and-optimize approaches, our method utilizes optimization to train learning models that establish connections between the system's current state and the expert's wait time. We apply our algorithm using data provided by our industrial partner to derive an operational policy. Results demonstrate the superiority of this policy over the current heuristic method in use.

4g - WB267 - 512E - **Parallel solvers** Chair: YUJI SHINANO

QuBowl: An exact solver for MaxCut and QUBO problems – DANIEL REHFELDT, ZUSE INSTITUTE BERLIN –

This talk introduces the solver QuBowl, which solves maximum-cut and quadratic unconstrained binary optimization (QUBO) problems to proven optimality. We describe key components, such as cutting planes, and also discuss parallelization. QuBowl was able to solve several maximum-cut and QUBO instances for the first time to optimality and is currently leading the Mittelman QUBO benchmark.

Progress on highly parallel ensemble solvers – YUJI SHINANO, ZUSE INSTITUTE BERLIN –

The Ubiquity Generator (UG) Framework was initially developed to parallelize advanced branch-and-bound based solvers. With the introduction of version 1.0, UG transitioned into a high-level task parallelization framework, enabling the parallelization of a broad range of state-of-the-art solvers. This

framework facilitates the integration of multiple algorithmic implementations into a parallel solver. This presentation will highlight several success stories where UG has been utilized to solve previously intractable instances of Mixed Integer Programming, Steiner Tree, Quadratic Assignment, and Quadratic Unconstrained Binary Problems, as well as to set new benchmarks for Shortest Vector Problem instances. Finally, the most recent developments in the UG Framework will be discussed.

Solving Large-Scale Quadratic Assignment Problems by a Parallelized Lagrangian-DNN-Based Branch-and-Bound – KOICHI FUJII, TOKYO INSTITUTE OF TECHNOLOGY – D. Mittelman Hans, Masakazu Kojima, Sunyoung Kim, Yuji Shinano

The Quadratic Assignment Problem (QAP) consists of assigning nodes to an equal number of facilities, where the cost associated with each assignment is quadratic. This classical NP-hard problem remains challenging to solve optimally, and several instances in QAPLIB are still open. This talk will present our advancements in tackling huge scale QAPs using a parallel branch-and-bound method equipped with the Ubiquity Generator framework (UG) and implemented in our solver, ParaQapNB, which utilizes over 100,000 cores. The efficacy of this approach relies heavily on the integrated lower bounding procedures. We employed the Lagrangian and doubly nonnegative (Lagrangian-DNN) relaxation and the Newton-bracketing method to obtain strong and accurate lower bounds. New features of UG, such as Enhanced Checkpoint and Huge Checkpoint File Split, help to efficiently utilize a massive number of cores and manage large branch-and-bound trees. We will present some preliminary numerical results from ParaQapNB, including the tai30a, sko42, and tho40 instances from QAPLIB, which were solved for the first time.

Parallel General-Purpose Dynamic Programming Solvers for Combinatorial Optimization – RYO KUROIWA, UNIVERSITY OF TORONTO – Yuji Shinano, J. Christopher Beck

Domain-independent dynamic programming (DIDP) is a recently proposed model-based paradigm for combinatorial optimization based on dynamic programming (DP). In DIDP, similar to mathematical programming, a user formulates a combinatorial optimization problem as a declarative DP model and then solves the formulated model by calling a general-purpose solver. The currently developed DIDP solvers are based on heuristic search algorithms studied in the artificial intelligence community. In this work, we develop parallel DIDP solvers using parallel heuristic search algorithms. First, we empirically confirm that the multi-thread versions of our parallel DIDP solvers outperform commercial multi-thread mixed-integer programming and constraint programming solvers in multiple combinatorial optimization problems. Then, we develop distributed parallel DIDP solvers using the Message Passing Interface. Our experimental evaluation demonstrates the high scalability of the parallel DIDP solvers in a massively parallel distributed environment.

4g - WC191 - 512E - **Optimization Modeling**
Chair: CARL LAIRD

From Algorithms to Applications: developing and deploying analytical and optimization

models with FICO®Xpress Insight – BRUNO VIEIRA, FICO –

In this talk we will start by presenting the latest developments in the Python interface for the Xpress Solver, outlining recent updates to modelling capabilities and discussing some best practice guidance for developing readable and yet efficient implementations of large-scale optimization models.

To facilitate the deployment and utilization of optimization or analytical solutions implemented in Python or Xpress Mosel, FICO®Xpress Insight provides a flexible environment for creating user interfaces with little to no programming effort. Its interface is designed for both technical and non-technical users interested in running several scenarios and comparing outcomes to understand trade-offs and sensitivities implicit to the optimization problem.

We will demonstrate how an Xpress Insight application can seamlessly be built from a standard Python or Mosel optimization model with intuitive drag and drop of front-end visualization components, including multi-scenario charts and tables. Furthermore, we will demo the latest advancements for interacting with the solver during an optimization run via the new Custom Progress Reporting functionalities.

OMLT: Solving inverse problems over trained graph neural networks – RUTH MISENER, IMPERIAL COLLEGE LONDON – Shiqiang Zhang, Juan S. Campos, Calvin Tsay

OMLT, the Optimization and Machine Learning Toolkit, is a Python package that represents machine learning models (neural networks, gradient-boosted trees, and linear decision trees) within the Pyomo optimization environment. This presentation discusses our recent extension of OMLT to incorporate graph neural networks and its possible application to computer-aided molecular design.

Democratization of High-performance Computing Approaches for Optimization – CARL LAIRD, CARNEGIE MELLON UNIVERSITY –

Large-scale optimization approaches often require tailored solutions and advanced computational implementations that are not possible with off-the-shelf computational tools. The implementation of high-performance computing (HPC) approaches requires significant expertise in low-level compiled programming languages and knowledge of complex, architecture specific parallelization frameworks. The success of high-level languages like Python and Julia, have helped "democratize" advanced HPC approaches, providing students and researchers with access to advanced, performant capabilities in high-level languages. This allows for much more rapid innovation, refinement, and adoption of HPC approaches. I will discuss recent software advances that support democratization of HPC and scientific computing solutions for challenging large-scale optimization problems.

4g - ThA181 - 512E - **A glimpse of computational optimization at Google II** Chair: JUAN PABLO VIELMA

Active Learning via Density-based Space Transformation – HOSSEIN ESFANDIARI, GOOGLE RESEARCH – Hossein Bateni, Samira Hosseinghorban, Alipasha Montaseri

In this work, we study active covering, a variant of the active-learning problem where we are required to find all of the examples with a positive label. We present an algorithm namely Density-Adjusted Adaptive (DAA) learner that queries the labels according to a similarity function based on a density-adjusted metric. Under mild assumptions, we prove that our algorithm discovers all of the positive labels while querying only a truly sublinear number of examples outside the support of positive examples. Our result extends the space of acceptable inputs by modifying the assumptions made by the previous work. Our experiments show that our algorithm significantly improves the prior work on a wide range of standard benchmark datasets, complementing our theoretical result. For example, when measuring the performance by AUC, our algorithm is the best in 19 out of 20 experiments.

An exponential penalization, network decomposition algorithm for large LPs – DANIEL ALJANDRO ESPINOZA GONZALEZ, GOOGLE RESEARCH

Large linear programs, in the order of hundred of millions or more variables and constraints, with large network-flow substructures, naturally arise in applications like mine planning, joint network design and routing problems, pricing problems, and in many other contexts. Most of these instances are impossible to solve with standard commercial LP solvers, but there are some specialized algorithms making headway for this class of problems.

In this presentation we motivate the importance of this class of problems, give a short literature review of the area, both of the theory and of previous implementations and alternative algorithms, and present our current results on an implementation of an exponential-penalization network-substructure-aware parallel algorithm; comparing it against both simplex and barrier methods on MIPLIB linear relaxations, and on a set of large instances collected from different applications.

Capacity management via network flows with nonlinear transformations at the nodes – JOEY HUCHETTE, GOOGLE RESEARCH

We define and study a class of directed network flow problems where nonlinear transformations are applied at the nodes in the network. We discuss how, given an appropriately selected set of nonlinearities, we can model this problem using mixed-integer programming; we also speculate on alternative algorithmic approaches that could bear fruit. We motivate this class of problems with a connection to how Google manages its compute resources, by showing how this abstraction is a useful substrate for modeling complex capacity management problems.

Network-aware scheduling of large ML workloads – CHRISTIAN TJANDRAATMADJA, GOOGLE – Given the scale of state-of-the-art models in machine learning, they are often trained in parallel across multiple hardware accelerators (GPUs or TPUs) that must communicate with each other. While smaller models can rely on dedicated links, larger models may partially need to use the broader data center network, where congestion can cause the training to stall and waste resources. To reduce the chance of network congestion, we can optimize the assignment of accelerator resources with network conditions in mind, which can be formulated as a

generalized quadratic assignment problem with capacity constraints and other requirements. We discuss a MIP-based approach and specialized heuristics to solve this problem from a practical perspective.

4g - FB165 - 512E - A glimpse of computational optimization at Google III Chair: JUAN PABLO VIELMA

ViolationLS: Constraint-Based Local Search in CP-SAT – TOBY DAVIES, GOOGLE –

ViolationLS is a new a Constraint-Based Local Search (CBLS) solver developed by Google's Operations Research team based on the Feasibility Jump Mixed-Integer Programming Heuristic. It uses a novel approach for finding multi-variable moves without requiring any "implicit" or "one-way" constraints traditionally used by CBLS solvers. The new solver significantly outperforms other Local Search solvers on the MiniZinc challenge Constraint Programming benchmarks, both with and without multi-variable moves. ViolationLS is integrated into the parallel portfolio framework of the state-of-the-art CP-SAT Lazy Clause Generation Constraint Programming solver.

MathOpt: Solver independent modeling and execution independent solving in Google's OR-Tools – ROSS ANDERSON, GOOGLE –

This talk introduces MathOpt, a software tool for modeling mathematical optimization problems (e.g., linear programs). MathOpt is part of Google's open source project OR-Tools and is used extensively within Google. MathOpt provides client libraries in various programming languages (including C++, Python, and Java) to build your optimization model, and then solves your model with a supported solver (e.g., Gurobi, SCIP, CP-SAT, GLOP, PDLP). A unified API solves models in process, in a subprocess, or remotely by HTTP or gRPC, which we refer to as execution independent solving. MathOpt supports many advanced features, including nonlinear constraints, duality, rays, LP basis, incremental solves, callbacks, and many solve parameters, all with solver independence and execution independence. The core library of MathOpt (and the C++ client) are written in portable C++ and can run on servers, mobile devices, or even in the browser via WASM. For non-C++ languages, the model building code in the client library is written entirely in that language. As a result, users solving by HTTP or gRPC can quickly get started without installing any C/C++ dependencies. Further, Google provides a (currently) free service for solving MathOpt models remotely.

Google's Operations Research API – DANIEL DUQUE, GOOGLE –

The Operations Research API aims to solve high-level operations research problems using linear programming, mixed integer programming, constraint programming and more. It provides optimization solutions for specific problems to external users through REST or gRPC calls (details at <https://developers.google.com/optimization/service>). In this talk, we present the current API functionalities and discuss some of the software engineering advantages to using such an API.

Duality and decomposition in OR-Tools' MathOpt – JUAN PABLO VIELMA, GOOGLE –

We describe and discuss the duality contracts of OR-Tools'

newest modeling language MathOpt. We study how these contracts can be used for testing of solvers that include linear and quadratic objectives, together with linear, conic and convex quadratic constraints. We also discuss how these contracts lead to abstract oracle-based decomposition methods.

4h - MB962 - 512C - **Game Theory 1** Chair: AH-MADREZA TAVASOLI

A new approach for solving linear programming problems with dense columns – AURELIO OLIVEIRA, UNICAMP – Catalina Jaramillo Villalba

The interior point methods for linear programming problems require solving linear systems in each iteration. Thus, the matrix structure becomes an important issue. For instance, the presence of dense columns in the constraints matrix, makes it crucial to employ specific strategies for solving such linear systems. In this work, we introduce a preconditioner aimed at addressing the linear systems arising from linear programming problems with dense columns. This preconditioner is applied to a modified augmented system to compute the search directions. We show that the resulting linear system remains uniformly bounded with the respect of the iterates as the primal dual interior point method converges to an optimal solution. To validate our approach, we perform computational experiments to linear programming problems with at least one dense column. In comparison with existing methods, our proposal stands out in terms of running time and total number of iteration. The proposed preconditioner successfully solves all the linear systems for the tested linear programming problems showing its robustness.

A New Strategy for Solving Cournot Games – NENA BATENBURG, UNIVERSITY OF COPENHAGEN – Trine Krogh Boomsma

It is well known that Cournot games involving heterogeneous players with affine and decreasing demand functions and convex quadratic production cost functions can equivalently be cast as a mixed linear complementarity problem or a quadratic optimization problem (QP). Market equilibria can thus be computed using a complementarity solver or, more efficiently, a convex optimization solver. When studying such games, it is common to aggregate the consumers and substitute their inverse demand function to eliminate the market price. We argue, however, that the market price should be treated as an explicit variable that is central to the game. We show that the problem can be reformulated as a fixed-point problem (FPP) for the price only, thereby greatly reducing its dimensionality. The FPP can therefore be solved even more efficiently than the QP and also yields a more accurate solution. The FPP further generalizes to spatial games and games involving multiple commodities. We envision that this could vastly improve solution efficiency and accuracy when incorporated into dynamic or multilevel games. We finally consider a Cournot game across the markets for two commodities that may be converted into each other. The FPP formulation reveals the structure of the joint equilibrium and allows an analytical stability and sensitivity analysis. We find that market imperfections in either of the markets extend to the opposite market and may significantly reduce the welfare gains from market coupling.

Mitigation mechanisms for externalities in in-

tertemporal multistage markets: The case of cascaded hydropower plants – GABRIEL CUNHA, PSR –

A well-known use case for intertemporal multistage stochastic optimization problems is the operation of hydro power plants: this is the type of optimization problem that led to the development of novel techniques such as Stochastic Dual Dynamic Programming (SDDP) (introduced by Pereira and Pinto, 1991), currently used as state of the art for a number of real-world applications. This type of strategy has proven highly effective for solving the central planner's optimization problem (cost minimization/welfare maximization), which can be shown to be equivalent to the Nash equilibrium achieved in a competitive market under certain regularity conditions (as demonstrated by Philpott, Ferris, Wets, 2016). However, the key hypotheses for this equivalence result are rarely valid in practice - in particular, for cascaded hydro power plants, the dependency between the water available for downstream plants and operating choices made at upstream power plants yields an externality, given that this coupling is not remunerated (as explored by Lino, Barroso, et al, 2003).

The main goal of this paper is twofold: firstly, to discuss possible strategies that could be used to "internalize" this externality as part of the electricity market design; and secondly, to present quantitative results for market equilibria involving hydropower cascades with and without the proposed mitigating mechanisms, and a varying number of agents participating in the market. Key strategies for the design of so-called "mitigation mechanisms" to correct for these externalities include vertical consolidation, wholesale water markets, and virtual reservoir mechanisms (see Resende and Cunha, 2022); and the quantitative results obtained can help illustrate the relative magnitudes of the effect of externalities VS the effect of market power for various benchmark systems.

Counterparty credit risk under netting agreements: A dynamic game Interpretation – AH-MADREZA TAVASOLI, HEC MONTRÉAL – Michèle Breton

We investigate counterparty credit risk and credit valuation adjustments in portfolios including derivatives with early-exercise opportunities, under a netting agreement. We show that credit risk and netting agreements have a significant impact on the way portfolios are managed (that is, on options' exercise strategies) and, therefore, on the value of the portfolio and on the price of counterparty risk. We derive the value of a netted portfolio as the solution of a zero-sum, finite horizon, discrete-time stochastic game. We show that this dynamic-game interpretation can be used to determine the value of the reglementary capital charges required of financial institutions to cover for counterparty credit risk and we propose a numerical valuation method. Numerical investigations show that currently used numerical approaches can grossly misestimate the value of credit valuation adjustments.

4h - ThA256 - 512F - **Applications of Nash Equilibrium Problems** Chair: ALEXANDRA SCHWARTZ

A Generalized Voting Game for Categorical Network Choices – YUEH LIN, IESEG SCHOOL OF MANAGEMENT – Stefano Nasini, Martin Labbe

This paper presents a game theoretical framework for data classification, based on the interplay of pairwise influences in multivariate choices. This consists of a voting game wherein individuals, connected through a weighted network, select features from a finite list. A voting rule captures the positive or negative influence of an individual's neighbours, categorized as attractive (friend-like relationships) or repulsive (enemy-like relationships). Payoffs are assigned based on the total number of matching choices from an individual's neighbours. We show that our approach constitutes a natural generalization of the K-nearest neighbours' method, establishing the proposed game as a theoretical framework for data classification. Computationally, we construct a mixed-integer linear programming formulation to approach the Nash equilibria of the game, facilitating their applicability to real-world data. Our results provide conditions for the existence of Nash equilibria and for the NP-completeness of its characterization. On the empirical side, we use the proposed approach to impute missing data and highlight its competitive advantage over the K-nearest neighbour's approach.

Strategy Investments in Zero-Sum Games – RAUL GARCIA, RICE UNIVERSITY – Seyedmohammadhossein Hosseini, Malleh Pai, Andrew Schaefer

We propose an extension of zero-sum games in which the row player may first select rows and remove columns from the associated matrix, subject to a budget constraint. Hence, the row player seeks to determine an optimal set of row selections and column removals to purchase and a corresponding (mixed) strategy for the game. We call this the matrix game designer problem (MGD) and show it can be formulated as a mixed-integer linear program (MILP). Our work is motivated by applications of MGD in the security domain and is similar to existing Bayesian Stackelberg game approaches to security scheduling. We differ from previous works in security games in that, whereas previous works focus solely on scheduling existing resources, we instead model the problem of determining which resources to invest into as well as how to employ them. We provide analytical results concerning our formulation, inequalities that strengthen our MILP formulation, and two heuristic approaches for obtaining solutions quickly. Our computational experiments then show that heuristic approaches on average obtain suboptimal solutions with at least a 20% relative gap with those obtained by our MILP formulation.

When Simple is Near-Optimal in Security Games – DEVANSH JALOTA, STANFORD UNIVERSITY – Michael Ostrovsky, Marco Pavone

Fraudulent or illegal activities are ubiquitous across many applications and involve users bypassing the rule of law, often with the strategic aim of obtaining some benefit that would otherwise be unattainable within the bounds of lawful conduct. However, user fraud is detrimental, as it may compromise safety or impose disproportionate negative externalities on particular population groups. To mitigate the potential harms of users engaging in fraudulent activities, we study the problem of policing such fraud as a security game between an administrator and users. In this game, an administrator deploys R security resources (e.g., police officers) across L locations and levies fines against users engaging in fraudulent or illegal activities at those locations. For this security game, we study both welfare and revenue maximization administrator objectives. In both settings, we show that the problem of

computing the optimal administrator strategy is NP-hard and develop natural greedy algorithm variants for the respective settings that achieve at least half the welfare or revenue as the welfare-maximizing or revenue-maximizing solutions, respectively. We also establish a resource augmentation guarantee that our proposed greedy algorithms with just one additional resource, i.e., $R + 1$ resources, achieve at least the same welfare (revenue) as the revenue-maximizing (welfare-maximizing) outcome with R resources that is NP-hard to compute. Finally, since the revenue and welfare maximizing outcomes can differ significantly and, in particular, a revenue-maximization administrator objective can substantially hamper welfare, we extend our security game framework to incorporate contracts. In this contract game, a welfare-maximizing principal offers contracts to a revenue-maximizing administrator to compensate it for the welfare it contributes to the system. For this contract game, we show that our developed algorithmic approaches and theoretical guarantees for the earlier studied welfare and revenue maximization administrator objectives naturally carry forward in studying optimal administrator strategies. Moreover, we introduce a dense-sampling approach to compute a near-optimal solution to the principal's problem of selecting an optimal contract to maximize its payoff. Finally, we present numerical experiments that highlight the effectiveness of contracts in bridging the gap between welfare and revenue-maximizing outcomes.

Dynamic pricing and strategic retailers in the energy sector: A multi-leader-follower approach – ALEXANDRA SCHWARTZ, TU DRESDEN – Giorgia Oggioni, Ann Kathrin Wiertz, Gregor Zöttl

We consider strategic retail pricing in markets, where retail companies buy commodities at fluctuating wholesale prices and resell them to final consumers by applying dynamic retail tariffs. This is of especially large relevance in the context of energy markets, where substantial wholesale price fluctuations are observed. Policy makers currently foster the introduction of such dynamic tariff schemes. From a modelling point of view, we propose a multi-leader-follower problem to investigate the implications of strategic retail pricing and we compare the impacts of implementing dynamic tariffs on retailers and final consumers. Our analysis tackles different aspects: first, we formulate the model and provide theoretical results. Second, we develop algorithms, which solve the multi-leader-follower problem and allow us to characterize the resulting market equilibria. Third, we calibrate and solve our framework based on data of the German retail electricity market for the years 2020 and 2021. This allows us to quantitatively assess the impact of introducing real time prices on retailers' profits and customers' benefits. As our results show, dynamic real-time pricing on the one hand typically increases market efficiency, which confirms previous results obtained without the explicit consideration of strategic behavior. On the other hand, however, as a novel aspect, dynamic real-time pricing turns out to significantly reduce equilibrium profits in case of strategic firms. This effect is especially large in environments with strongly fluctuating wholesale prices.

4h - FA330 - 512F - Computing Equilibria of Games Chair: YICHUAN DING

Learning in Multi-Leader Single-Follower Stackelberg Games – FRANCESCO MORRI, IN-

RIA LILLE – Le Cadre Hélène, Luce Brotcorne

In this work we want to tackle the problem of learning in multi leader single follower (MLSF) Stackelberg games. These games are less studied in the literature than other formulations, even though they are very relevant in a multitude of applications, such as security games, strategic bidding in markets or electricity production. The game we study couples all the leaders between each others and with the follower, through their objective functions. The complexity arises from this coupling, which we encode using a reaction function. This function represents how the other players react to the actions proposed by the leaders. In a standard Stackelberg game, the reaction function is the optimal response; in this presentation, we consider different possibilities, aimed at hiding information between the players. In particular each leader will only observe the reaction function output, and it will have to learn its strategy through that. The learning algorithm is based on the concepts of multi-agent reinforcement learning, and the structure of the algorithm relies on the game properties, since the goal is to have an interpretable and provably convergent algorithm. In particular we also focus on how much information needs to be shared between agents in order to prove the convergence.

Nash Equilibrium of Code-Share Mechanism under Airline Frequency Competition Game with Second Order Cone Programming – ZI YI HUANG, NATIONAL TSING HUA UNIVERSITY – Yu Ching Lee, Chun Han Wang

In the modern aviation industry, strategically setting flight frequencies between airports is critical for increasing market shares, thus playing a pivotal role in the competitive dynamics among airlines. With the growth of the aviation market, the nature of commercial relationships between airlines has evolved from mere competition to the formation of alliances and collaborations. Among various cooperative strategies, the code-sharing agreement emerges as a particularly essential approach.

Code-sharing serves as a mechanism that allows airlines to sell seats on flights they do not operationally control, fostering participation through commercial agreements. This practice is frequently utilized in hub airports to improve connectivity between long-haul or high-frequency short-haul flights. Statistics indicate that more than half of U.S. airlines have embraced code-sharing strategies, with a notable prevalence on routes neighboring popular hub airports.

Despite the prevalence of codeshare in the airline industry, it is worth noting that there is no comprehensive quantitative model that addresses both codeshare strategy and flight frequency competition. Given its importance, this study attempts to develop a frequency competition model that incorporates code-sharing strategies in a game-theoretic context and uses second-order cones to characterize some of these constraints. To find the Nash Equilibrium for this problem, we apply the nature of the optimality conditions of second-order cone programming to build a KKT model containing primal, dual, and complementary slackness constraints. We hope to utilize the proposed model to perform calculations on available data to determine the optimal frequency for each segment and provide strategic information for decisions related to code sharing, suitable partners, and profit sharing among airlines.

Computing Quasi-Proper Equilibria: From a Characterization to a Differentiable Path-

Following Method – CHUANGYIN DANG, CITY UNIVERSITY OF HONG KONG – Yiyin Cao

As a different paradigm from Myerson's properness for rationality on strategy perturbation, the concept of quasi-proper equilibrium was formulated by van Damme for finite extensive-form games with perfect recall. As for quasi-perfect equilibrium, the admissibility of quasi-proper equilibrium gains itself an advantage over Myerson's proper equilibrium. Nevertheless, the formulation contains insufficient information on how to find such an equilibrium. To address this issue, this paper develops an equivalent definition of quasi-proper equilibrium through the introduction of an auxiliary behavioral strategy profile into the formation of an epsilon-quasi-proper equilibrium. To show its computational effectiveness, we illustrate with simple examples how one can utilize the definition to analytically find all quasi-proper equilibria. The exploitation of the definition leads to a differentiable path-following method to compute quasi-proper equilibria. In the development of the method, we incorporate with an extra variable ranging between zero and one a logarithmic-barrier term into the payoff function of a player at each information set and constitute a logarithmic-barrier extensive-form game. At each information set in this barrier game, a player solves against any given behavioral strategies of other players and the payer herself at other information sets a strictly convex optimization problem. An application of the optimality conditions to the barrier game together with the equilibrium condition yields a polynomial equilibrium system. With this equilibrium system, we establish the existence of a smooth path that starts from an arbitrary totally mixed behavioral strategy profile and ends at a proper equilibrium as the extra variable vanishes. Numerical results further confirm the effectiveness and efficiency of the method.

Tight Bounds for the Price of Fairness – YICHUAN DING, DESAUTELS FACULTY OF MANAGEMENT, MCGILL UNIVERSITY – Yifeng Cao, Daniel Granot

A central decision maker (CDM), who seeks an efficient allocation of scarce resources among a finite number of players, often has to incorporate fairness criteria to avoid unfair outcomes. Indeed, the Price of Fairness (POF), a term coined in Bertsimas et al (2011), refers to the efficiency loss due to the incorporation of fairness criteria into the allocation method. Quantifying the POF would help the CDM strike an appropriate balance between efficiency and fairness. In this paper we improve upon existing results in the literature, by providing tight bounds for the POF for the proportional fairness criterion for any n , when the maximum achievable utilities of the players are equal or are not equal. Further, while Bertsimas et al (2011) have already derived a tight bound for the max-min fairness criterion for the case that all players have equal maximum achievable utilities, we also provide a tight bound in scenarios where these utilities are not equal. Finally, we investigate the sensitivity of our bounds and bounds of Bertsimas et al (2011) for the POF to the variability of the maximum achievable utilities, and identify certain patterns of the maximum achievable utilities for which the POF attains its maximum of $1-1/n$.

4i - TA276 - 512F - Large Scale Supply-Chain Network Design I Chair: SEMIH ATAKAN

Relay-Hub Network Design for Consolidation Planning Under Demand Variability – MATH-

IEU DAHAN, GEORGIA INSTITUTE OF TECHNOLOGY
– Onkar Kulkarni, Benoit Montreuil

We introduce the problem of Capacitated Relay Network Design under Stochastic Demand and Consolidation-Based Routing (CRND-SDCR), which focuses on improving the efficiency and resiliency of large-scale relay logistics hub networks against demand variability by integrating tactical planning decisions. We model the commodity demand uncertainty through a finite set of stochastic scenarios and formulate CRND-SDCR as a two-stage stochastic program with mixed-integer recourse. The first stage comprises long-term decisions where we locate relay hubs and decide their respective sizes whereas the second stage involves tactical decisions, where we design a minimum-cost consolidation plan on a multi-commodity flow network for a given stochastic commodity demand realization.

We develop a decomposition-based branch-and-cut algorithm that uses nested Benders decomposition and Integer L-shaped method coupled with sample average approximation. We employ Benders decomposition at two stages: Akin to classical stochastic programs, we solve for each stochastic scenario, the dual of the linear programming relaxation of our second-stage subproblem to add Benders feedback cuts to the first-stage master problem. However, as the LP relaxation of the second-stage problem entails solving a capacitated multi-commodity minimum-cost network flow problem, solving it and its dual by directly feeding it to an off-the-shelf optimization solver is not computationally efficient. Hence, we use Benders decomposition (again) to solve the relaxed second-stage subproblem by decomposing it on the basis of each O-D pair. Here, we leverage the underlying network-flow structure to generate Benders cuts in polynomial time using shortest-path routines. In order to enhance the computational efficiency of our solution approach even further, we add valid inequalities and strengthen the formulations by adding specific dummy variables. Finally, to guarantee the exactness of our solution approach, we add integer L-shaped cuts by solving the second-stage subproblem exactly through Benders decomposition as well.

We conduct an extensive case study pertaining to the design of large-scale relay networks to be used for finished vehicle deliveries of a US-based car manufacturer company that partnered with our research team. We show that our solution approach can obtain provably (near) optimal solutions for practically relevant instances and that our tailored implementation of decomposition-based branch-and-cut scales better in comparison to its classical counterpart with an increase in demand scenarios and network size. We assess the performance of the generated networks by determining minimum-cost consolidation plans for all available stochastic demand scenarios, and compare it with that of relay networks designed for an average parcel demand scenario. Our designed networks show-case significantly higher robustness against demand variability, as they can fulfill more demand by short-haul transportation and at a lower cost. Finally, we benchmark the performance of our generated relay networks against literature-proposed relay networks designed by approximating the tactical planning decisions and depict the importance of considering such tactical planning decisions at the network design stage.

Dynamic discretization discovery for region-based networks – JOCHEN KOENEMANN, AMAZON
– Madison Van Dyk

Practical network design applications are often complicated by temporal considerations such as transit times and time-varying network parameters. While these problems can be modeled with time-indexed formulations, this often results in huge time-expanded graphs. The recently-developed dynamic discretization discovery (DDD) method allows many time-dependent problems to become more tractable by iteratively solving instances of the problem on smaller networks where each node has its own discrete set of departure times.

However, in the current implementation of DDD, all arcs departing a common node share the same set of departure times. This causes DDD to be ineffective for solving problems where all near-optimal solutions require many distinct departure times at the majority of the nodes in the network. To address this shortcoming, we develop a DDD framework where the set of departure times is determined on the arc level rather than the node level. This approach is particularly effective for region-based networks.

Fast Combinatorial Algorithms for Efficient Sortation – MADISON VAN DYK, UNIVERSITY OF WATERLOO; MODELING AND OPTIMIZATION, AMAZON.COM – Kim Klause, Jochen Koenemann, Nicole Megow

Modern parcel logistic networks are designed to ship demand between given origin, destination pairs of nodes in an underlying directed network. Efficiency dictates that volume needs to be consolidated at intermediate nodes. In practice, such consolidation requires parcel-sortation. In this work, we propose a mathematical model for the physical requirements, and limitations of parcel sortation. We then show that, in the general model, it is NP-hard to determine whether a feasible sortation plan exists. We discuss several special settings, where (near-)feasibility of a given sortation instance can be determined efficiently. The algorithms we propose are fast and build on combinatorial witness set type lower bounds that are reminiscent and extend those used in earlier work on degree-bounded spanning trees and arborescences.

Designing Optimization Software for Use Case Scalability – SEMIH ATAKAN, AMAZON.COM –

The scalability of optimization models is a common concern in academic literature, but in industry, scalability to multiple use cases is equally critical. Often overlooked, this aspect results in a proliferation of optimization models, each with similar mathematical structures, but specialized software for solving distinct problems. This presentation delves into the design of software solutions that enable resolution of diverse business challenges at scale.

4i - WA273 - 512G - Large Scale Supply-Chain
Network Design II Chair: ARASH HADDADAN [↩](#)

Multi-Modeling Approach to Real-World Optimization Problems – EUGENE ZAK, INDEPENDENT CONSULTANT –

A distinguishing factor between textbook optimization problems and their real-world counterparts is the necessity of exploring multiple alternative formulations. In this study, I demonstrate the effectiveness of a multi-modeling approach by examining three real-world problems: the Cutting Stock Problem (CSP), the Unit Commitment Problem (UCP), and the Warehouse Storage Space Problem (WSSP). For the CSP, I

investigate both Bin-Packing and Column Generation models. For the UCP, I analyze Mixed Integer Programming (MIP) and Pseudo-Analytical models. Finally, for the WSP, I evaluate both Continuous and Discrete models. Practical applications reveal that embracing a multi-modeling approach significantly expands the arsenal of solving techniques, effectively addressing multiple functional requirements and criteria. Furthermore, it facilitates the delivery of multiple optimal or near-optimal solutions, which is crucial for the end-user's decision-making process.

A decomposition approach for large scale service network design – ARASH HADDADAN, AMAZON – Jochen Koenemann, Daniel Ulch, Marcus Poggi, Madison Van Dyk

We consider the service network design problem (SND). SND-like problems are essential to the operation and the design of supply chain network of online retailers such as Amazon. SND is a well studied problem and is known to be difficult to solve exactly at scale. In this work, we start with a simple decomposition approach that separates the problem into two stages of topology design and timing optimization. Using this decomposition approach we obtain lower bounds and upper bounds on the solutions of the SND problem, however gap between the two bounds could be large. To address this issue, we introduce a family of cuts, called the time-window cuts, to provide the topology design problem with partial timing awareness. We experiment on a family of publicly available service network design instances (Boland et al. 2017), and a class of large instances arising in operations at Amazon. Our experiments show that adding time window cuts increases the quality of the decomposition significantly for both.

Practical stochastic network design for long-term network planning – DAVID MILDEBRATH, AMAZON –

Any real-world service network faces uncertainty from many sources. These uncertainties must be accounted for in day-to-day execution planning, but also in longer-term network planning. Tactical day-to-day planning typically has fewer degrees of freedom, sometimes reducing the complexity of uncertainty modeling. In contrast, long-term network plans must be adaptable to potentially large changes in network structure, in order to provide better plans for tactical execution. This adaptability raises substantial computational challenges. In this talk, we describe practical methods for incorporating stochasticity into long-term service network plans.

Velocity-aware inventory placement – R RAVI, AMAZON SCHOLAR – Cristiana Lara, David Mildebrath Amazon offers millions of unique items for sale on its website. The popularity of these items is not evenly distributed, and the mean weekly demand for different products on Amazon spans multiple orders of magnitude. Because of this enormous scale and selection, deciding how to optimally place this inventory throughout our fulfillment network is very challenging. Directly modeling all products is not feasible without advanced decomposition techniques and massive compute power. On the other hand, down-sampling products typically does not adequately capture the behavior of the “long tail”. We propose a scalable method for modeling both head and tail demand, and placing the inventory to best utilize our network assets. Our model matches the ship and storage capacity of nodes in our

network with the velocity profile of demand in order to minimize fulfillment cost. We use this framework to derive insights on the relationship between placement strategy and customer experience.

4i - ThB963 - 512G - **Industrial applications 1**

Chair: THIBAUT BARBIER



Faster Infeasibility Analysis for Linear Programs – JOHN CHINNECK, SYSTEMS AND COMPUTER ENG., CARLETON UNIVERSITY – Cho Ho (Peter) Lam, Zirui Zhou

Presolving a linear program is important for fast solution and can sometimes detect infeasibility before the reduced model is solved. But presolving typically interferes with finding an Irreducible Infeasible Subset (IIS) of row constraints and variable bounds, the main way to analyze infeasibility. Early attempts to backtrack the set of logical model reductions when the presolver detects infeasibility, with the goal of finding an IIS, were abandoned as impractical. However, the OptVerse solver has now implemented a very fast backtrack capability that greatly speeds IIS isolation whether or not the presolver detects infeasibility. In both cases, the backtracker isolates a small subset of the model that is then subjected to typical IIS isolation procedures. The speed advantage is demonstrated experimentally vs. other major LP solvers.

Novel Branch & Bound Tree Management for Fast Convergence of MIPs on Parallel Computers – DAVID LANDAY, OPTIMAL SOLUTIONS INC, LLC – Vijay Hanagandi, Andrea Lodi, Willem Van Osselaer, Gaurav Singh

Due to the computational complexity of solving large Mixed Integer Programming (MIP) problems, many industry specific problems – requiring complex formulations and having many variables and constraints – are routinely left unsolved or solved sub-optimally, resulting in lost profits and productivity. Parallelization combined with techniques for managing the “Branch and Bound” (BB) tree search and “branching variable selection” (BVS) are often required to reduce the time it takes to solve such problems. We seek to improve upon solutions that leverage parallel computing to solve large MIPs relevant to industrial use cases by introducing two key features: distributed data structures with implicit concurrency semantics and Machine learning-powered branching variable selection. We explore the integration of conflict-free replicated data types (CRDTs) into a parallel implementation of the Branch and Bound algorithm. Previous work has shown that an intelligent BVS technique can improve search efficiency, and the addition of CRDTs can simplify data management in a parallel compute environment; improving sub-problem load-balancing and ensuring consistent replication of global information across distributed resources involved in solving a MIP. In this presentation, we demonstrate the feasibility of our approach on MIPLIB problem instances and we report preliminary results from a proof-of-concept implementation of our framework that uses the open-source HiGHS solver.

Joint Transshipment, Markdown, and Clearance Decisions at a Fast-Fashion Retailer – SIAMAK NADERI, TEACHING FELLOW – Abdullah Dasci, Kemal Kilic

This paper considers a joint transshipment, markdown, and

clearance sales decisions of multiple products for a fast-fashion retailer that owns and operates a large network of retail stores. This problem is motivated by the logistics operations of the largest apparel retailer in Turkey, namely LC Waikiki. Particularly, the emphasis is on the following questions: How should the products be re-distributed among retail stores after a few weeks of demand observation? What should be the price level for each product? And, finally, which products should be marked for clearance sales? To answer these questions, a mathematical model is proposed considering the business rules and the practices of the retailer encounters. We develop a Bender-Decomposition-based heuristic to find upper bounds and a simulated annealing-based metaheuristic to find incumbent solutions, both of which have proven to be quite effective. A Covering-Cut-Bundle approach is used to accelerate the convergence of the Bender-Decomposition algorithm. We have also conducted a set of experiments to uncover the impact of business rules on the retailer's operations and the effectiveness of joint markdown and transshipment decisions.

Optimizing Ticket Pricing with ExPretio's AI ? – THIBAUT BARBIER, EXPRETIO TECHNOLOGIES INC. – Vilmar De Sousa

ExPretio emerges from two decades of academic research in pricing and revenue optimization, integrating expertise from revenue management practitioners and operations research / data science specialists. Focused on helping rail passenger carriers better manage their capacity and pricing to maximize revenue, our challenge lay in understanding the intricate relationships between price, business rules, and passenger buying behaviors. Our AI-powered solution employs an interconnection of mathematical programming, machine learning, and simulation. Every day, we solve more than 125,000 mathematical programs, generating 60 million forecasts for over 20 global clients. This enables precise resource allocation, empowering rail companies to achieve revenue goals while considering broader business objectives. In our presentation, we will discuss our AI framework, the diverse challenges it addresses, and how we continue to collaborate with universities to improve our chain.

4i - FA288 - 512G - Optimization and Validation of Execution Models at Amazon Scale

Chair: QI CHEN

Ideal inventory placement with demand spillover – ZIHAO LI, AMAZON.COM – Zhongxiang Wang

We share an ideal inventory placement strategy for designing an effective inbound network of an online retailer. In an online retailer's fulfillment network, demand is ideally fulfilled from the nearest warehouse with inventory and ship capacity before other, further-away warehouses (defined as spillover). The spillover of demand leads to variable (inventory) turn (measured as the ratio of the average daily shipment units to the average on-hand units) for each item at different warehouses. We build a two-stage stochastic model to capture the variable turns of an item at different warehouses in inventory placement and a discrete-event simulation model to assess its impact on fulfillment. We generate two inventory allocation plans under two inventory placement strategies: assuming all warehouses have the same network turn (Constant Turn) for an item or in-

corporating the variable turns (Variable Turn) at warehouses for an item. Simulation is used to evaluate the fulfillment of the two inventory placement strategies, showing that Variable Turn reduces total fulfillment distance.

Validating Network Planning via Customer Centric Baselines – QI CHEN, AMAZON – Granville Paules

To align resourcing decisions in a large-scale supply chain network, Sales and Operations Planning generates optimal aggregate flow and capacity plans given customer demand and vendor supply projections. These plans must strike a balance between incentives of disparate business units. We evaluate common baseline-generation approaches for assessing plan and execution quality against instances of a network model. We show that our choice of methodology may lead to conflicting interpretations of plan quality.

Dynamic Resource Allocation: Algorithmic Design Principles and Spectrum of Achievable Performances – AKSHIT KUMAR, COLUMBIA UNIVERSITY, GRADUATE SCHOOL OF BUSINESS – Omar Besbes, Yash Kanoria

In this work, we consider a broad class of dynamic resource allocation problems, and study the performance of practical algorithms. In particular, we focus on the interplay between the distribution of request types and achievable performance, given the broad set of configurations that can be encountered in practical settings. While prior literature studied either a small number of request types or a continuum of types with no gaps, our work allows for a large class of type distributions. Using initially the prototypical multi-secretary problem to explore fundamental performance limits as a function of type distribution properties, we develop a new algorithmic property "conservativeness with respect to gaps," that guarantees near-optimal performance. In turn, we introduce a practically-motivated, simulation-based algorithm called RAMS, and establish its near-optimal performance, not only for multi-secretary problems, but also for general dynamic resource allocation problems.

Solving Large Scale Linear Problems with Lagrangian Decomposition – OZLEM BILGINER, AMAZON.COM –

Various Amazon systems optimize shipment-path assignment decisions with capacitated resources. In this process, every shipment is unique, the shipments are assigned to paths, and paths have many capacity constraints. As such, these are very large scale assignment problems, not easily solvable in their canonical form. They consume a lot of memory and they take a lot of time to solve due to their single-node memory and CPU footprint. As Amazon is constantly growing, the challenge is to keep the solution time and memory under control, without compromising the solution quality. In this talk, we present an algorithm to solve these large scale linear assignment problems with Lagrangian decomposition techniques. We discuss the strategies for fast convergence, including problem reformulation and mixed use of cutting planes and subgradient updates.

5c - TA280 - 511F - A.W. Tucker Prize Chair: FRANK E. CURTIS

A.W. Tucker Prize – TO BE ANNOUNCED ,

Tucker Prize Finalists – , – To Be Announced

The speakers and titles will be announced during the opening session of the symposium.

5c - WA316 - 511F - **In memory of Jean-Louis Goffin 1: Recent advances in large scale optimization** Chair: SAMIR ELHEDHLI 

Lifting nonnegativity over general semialgebraic sets to nonnegativity over simple sets –
LUIS ZULUAGA, LEHIGH UNIVERSITY –

In this talk, we propose a universal approach to derive new Positivstellensätze for general semialgebraic sets from ones developed for simpler sets, such as a box, or a simplex. We illustrate the approach by first considering Handelman’s Positivstellensatz over a box, to construct non-SOS Positivstellensätze over any compact semialgebraic set; that is, Positivstellensätze where instead of SOS polynomials, any class of polynomials containing the nonnegative constants, such as SONC, DSOS/SDSOS polynomials can be used to certify nonnegativity. Secondly, by considering the simplex as the simple set, we derive a sparse Positivstellensatz over general compact semialgebraic sets, which does not rely on any structural assumptions of the set. Finally, we show how these results can be extended to the case of unbounded semialgebraic sets. Throughout the talk, we illustrate our results with relevant examples and numerical experiments.

A low-rank method for large-scale semidefinite programs –
RENATO MONTEIRO, GEORGIA TECH –

This talk discusses the theory of a new first-order method for solving large-scale semidefinite programs (SDPs) with bounded domain. It is an inexact augmented Lagrangian (AL) method where the AL subproblems are solved by a hybrid low-rank method. In contrast to the classical low-rank method, the new method finds a near-optimal solution (with provable complexity bounds) of SDP instances satisfying strong duality. Computational results comparing the new method to state-of-the-art solvers on several large SDP instances show that the former finds higher accurate solutions in substantially less CPU time than the latter. For example, it can solve a maximum stable set SDP with a million vertices and 10 million edges with 10^{-10} accuracy in less than 21 minutes

Rectified relaxation: A trustworthy and interpretable deep learning approach to linear optimization with complementary constraints –
JIMING PENG, UNIVERSITY OF HOUSTON –
Wisam Alali

In this talk, we introduce a new deep learning framework, called rectified relaxation for linear optimization with complementary constraints (LOLCC). For this, we first propose to solve LOLCC via solving a series of feasibility problems for systems defined by linear and complementary constraints (SLCC). Then we introduce the new universal relaxation theory, which shows SLCC can be addressed via solving another equivalent bilinear optimization problem. We explore various properties of the bilinear model to derive different sufficient/necessary

conditions at the global optimal solution of the bilinear model itself. We then combine the new theoretical insights with the backward propagation method to design a new learning scheme, called rectified relaxation (ReCR). Global convergence of the ReCR is established under mild conditions and preliminary numerical results will be reported to demonstrate the efficacy of the new DL approach.

5c - WB317 - 511F - **In memory of Jean-Louis Goffin 2: Analytic Centre Cutting Plane Methods: The contributions of J-L Goffin’s work** Chair: TAMÁS TERLAKY 

From the Relaxation Method to Chubanov’s Algorithm, from ACCPM to Tall LPs. –
TAMÁS TERLAKY, LEHIGH UNIVERSITY –

Jean-Luis Goffin made numerous notable contributions through his academic career. In this talk we highlight two themes: (1) The Relaxation Method, that was more recently rediscovered and in variants proved to be polynomial time algorithms to solve linear inequalities and linear optimization (LO) problems; (2) The Analytic Center Cutting Plane Method (ACCPM), that can be put in the context of the recently expanding interest in solving “Tall LO” problems.

A personal view of primal-dual interior-point algorithms for large-scale optimization and some recent developments –
LEVENT TUNCEL, UNIVERSITY OF WATERLOO –

This presentation will have two main ingredients. One ingredient lies in the in the area of primal-dual interior-point algorithms for large-scale convex optimization and it will be a summary of recent research results my co-authors and I have obtained. The other ingredient will be a summary of how some of my research work over the last 30+ years relates to Jean-Louis Goffin’s outstanding contributions.

Analytic Centres and Mixed Integer Programming –
SAMIR ELHEDHLI, UNIVERSITY OF WATERLOO –

J-L Goffin’s work on ACCPM inspired a number of approaches in solving mixed integer programming problems. We explore, exact, heuristic, and decomposition based approaches, as well as recent work on inverse integer optimization.

Local center cutting plane methods framework –
MOHAMED EL FASSI, POLYTECHNIQUE MONTRÉAL –

Centering methods are best known for solving convex nondifferentiable optimization problems which have a polynomial bound on the number of iterations, i.e., the number of calls to the separation oracle. These methods solve, at each iteration, an optimization problem that is often nonlinear, aiming to find a primal-dual center of the approximation polyhedron. While preserving the advantages of these methods, this paper presents a novel primal approach by integrating the Centering (interior point) methods into the Improved Primal Simplex decomposition. We propose a new procedure to generate columns/cutting planes. Specifically, we optimize in the subspace spanned by columns associated with non-degenerate variables of the current solution, where the optimization is efficient and we do not need to calculate a center. If we cannot improve the objective value, we enhance the quality of generated columns in the com-

plementary subspace using a "local center". This novel way of generating columns (or cuts) allows us to reduce the number of calls to the procedure for finding centers.

5c - WC81 - 511F - **In memory of Jean-Louis Goffin 3: A tribute to J-L Goffin's academic influence** Chair: FATMA GZARA 

The value of structure and decomposition in large scale optimization – FATMA GZARA, UNIVERSITY OF WATERLOO –

In this talk, we will describe innovative applications of two key ideas in the solvability of large scale optimization problems: problem structure and primal and dual decomposition. These two concepts are among the favorite to Jean-Louis Goffin and have been very useful in my research work in a wide range of optimization frameworks. We will highlight some successful uses of structure and decomposition to address both scale and difficulty in mixed-integer, nonlinear, stochastic and robust optimization.

Beyond Academia: Remembering the Human Essence of a Great Scholar – MOHAMMAD OSKOOROUCHI, CSUSM –

In this session, we pay tribute to the remarkable life and legacy of Jean-Louis Goffin, a scholar whose influence transcended the boundaries of academia. While his contributions to our field are well-documented, this talk seeks to illuminate a less discussed, yet equally significant, aspect of his legacy: his profound humanity and mentorship.

The power transition: optimization meets policy – MICHEL DENAULT, HEC MONTRÉAL – Florian Mitjana

The transition to a zero-carbon power grid requires far-reaching decisions that will have large financial, societal and environmental impacts. Using very simple optimization tools, we highlight transition pathways and the urgency of moving now. This is joint work with F. Mitjana and P.-O. Pineau at HEC Montréal

An Analytic Center self-Concordant Cut method for the Convex Feasibility Problem – FARANAK MOKHTARIAN, JOHN ABBOTT COLLEGE –

A case of the convex feasibility problem where the set is defined by an infinite number of certain strongly convex self-concordant inequalities is considered. At each iteration, the analytic centre based algorithm adds a self-concordant cut through an approximate analytic center of the current set of localization until a feasible point is found. The algorithm is proven to be a fully polynomial approximation scheme.

5c - ThC978 - 511F - **Special sessions 1** Chair: TBA 

Paul Y. Tseng Memorial Lecture – , – To Be Announced

The speaker and title will be announced during the opening session of the symposium.

Participants

Warning: The schedule is subject to changes. Please consult the [online version](#).

A	Aardal, Karen	WA957 Th1	Anderson, Eddie	WA206	Bahn, Olivier	MC337 F2
	Abada, Ibrahim	WA206	Anderson, Ross	FB165	Baiou, Mourad	WA923 ThA19
	Abbasiparizi, Saeid	WA941	Andreana, Gianmarco	FB172	Bajic, Sanja	TA117
	Abbaszadehpelivasti, Hadi	WB18	Andrianesis, Panagiotis	ThC255	Balasubramanian, Krishnakumar	TA31
	Abi Zeid, Irène	WA941	Ang, Andersen (Man Shun)	FA25		FA113
	Abouee Mehrizi, Hossein	TB168	Angeris, Guillermo	WA251	Balch, Tucker	TA99
	Abueida, Atif	WB185	Angulo, Alejandro	WB958	Balcik, Burcu	FB164
	Adhau, Saket	TB945	Angulo, Gustavo	FB151	Balkanski, Eric	FB295
	Adler, Nicole	FB172	Anitescu, Mihai	FA260	Bambi, Bienvenu	ThC951
	Adly, Samir	ThB907	Anjos, Miguel	OC MP MB119 TB915 TB292	Bandi, Chaithanya	ThA277
	Adulyasak, Yossiri	WC241 ThB155		WC346 WC171	Banerjee, Imon	ThB186
	Afroz Jalilzadeh, Afroz	WB227	Ankem, Venkat Akhil	ThB918	Bangerter, Elise	WA947
	Afsar, Sezin	MB107	Anru, Zhang	FB75	Bani, Abderrahman	TA117
	Agarwal, Arpit	WB146	Ansarlari, Zahra	TB945	Bansal, Akshay	MC163 TC282
	Agarwal, Yogesh Kumar	WB944 ThA948	Anstreicher, Kurt	TB195	Bansal, Ishan	FA150
	Aghaei, Sina	ThA211 FA234 FB141	Antil, Harbir	MB88 TC320	Baradaran Noveiri, Borna	MB219
	Aghasi, Alireza	WA309	Antoniadis, Antonios	MB153	Barahona, Francisco	WA923 ThA19
	Agrawal, Priyank	WA941	Applegate, David	TB207	Baraldi, Robert J	TA90 TC320 WA333
	Agrawalla, Bhavya	TA31	Aprile, Manuel	ThC180 FA228	Barbalho, Hugo	MC177
	Ahmed, Cheikh	WB230	Arash, Rafey	WA169	Barbier, Thibault	ThB963
	Ahmed, Sultan	TA922	Araujo, Kennedy A.G.	FA955	Bargetto, Roberto	FA193
	Ahn, Hyun Soo	MC177	Araujo, Silvio	TB952	Barkley, Peter	ThC255
	Ahn, Miju	FB114	Araujo Pardo, M. Gabriela	WA169	Barré, Théo	FB225
	Aigner, Kevin Martin	TA297 WB939 WB958	Aravkin, Aleksander	WA333	Barry, Alpha Saliou	WA281
	Aissi, Hassan	TB943	Archer, Aaron	MB209	Bartz, Sedi	MB312
	Ajudarte De Campos, Rafael	MB157	Archetti, Claudia	TA938 FA925	Basciftci, Beste	ThA174
	Akbas, Beste	WA206	Aretz, Nicole	ThA361	Bassett, Robert	TA297
	Akgun, Ozgur	FB242	Arjevani, Yossi	ThB93	Bastin, Fabian	MB960 TA59 WB41 W3 ThB935
	Alacaoglu, Ahmet	MC44 TA40	Armbruster, Susanne	ThB244	Basu, Amitabh	TC289 FB9
	Alali, Wissam	WA316	Arsalan, Nazlican	MC57 WP	Basu, Saugata	WA251
	Alarie, Stéphane	WA309	Arsalan, Okan	ThA949 ThB155	Batenburg, Nena	MB962
	Alasseur, Clémence	TB292	Asad, Reza	MB62	Bateni, Hossein	MB209 ThA181
	Alavi, Alireza	ThA954	Asadi, Kavosh	ThA246	Battista, Federico	WA231
	Al Baali, Mehiddin	WB49	Asai, Kensuke	FB114	Baumgärtner, Katrin	FA260
	Albert, Laura	TC953	Aspman, Johannes	ThC38	Baumgärtner, Lukas	WA914
	Albey, Erinc	TB946	Atakan, Semih	TA276	Bauschke, Heinz	TA313
	Alcantara, Jan Harold	FA103	Atallah, Ribal	MB119	Bayraksan, Guzin	WB82 F3
	Alfandari, Laurent	TB168	Atenas, Felipe	MB92 WA338	Bazerghi, Audrey	TB67
	Alioui, Kenza	TC953	Attila, öykü Naz	FB172	Bazotte Corgozinho, Maria Carolina	ThA937
	Alizadeh, Pegah	WC346 WC212	Aubin, Pierre Cyril	WB230	Beach, Ben	FB1
	Alker, Jonas	TA263	Audet, Charles	OC MB302 TA306 TB322	Beaudoire, Grégoire	MC250
	Almeida, Warley	MC202		WA309 ThA954 ThC303	Beauthier, Charlotte	ThC303
	Almen, Aray	WA102	Auger, Anne	ThB300 FA326	Beck, J. Christopher	WB267 FB242
	Alnaggar, Aliaa	ThB296	Augustino, Brandon	MB162 MC163	Beck, Yasmine	WA231
	Aloise, Daniel	TC289 WA281 W1 FA955 FP	Aunsmo, Sigrid	TB945	Becker, Stephen	ThA325
	Al Siyabi, Ahmed Azzan	WB49	Austin, Anthony	TA297	Becu, Berkay	ThB161
	Altschuler, Jason	WB18	Avadhanula, Vashist	WA941	Bednarczuk, Ewa	TB314 TC911
	Alvarez, Ricardo	WC160	Avinash Bhardwaj, Avinash	FA348	Behling, Roger	MB92
	Alves Ribeiro, Ademir	ThC904	Avishan, Farzad	ThA174	Bektas, Tolga	ThC270
	Amaldi, Edoardo	MC298	Axiotis, Kyriakos	MB209	Bélanger, Olivier	WB944
	Amerini, Irene	TA130	Aydin, Nursen	WC942	Bélanger, Valérie	FB164
	Amin, Saurabh	MB254 ThA174	Ayotte Sauv��, ��tienne	MC202	Belbasi, Mohammad Reza	WC942
	Amiri, Arezoo	ThB123	Azizadenesheli, Kamyar	ThA246	Belieres, Simon	TA188
	Amiri, Elahe	FA245	Azuma, Godai	FB36	Belilovsky, Eugene	TC911
	Amirmoeini, Kamyar	MC337			Bellavia, Stephania	MB39 TB60
	Ammann, Sabrina	MB932	B		Bello Cruz, Yunier	MB92
	Amorosi, Lavinia	MC339			Bellon, Antonio	WB329
	Amos, Brandon	TC198			Belotti, Pietro	WB7
	Amrouche, Karim	TC953			Belov, Gleb	MB287
	Anandkumar, Anima	ThA246	Baardman, Lennart	MC177	Benammour, Mohamed Faouzi	TA274
			Babanezhad, Reza	MB62	Benatti, Kleber	ThC904
			Babonneau, Frederic	MC337	Benedetto Proen��a, Nathan	FB919
			Bae, Juhan	TB67		
			Baharlouei, Sina	FA234		

Beneventano, Pierfrancesco	FB154	Boudhar, Mourad	TC953	Cerulli, Martina	WA231 FA925
Bene Watts, Adam	WA97	Boumal, Nicolas	TC299 ThA30 ThB93 FA94	Chai, Simin	FB172
Benfield, David	FB336	Bouras, Ikram	TB292	Chakraborty, Abhishek	ThA42
Bennouna, Omar	ThA174	Bourdin, Loïc	TA306	Chan, Timothy	MB108 FB104
Benson, Hande	FB959	Bousselmi, Nizar	WC903	Chancelier, Jean Philippe	MB107
Ben Tal, Aharon	WB125	Boxiao, Chen	MC177	Chandrasekaran, Karthekeyan	MB223 FA150
Ben Yahia, Olfa	WB944	Brandstetter, Christof	ThB123	Chandrasekaran, Venkat	TB87 ThA139
Berahas, Albert	MB11 MC55 TA40 TB60	Braun, Gábor	WA284 ThB161	Chang, Jingya	FB52
	WB41 ThA42	Braun, Kristin	FB1	Charlin, Laurent	WA957
Bergman, David	WA243 WB236	Braverman, Mark	FA259	Charly, Alizadeh	WC346
Bergou, Elhoucine	TA306	Breitkopf, Jannik	WB321	Chavardès, Raphaël D.	FB164
Berk, Aaron	TB311	Brenner, Aron	MB254	Chavdarova, Tatjana	ThA912
Bernal Neira, David	MC339 TC282 FB151	Brenner, Ulrich	ThB244	Che, Ethan	MB167 ThB186
Bernard, Antoine	WB944	Breton, Michèle	MB962	Chekuri, Chandra	FA150 FB926
Bernardelli, Ambrogio Maria	WC203	Brilli, Andrea	TB322 WC323	Chelbani, Mohand Ryad	TA274
Bernhard, Daniela	WC258	Brockhoff, Dimo	TC305	Chen, Caihua	TC940 WA106 WB939 ThB194
Berthold, Timo	TA269	Brosch, Daniel	ThA214	Chen, Cairong	FA51
Bertsimas, Dimitris	MB6 TA31 TB199 TB168	Brotcorne, Luce	MB119 TB292 FA330	Chen, Feng	MB157
	FB184	Brown, David	TA53 FA259	Chen, Jein Shan	TC911
Besançon, Mathieu	WA97	Brugiaglia, Simone	TB311	Chen, Juntao	ThC252
Besbes, Omar	FA288	Brugman, Judith	FA133	Chen, Kaihuang	WA285
Besner, Alexandre	TA117	Brun, Matthew	ThA204	Chen, Li	MB108
Bestuzheva, Ksenia	FA224	Bruni, Maria Elena	FA193	Chen, Li	TB73
Bettiol, Enrico	MB12	Bucarey, Víctor	WA231 WB232	Chen, Lin	MB209
Beugnot, Gaspard	ThB37	Bueler, Ed	ThC319	Chen, Louis	TA170
Bhathena, Aaresh	TA31	Bueno, L Felipe	FB154	Chen, Qi	FA288
Bhatt, Shubhankar	MB312	Bui, Quang Minh	WB236	Chen, Rui	ThB124
Bhattarai, Sudhan	FB109	Burachik, Regina	WA338	Chen, Rui	TC289 ThB124
Bian, Wei	TA65	Burer, Samuel	MB6	Chen, Sheng Jie	ThB342
Bidkhor, Hoda	FA259	Burgard, Jan Pablo	MB960 MC298	Chen, Sheng Yin	TC96
Bienstock, Daniel	TA152 TB187	Burke, James	WA333	Chen, Wei Kun	ThB342
Bilegan, Ioana C	TB166	Burkowski, Forbes	FB275	Chen, Xiaojun	M3
Bilginer, Ozlem	FA288	Burlacu, Robert	FA348 FB1	Chen, Xiaojun	ThA56
Billen, Malte	ThA949	Busch, Arthur	WB185	Chen, Xin	WC328
Birbil, Ilker	WC942 ThA211	Bustamante Faundez, Pamela	WA231	Chen, Xuxing	TA31
Birgin, Ernesto	TA59 FA955	Butenko, Sergiy	TA324 ThB26	Chen, Yannan	ThB50
Bisaillon, Serge	MB190	Butler, Andrew	MB960	Chen, Yiwen	ThB300
Blais, Bruno	ThA954	Byeon, Geunyeong	ThC116	Chen, Zherui	MB162
Blanchet, Jose	WA106			Chen, Zhi	MB108
Blanco, Víctor	FA351			Chen, Zhi	TB100
Blankenburg, Daniel	ThB244			Chen, Zhiping	MB107
Blanquero, Rafael	MB248			Cheng, Ching An	MC177
Blauth, Jannis	FA245			Cheng, Chun	FA140
Blekherman, Grigoriy	TB195 WA251	Cabo, Marta	FB926	Cheng, Hongyu	TC289 FB9
Blomvall, Jörgen	MB960	Cafieri, Sonia	ThC159	Cheng, Jianqiang	FA140
Blondin Massé, Alexandre	TA117	Cai, Xingju	FA51 FB52	Chenreddy, Abhilash	TA99
Bobbio, Federico Bobbio	FB936	Cai, Xufeng	MC44	Cheriyen, Joseph	FA150
Bodard, Alexander	ThB226	Caldwell, Bethany	WA338	Cherkesly, Marilene	WA947
Bodur, Merve	MB108 TB945 TB952 ThB296	Callebaut, Hugo	WC212	Chicoisne, Renaud	ThB918
	FB109	Camelo, Sergio	TA99 WB147	Chinneck, John	ThB963
Boisneault, Luca	MB302	Cameron, Ben	WB185	Chiu, Hong Ming	WA208
Bok, Jan	WA169	Cameron, Kathie	MB223 WA169	Chmiela, Antonia	MC339 ThB161
Bollapragada, Raghu	MB11 TA40 WB41	Campos, Juan S.	WC346 WC191	Cho, Nathan	ThB33
	ThA42	Candogan, Ozan	MB254	Choi, Jiyoung	TA138
Bollapragada, Raghu	TA40 TB60 ThA325	Canelas, Alfredo	FA140	Chowdhury, Anirban	WA97
Bolte, Jerome	TB915 WB331	Canrong, Zhang	ThA954	Christensen, Adam	MB287
Bomze, Immanuel	MB12 TB5 FB929	Cao, Liyuan	ThA325	Christophel, Philipp	TB340
Bonami, Pierre	WA927 WB341	Cao, Nairen	WA931	Chu, Hong T.M.	TA65
Bondarev, Anton	WB321	Cao, Yifeng	FA330	Chu, Ya Chi	FB141
Boob, Digvijay	FB114	Cao, Yiyin	FA330	Chzhen, Evgenii	ThC129
Boomsma, Trine Krogh	MB962	Cappart, Quentin	WA243 FB172	Cipolla, Stefano	MB39
Borgwardt, Steffen	TC359	Caramanis, Michael	ThC255	Cire, Andre Augusto	TA53 WB236
Bornay, Bahman	WA947	Cardonha, Carlos	ThB935	Clarke, Stefan	WC203
Borndörfer, Ralf	FB268	Carmon, Yair	ThC129	Clautiaux, François	TA188
Börner, Pascal	ThB118	Carpentier, Pierre	MB107	Clow, Alexander	TA263
Borst, Sander	ThC180	Carrasco, Miguel	FA140	Cocchi, Lorenzo	MB302
Borst, Sander	TA930	Carrizosa, Emilio	MB248 MC298	Coelho, Leandro	MB157 ThC270
Bortoletto, Enrico	FB268	Cartis, Coralía	TP ThB35 FB70	Coelho, Leandro C.	WA931
Bot, Radu Ioan	TA313 WA251	Carvalho, Margarida	MB119 MC202 TB915	Cohen Addad, Vincent	WA931
Boucher, Jonathan	FB164		WB236 ThA937 ThP Th1 FB936	Cole, David	FA260
Bouchet, Pierre Yves	TA306	Catanzaro, Daniele	ThB352	Concas, Emanuele	ThA937
		Cebulla, Dominik H.	WA914		

Condat, Laurent	TA130	Dellacherie, Stéphane	WB41	Dong, Hang	ThB155
Coniglio, Stefano	WC353	De Loera, Jesus	TB195	Dong, Wanping	TB60
Consolo, Antonio	MC298	Del Pia, Alberto	FA348	Dong, Zhilong	MC14
Contardo, Claudio	WA243 FA955	Demir, sebnem Manolya	ThA949	D'onofrio, Federico	FB929
Contreras, Juan Pablo	WB227	De Moor, Bart	MC908	D'orazio, Ryan	ThA128
Contu, Francesco	TA192	De Moor, Danique	MB6 FB184	Dos Reis, Goncalo	ThA949
Coppola, Corrado	TA130	De Moor, Danique	FB184	Dowson, Oscar	MB344 WP
Corbean, Elisa	WB958	Dems, Amira	WC171 ThB155	Draganescu, Andrei	ThA361
Corbett-Davies, Sam	FA234	De Muinck Keizer, Willem	TC23	Dragotto, Gabriele	WC203
Cordeau, Jean François	ThA948 ThB155 Th2	Denault, Michel	MC337 TC142 WC81	Drees, Martin	FA351
Cormier, Lindsay	MB219	Deng, Qi	WB120 TC940 FB114	Dressler, Mareike	WB906
Correcher, Juan Francisco	WB917	Deng, Zhanwang	WA208	Drusvyatskiy, Dmitriy	TA115 ThA354 ThB93
Cory Wright, Ryan	TA31	Den Hertog, Dick	FA133 FB184	Dumochelle, Justin	FB104
Costa, M. Fernanda P.	ThA24	Den Hertog, Dick	MB6 TB168 ThA211 FB184	Duque, Daniel	FB165
Côté, Alain	ThC261	Dentcheva, Darinka	WA102	Durante, Valentin	ThA214
Côté, Jean François	TB945 WA931 ThA954	Denzler, Sebastian	WB939	Dursunoglu, Cagla	ThA949
Cotton Gagnon, Anne	FB164	Deo, Anand	WC258	Dussault, Jean Pierre	ThA354
Crainic, Teodor Gabriel	MB190 TA192 TB166	De Oliveira, Wellington	TA112	Dvorkin, Yury	TB187
	WA947 FA193	Derakhshan, Mahsa	FA259	Dvurechensky, Pavel	ThB226
Crimmins, Braden L	TC96	Dereich, Steffen	WA941	Dzahini, Kwassi Joseph	FB327
Criscitiello, Christopher	ThB93	Derezinski, Michal	MC55		
Cristian, Rares	TB73	Derval, Guillaume	FA933		
Cuadrado, Marlyn D.	WC171	De Santis, Marianna	MC339 ThB34		
Cui, Chunfeng	ThB907 FA51	Desaulniers, Guy	MB12 TA274 TB952 T3		
Cui, Ying	TA112 TB87 FA94		ThA949 ThB918 Th2 FA955	Eagles, Nancy Mae	WB185
Cunha, Gabriel	MB962	Designolle, Sébastien	WA97	Eckstein, Jonathan	WA338 ThB200
Curcic, Mihajlo	TA117	Deslauriers, Alexandre	WB41	Edelman, Alan	WA251
Curtis, Frank E.	M1 TA40 TA280 ThB200	De Sousa, Vilmar	ThB963	Ehrenmann, Andreas	WA206
Custodio, Ana Luisa	TC305 WC323	De Sterck, Hans	FA25	Eichfelder, Gabriele	ThB34
Cutler, Joshua	ThA354 ThB93	Devera, Ryan	FA94	El Fassi, Mohamed	WB317
		De Vicente Jiang, Daniel	TB5	Elgersma, Maaïke	WA957
		De Vries, Sven	FA150	El Halabi, Marwa	TB199 ThA19
		De Weerd, Mathijs	WA957	Elhedhli, Samir	WA316 WB317
		De Wolff, Timo	MB932	El Housni, Omar	TA229 ThB124 FA133 FB225
		Dey, Santanu	WA251 WA8 ThB161 ThB124	Elyasi, Milad	ThB155
		Dey, Shibshankar	FB936	Emadikhia, Mohsen	WA243
		Deza, Anna	ThB200	Emine, Youssef	WA281 ThC928 FB926
		De Zegher, Joann	TA99	Ent, Ryan	WA206
		Dhulipala, Laxman	MB209	Epperly, Aidan	WA97
		Diago, Miguel	WA309	Erazo, Ignacio	ThB950
		Diago, Miguel	ThC261	Erdos, Péter	ThC355
		Diakonikolas, Ilias	WA196	Ergun, Ozlem	ThP
		Diakonikolas, Jelena	MB11	Erlebach, Thomas	WA931
		Diakonikolas, Jelena	MC44 WA196	Ernst, Damien	FA933
		Diamandis, Theo	WA251	Errico, Fausto	MB12 MB119
		Diamant, Adam	TA53	Eschenhagen, Runa	TB67
		Dias, Fernando	ThB918	Escobar Vargas, David	TA192
		Dias Garcia, Joaquim	MB344 MC339 TB22	Escobedo, Adolfo	TA922
		Diaz, Mateo	TB207 WC45 ThA354 ThB93	Esfandiari, Hossein	ThA181
		Diego, Cifuentes	WA8 WB222 ThB93 FA89	Esperet, Louis	MB223
			FA351	Espinoza Gonzalez, Daniel A.	ThA181
		Diego, Ponce	MC286	Essodaigui, Siham	TA188
		Diehl, Moritz	TC920 FA260	Esteban Perez, Adrian	TA134 ThA46
		Digalakis, Vassilis	TB199 TB168 ThB200	Estrada Garcia, Juan Alberto	TC96
		D'ignazio, Kyra Ann	TB292	Euler, Ricardo	FA925
		Di Gregorio, Silvia	FB9	Evans, Brecht	TB131
		Dilkina, Bistra	WA243		
		Dimitrakopoulos, Roussos	WA927		
		Ding, Kuangyu	MB71 FA83		
		Ding, Lijun	TB136 ThA74		
		Ding, Yichuan	FA330	Faenza, Yuri	FB929
		Dinh, Khanh	MC283	Fahrbach, Matthew	MB209
		Dinitz, Michael	FB295	Fakoor, Rasool	ThA246
		Dionísio, João	FA224	Fampa, Marcia	WB15 FP
		Diouane, Youssef	M1 TA306 WA333 WA309	Fang, Cong	TB131
			ThC303	Fang, Kaiwen	ThC116
		Dirkse, Steven	MB287	Fantuzzi, Giovanni	ThB37
		Djeumou Fomeni, Franklin	WC353 ThC270	Fanzeres, Bruno	ThB33
		Dkhili, Nouha	TA938	Farrell, Patrick	ThC319
		Dogeas, Konstantinos	WA931	Fathollahifard, Amirmohammad	MB190
		Dommel, Paul	TA956	Fatkhullin, Ilyas	TB131

D

Dadush, Daniel	TA930 ThC180
Dagdougui, Hanane	WC171
Dahal, Hari	WA284
Dahan, Mathieu	TA276
Dahdah, Steven	FB36 FB929
Dahl, Joachim	WB120
Dai, Yu Hong	MC902 MC14 ThB342
Dai, Yutong	ThB200
Dan Dije, Ibrahim	ThC270
Daneshvar, Mohammad	MC202
Dang, Chuangyin	FA330
Dang, Nguyen	FB242
Dangel, Felix	TB67
Dansou, Harcénage	FA955
Dao, Minh	WB906
Darlay, Julien	TB340
Daryalal, Maryam	ThA937
Dasci, Abdullah	ThB963
Davanloo Tajbakhsh, Sam	MB71
Davarnia, Danial	WC241
Davies, Toby	FB165
Davis, Damek	TA115
De Boeck, Jérôme	WC212
De Carli Silva, Marcel K.	FB919
De Carvalho Ramos, Andre Felipe	ThB33
De Castro, Philip	ThB34
De Castro, Yohann	WB230
Defrain, Oscar	MB223
Dehghani, Saeedeh	TB166
De Jong, Gerben	FB172
Deka, Deepjyoti	ThA77 FB1
De Klerk, Etienne	WB18
Delage, Erick	TA99 TA170 ThB155 ThC252
	FB104 F2
De Lara, Michel	MB107
Delarue, Arthur	WB147
Deligny, Carlyle	ThA19

E

Eagles, Nancy Mae	WB185
Eckstein, Jonathan	WA338 ThB200
Edelman, Alan	WA251
Ehrenmann, Andreas	WA206
Eichfelder, Gabriele	ThB34
El Fassi, Mohamed	WB317
Elgersma, Maaïke	WA957
El Halabi, Marwa	TB199 ThA19
Elhedhli, Samir	WA316 WB317
El Housni, Omar	TA229 ThB124 FA133 FB225
Elyasi, Milad	ThB155
Emadikhia, Mohsen	WA243
Emine, Youssef	WA281 ThC928 FB926
Ent, Ryan	WA206
Epperly, Aidan	WA97
Erazo, Ignacio	ThB950
Erdos, Péter	ThC355
Ergun, Ozlem	ThP
Erlebach, Thomas	WA931
Ernst, Damien	FA933
Errico, Fausto	MB12 MB119
Eschenhagen, Runa	TB67
Escobar Vargas, David	TA192
Escobedo, Adolfo	TA922
Esfandiari, Hossein	ThA181
Esperet, Louis	MB223
Espinoza Gonzalez, Daniel A.	ThA181
Essodaigui, Siham	TA188
Esteban Perez, Adrian	TA134 ThA46
Estrada Garcia, Juan Alberto	TC96
Euler, Ricardo	FA925
Evans, Brecht	TB131

F

Faenza, Yuri	FB929
Fahrbach, Matthew	MB209
Fakoor, Rasool	ThA246
Fampa, Marcia	WB15 FP
Fang, Cong	TB131
Fang, Kaiwen	ThC116
Fantuzzi, Giovanni	ThB37
Fanzeres, Bruno	ThB33
Farrell, Patrick	ThC319
Fathollahifard, Amirmohammad	MB190
Fatkhullin, Ilyas	TB131

Fattahi, Salar		TA31	Gao, Wenzhi		ThB194	Grady, Daniel		WB350
Favrel, Arthur		ThC261	Gao, Ying		FA51	Graham, Naomi		WC45
Fawzi, Hamza	WA97	ThB37 FA83	Garaix, Thierry		TC920	Granade, Clémence		WB41
Feldmann, Christian		WC346	Garcia, Raul		ThA256	Granot, Daniel		FA330
Feng, Yasong		ThB137	Garcia, Sergio		ThA949	Grapiglia, Geovani		FA326
Ferber, Aaron		WA243	Gardi, Fred		MB932	Grappe, Roland		ThA924
Fercoq, Olivier		ThC904	Garg, Pranav		ThA214	Gratton, Serge	ThB35	FB70
Fernandes, Edite M.G.P.		ThA24	Garg, Sachin		MC55	Grigas, Paul		TB73
Fernández De Dios, Manuel		ThA77	Garner, Casey		FA17	Grigoriev, Alexander		WA923
Fernández Fisac, Jaime		WC203	Gaspar Cunha, A.		ThA24	Grimmer, Benjamin	MC44	WB18
Ferrarini, Luca		FA228	Gatmiry, Khashayar		MB153	Gross, David		TC23
Ferris, Michael		MB287	Gawas, Prakash		WA961	Grossmann, Ignacio	TA956	FB151
Ferry, Julien		TA297	Ge, Dongdong	TC940 WB120 ThA135	ThB194	Grout, Logan		FA150
Fiedorowicz, Barbara Anna		FB9	Ge, Yuchen		WB178	Grübel, Julia		ThB118
Filomeno Coelho, Rajan		ThC303	Geiersbach, Caroline		TB111	Gruber, Aritanan		ThB935
Fiorini, Samuel	ThC180	FA228 FA150	Gelß, Patrick		WA97	Grunau, Lukas		ThA277
Fischer, Andreas		WB906	Gemander, Patrick		FA348	Gruson, Matthieu		FA955
Fischer, Anja	WA923	ThB918	Gendreau, Michel	WA947 WP	FA193	Grütter, Adrian		ThA949
Fischer, Frank	WA923	FA113	Gendreau, Michel	TA930	ThA949	Gschwind, Timo		TA192
Fischer, Vera		ThA343	Gendron, Bernard		WA947	Gualandi, Stefano		WC203
F. Javier, Heredia		WC171	George, Erin		FB1	Guedes Ayala, Monse		WB222
Flatberg, Truls		FB959	George, Ian		MC163	Guillot, Gaël		MB119
Florea, Mihai I.		WB350	Georgiou, Angelos	TA170	ThC252	Gumus, Mehmet		ThA954
Flores Quiroz, Angela		ThA24	Georgia, Perakis		TB73	Guo, Bingni		ThA139
Foley, Angele		WB185	Gervais Dubé, Mathieu		TC953	Guo, Cheng		TB187
Foos, Josefine		ThB244	Gess, Benjamin		MB71	Guo, Xinyi		ThA954
Forbes, James Richard	WA914	FB36 FB929	Gfrerer, Helmut		TB311	Gupta, Anupam		WB146
Forel, Alexandre	ThB935	ThC928	Ghaboulian Zare, Sara		MC337	Gupta, Shagun	MB11 TA40	ThA42
Fornier, Zoe		TB952	Ghaddar, Bissan	WC160	ThA77	Gupta, Swati		FA182
Fortz, Bernard	TB292	WC212	Ghadimi, Saeed	WA309	FA182	Gupta, Vishal		ThB194
Fourer, Robert		MB287	Ghadiri, Mehrdad		FA182	Gupte, Akshay	MC286	TB66
Foussoul, Ayoub		FA133	Ghannam, Mohammed		ThB342	Gurbuzbalaban, Mert		MB11
Fowkes, Jaroslav		WA931	Ghassemi, Mohsen		TA99	Gutman, David	WA905	WB350
Franchini, Giorgia		TB60	Ghazouani, Sami		TA938	Guyslain, Naves		FB926
Francobaldi, Matteo		TA297	Ghezzi, Andrea		TC920	Guzmán, Cristóbal	MB11	WC45
Frappier, Mathieu		ThA354	Ghosal, Promit		TA31	Gzara, Fatma		WC81
Freitas Paulo Da Costa, Bernardo		TB22 TB952	Ghosh, Soumik		MC163			
			Ghuge, Rohan		WB146			
Frejinger, Emma		MB190 T1	Giang Tran, Khanh Hung		ThB26	H		
Freund, Robert	TB207	WC45	Gilbert, Jean Charles		ThA354	Ha, Yunsoo	WA309	ThA325
Frey, Jonathan		FA260	Giovanidis, Anastasios		WC212	Haas, James		ThA214
Friedlander, Michael		WC45	Gissler, Armand		ThB300	Habiboullah, Mohamed Laghdaf		WA333
Friedrich, Ulf		ThA924	Gläser, Max		ThB161	Haddadan, Arash		WA273
Frujeri, Felipe		MC177	Gleixner, Ambros	WC266	ThB342 FA224	Hadian, Majid		MB209
Fu, Chenyi		MB108	Glineur, François		TB340 WC903	Hager, Lukas		FB1
Fu, Mingyang		TA170	Glista, Elizabeth		MC783	Hajiaghayi, Mohammad		MB209
Fu, Thomas		MB209	Goedhart, Rob		ThA211	Halderman, Alex		TC96
Fügenschuh, Armin		FA224	Goel, Asvin		FC353	Hall, Georgina	TB199	WB222
Fuji, Terunari		WA334	Goel, Sharad		FA234	Hall, Julian	TA269	WA961
Fujii, Koichi	WA961	WB267	Goerigk, Marc	TA297	FA133	Hallé Hannan, Edward		ThC303
Fukasawa, Ricardo	MB157	TA297 WA243	Golak, Julian		WA923	Halskov, Thomas		MB248
		WB236 WB15	Goldberg, Noam	ThB200	ThB296	Hamad, Fadi		TB69
Fukuda, Ellen Hidemi		MC908	Gomez, Andres	TA31 TB195	WA8 WC346	Hamid, Faiz		ThA948
Furini, Fabio		WC353		ThA211	FB141	Hammadia, Abdelghani		TA117
Furman, Kevin		FB151	Gómez Casares, Ignacio		WB7 ThA77	Han, Dahye		ThB124
			Gómez Vargas, Nuria		MB248	Han, Deren	WB49	FA51
			Gonçalves, Maurício		TB952	Han, Lei		ThA246
			Gondzio, Jacek	MB39	WB906	Han, Qiushi		TC940
			Gong, Rui		FA351	Han, Shaoning		TB195
Gaar, Elisabeth		ThB123	González Díaz, Julio		WB7 ThA77	Han, X.Y.		WB331
Gabl, Markus	TB5	TB195	González Rodríguez, Brais		WB7 ThA77	Hanagandi, Vijay		ThB963
Gabriel, Steven		WA206	González Rueda, ángel Manuel		ThA77	Hanasusanto, Grani A	TB100	TB22
Gachechiladze, Mariami		TC23	Goodwin, Ariel		MB312	Handa, Marouan		ThB910
Gaebler, Johann		FA234	Göß, Adrian		FA27	Hans, D. Mittelmann		WB267
Gagné, Nicolas		ThB935	Gotoh, Jun Ya		FB114	Hansen, Nikolaus		ThB300
Gagnon, Martin		ThC261	Gottlieb, Robert		TA324	Hao, Hao		WB82
Galli, Leonardo		WB68	Gould, Nick		WA931	Hao, Zhaowei		FA140
Gamache, Michel		ThB918	Gounaris, Chrysanthos		FB1	Hare, Warren	TB322	ThB300
Gamboa, Carlos		ThB33	Gouveia, João		TC359	Harris, Mitchell		FB87
Gao, Bin		MB62	Goyal, Vineet	TA229	FA133	Harsha, Pavithra		TB73
Gao, Rui	MB108	TB187 WA102 WA106	Govens, Florentin		TA59			

G

Gaar, Elisabeth		ThB123	Gao, Wenzhi		ThB194	Grady, Daniel		WB350
Gabl, Markus	TB5	TB195	Gao, Ying		FA51	Graham, Naomi		WC45
Gabriel, Steven		WA206	Garaix, Thierry		TC920	Granade, Clémence		WB41
Gachechiladze, Mariami		TC23	Garcia, Raul		ThA256	Granot, Daniel		FA330
Gaebler, Johann		FA234	Garcia, Sergio		ThA949	Grapiglia, Geovani		FA326
Gagné, Nicolas		ThB935	Gardi, Fred		MB932	Grappe, Roland		ThA924
Gagnon, Martin		ThC261	Garg, Pranav		ThA214	Gratton, Serge	ThB35	FB70
Galli, Leonardo		WB68	Garg, Sachin		MC55	Grigas, Paul		TB73
Gamache, Michel		ThB918	Garner, Casey		FA17	Grigoriev, Alexander		WA923
Gamboa, Carlos		ThB33	Gaspar Cunha, A.		ThA24	Grimmer, Benjamin	MC44	WB18
Gao, Bin		MB62	Gatmiry, Khashayar		MB153	Gross, David		TC23
Gao, Rui	MB108	TB187 WA102 WA106	Gawas, Prakash		WA961	Grossmann, Ignacio	TA956	FB151

Hartisch, Michael		TA297	Huang, Cunxin		FB327	Jiang, Xiaotian		MC902
Harvey, Nick		WB68	Huang, Jing		WA169	Jiang, Xin	TA922	TB943
Hasan, Faruque		ThA24	Huang, Michael		ThB194	Jiang, Yuntian		ThA135
Haskell, William B.		ThA101	Huang, Queenie		TB5	Jiawei, Xu		ThA246
Hassani Goodarzi, Asefeh		ThB950	Huang, Wenjie		ThA101	Jie, Kewei		TB946
Hassanzadeh, Alborz		FB164	Huang, Yizhe		WC328	Jin, Billy		ThA293
Haubner, Johannes		WA914	Huang, Zi Yi		FA330	Jin, Qiujiang	MC55	ThA135
Haurie, Alain		MC337	Huber, Olivier		MB287	Jin, Zhang		ThB913
Hauser, Raphael		ThB35	Huchette, Joey	MB344	ThA181	Jo, Nathanael		ThA211
He, Chang	ThA135	FA94	Huguet, Thomas		ThB318	Jo, Nathanael		ThA211
He, Kerry		FA83	Huiberts, Sophie		TA930	Johansson, Mikael		ThA128
He, Long	FA140	FB141	Huot Chantal, Francis		MB960	Johansson S, Ingemar		WC212
He, Niao		TB131	Hurkens, Cor		FA262	Joormann, Imke		MB932
He, Ye		WB350				Joret, Gwenaël	ThC180	FA228
Heimann, Sophia		FA245	I			Josz, Cedric	TA138	ThB907
Heinkenschloss Matthias, Matthias		TB111				Jourani, Abderrahim		ThB913
Held, Stephan	ThB244	FA245				Julien, Esther		WB230
Hélène, Le Cadre		FA330	Iancu, Dan	TA99	TA134	Jung, Hojin		TA956
Helisto, Niina		WA957	Ibn Brahim, Marouane		TA229	Justin, Nathan	ThA211	FB141
Hell, Pavol		WA169	Ibrahimpur, Sharat		MB153			
Hellani, Ali		TA938	Ickstadt, Constantin		ThB123	K		
Hellemo, Lars		TB945	Im, Haesol		FB275	Kai, Zhou		WA285
Helton, Bill		WA97	Im, Hyunki		TB73	Kalcsics, Jörg		ThA949
Hendel, Gregor		MB344	Im, Sungjin		FB295	Kalogerias, Dionysis	WB906	ThB935
Henderson, Shane		ThB35	Ioannis, Mitliagkas		ThA128	Kamil Khan, Kamil		FA27
Hendrickx, Julien M.		WC903	Iommazzo, Gabriele		WA97	Käming, Isabella		WB906
Hendrix, Eligius Mt		ThA24	Irnich, Stefan		T3	Kamp, Dominik		FA224
Hennebelle, Ulysse		ThB124	Ishfaq, Haque		ThA246	Kamri, Yassine		WC903
Henrion, Didier		WB329	Ivaturi, Vijay		MC283	Kandora, Paul Niklas		TB5
Henrion, René		TB111	Iwata, Satoru		WA923	Kannan, Rohit	ThA77	FB1
Hernández Cruz, César	MB223	WA169	Iyengar, Garud		ThB194	Kannengießer, Lukas		WB321
Hershkovitz, Ellis		MB239	J			Kanno, Yoshihiro		WA333
Hertz, Alain	MC250	TC953				Kanoria, Yash		FA288
Herzog, Roland		ThB318	Jabbarzadeh, Armin		ThB950	Karabulut Kurt, Gunes		WB944
Hewitt, Mike		TA188	Jacob, Isabel		FB52	Karaca, Utku		WC942
Hidayatullah, Alvian Alif		MB29	Jacquillat, Alexandre	TA188	TB199	Karagoz, Aysenur		WB178
Hijazi, Hassan		TA152	Jaeger, Annika		FA262	Karahalios, Anthony		FA933
Hildebrand, Robert	FA228	FB1	Jafari, Ali		MB219	Karakolios, Kleanthis		WB147
Hinder, Oliver	TB69	TB207	Jain, Rhea		MB239	Karamanli, Cem		ThA325
Hintermüller, Michael		TA90	Jalota, Devansh		ThA256	Karangozishvili, Elene		WB185
Hnatiuk, Arsen		ThA361	Jans, Raf	TB952	TC289	Karimi, Mehdi		FA17
Ho, Chin Pang		TB100	Jansen, Erik		WB917	Karimi, Mohammad	ThA949	
Hoàng, Chinh		WB185	Jaramillo Villalba, Catalina		MB962	Karmarkar, Tanmaya		FA27
Hoang, Hung P. Hoang		FA245	Jara Moroni, Francisco		WB232	Karrenbauer, Andreas		TA930
Hoeksma, Ruben		MB153	Jarry Bolduc, Gabriel		TB322	Kashaev, Danish		TA930
Hoen, Alexander		WC266	Jayaram, Rajesh		MB209	Kassem, Zeyad		TA922
Hoheisel, Tim	M2	M3	Jayaswal, Sachin		WB944	Kassing, Sebastian		MB71
Hojny, Christopher		FA262	Jena, Sanjay Dominik	MC202	FB172	Kawasaki, Hidefumi		ThA19
Homem De Mello, Tito		WB232	Jensen, Bjørn		MB88	Kaya, Yalçın		WA338
Homsí, Gabriel		MC202	Jentzen, Arnulf		WA941	Kazachkov, Aleksandr M.		MB157
Hong, Mingyi		MC902	Jeong, Jaehee		WC160	Kazda, Kody		WA905
Ho Nguyen, Nam		WB350	Jeong, Juyoung		WB906	Ke, Chengyu		FB114
Hooker, John		WC241	Jha, Devesh		WC258	Keehan, Dominic	TB22	WB176
Horländer, Andreas		WB236	Ji, Yao		WA196	Keller, Nathan		MC908
Hosoda, Junko		ThB155	Jia, Menglei		MB157	Kelly, Sarah		MB6
Hosseinghorban, Samira		ThA181	Jia, Zhichao	MC44	WB176	Kempke, Nils Christian		TA269
Hosseini, Seyede Saeede		WC241	Jia, Zhuangzhuang		TB22	Kermani, Ali	ThA948	
Hosseinian, Seyedmohammadhossein		ThA256	Jiang, Bo		ThA135	Kerrigan, Eric C.		WA914
Houdouin, Pierre		FB929	Jiang, Daniel	MB167	TA53	Kesselheim, Thomas		MB153
Hougardy, Stefan	ThB244	FA245	Jiang, Hansheng		WA941	Khademi, Abbas		FA133
Hough, Matthew		ThB54	Jiang, Jiashuo		MC177	Khajavirad, Aida		FB9
Houta, Zakaria		ThB318	Jiang, Jie		TB63	Khalife, Sammy		FB9
Hrga, Timotej		TB66	Jiang, Jun		FA140	Khalil, Elias		WA243
Hsu, Alexander		WA333	Jiang, Li		FB154	Khandelwal, Rohit		TB91
Hu, Hao		FB275	Jiang, Liwei		TA115	Khanduri, Prashant		MC902
Hu, Jie		MB108	Jiang, Nan		ThB124	Khodaparasti, Sara		FA193
Hu, Xiaoyin		ThA135	Jiang, Ruichen	MC55	ThA135	Khorramfar, Rahman		MB254
Hu, Yifan	TB131	WB176	Jiang, Ruiwei	TC96	ThB124	Kiem, Aldo		WC266
Hu, Zhenyu		FA140	Jiang, Shiyi		FA140	Kienzle, Julie		MB190
Huang, Alice		WB185						
Huang, Bo		FA51						

Kiessling, David	FA260	Kuznetsov, Anatoliy	TA324	Legault, Robin	ThA204
Kießling, Miriam	FA224	Kwok, Wing Hong Felix	ThC319	Legavre, Julien	FA955
Kileel, Joe	ThA139	Kwon, Jeongyeol	FA103	Legrain, Antoine	MB219 TC953 WB944
Kilic, Kemal	ThB963	Kwon, Roy	MB960		ThA343 Th3 FB172
Kilinc Karzan, Fatma	WA8	L		Leite, M.	FB154
Kim, Donghwan	TA40			Lerman, Gilad	FA17
Kim, Jongeun	FB1			Le Roux, Nicolas	MB62 TB67
Kim, Junyoung	WA916			Lesage Landry, Antoine	TB322 WA927 WB944
Kim, Kibaek	TA956 ThA343 ThC116		WA231 ThA256		WC171 ThA46
Kim, Sunyoung	WB267 FB36	Labbe, Martin	WB232	Letendre, Nathalie	TA117
Kim, Woojin	WA58 WC203	Lagos, Felipe	WB232	Létocart, Lucas	MB12 WC346
Kimiaei, Morteza	TB322	Lagos, Guido	MB223	Levin, Eitan	TB87
Kirches, Christian	WA914 ThC904	Lagoutte, Aurélie	ThC904	Levin, Michael	MC286 ThB950
Kiviluoma, Juha	WA957	Lahmann, Michel	MB219 WA58	Lewis, Adrian	MB312 TB136 WB331
Klause, Kim	TA276	Lahrichi, Nadia	ThB907 FB154	Ley, Eva	WB939
Kleer, Pieter	ThA937	Lai, Lexiao	TA138 TB136 ThB907 FB154	Leyffer, Sven	T2 TC282 ThA361 FA260
Klein, Nathan	MB239 ThA293	Lai, Zhijian	MB29	Li, Can	MB344 FB151
Klein, Ohad	MC908	Laird, Carl	WC191	Li, Chenxi	TC940
Kleinert, Thomas	WA231 FB1	Lam, Cho Ho (Peter)	ThB963	Li, Guoyin	WA251
Klep, Igor	WA97	Lam, Edward	ThB342	Li, Haitian	TA306
Klibi, Walid	TB166	Lam, Henry	ThB194	Li, Jiajin	TA158 FB257
Knebel, Sebastian	WA97	Lambert, Amélie	FA348	Li, Jingling	MC177
Knudsen, Brage Rugstad	TA956 TB945	Lamothe, François	FA955	Li, Jingyang	FA83
Knueven, Bernard	MC783	Lamperski, Jourdain	TA324	Li, Jonathan	TA158 ThA101 FA247
Kober, Stefan	ThC180 FA228	Lan, Guanghui	TA115 WB176 ThB50 FB114	Li, Mengmeng	TB100 WB125
Koch, Thorsten	MC339	Lan, Qingfeng	ThA246	Li, Michael	TB199 WB147
Kocuk, Burak	FB151	Landay, David	ThB963	Li, Minghua	ThA56
Kocvara, Michal	WA208	Landete, Mercedes	WB917	Li, Qian	WA285
Kocyigit, Cagil	TA134	Lane, Amelia	MB219	Li, Shi	WA931
Koehler, Jana	TA930	Lange, Julia	TA192	Li, Shukai	TA229
Koenemann, Jochen	TA276 WA273	Langlois, Hélène	ThC355	Li, Shuyao	TB69
Koepl, Dominik	ThB352	Lanza, Giacomo	FA193	Li, Su	MB219
Koh, Zhuan Khye	ThB352	Laporte, Gilbert	ThA949 FB164	Li, Tianjiao	TA115
Kojima, Masakazu	WB267 FB36	Lapprès Chartrand, Jean	TA59	Li, Ting	FB52
Kolar, Mladen	FB70	Lara, Cristiana	WA273	Li, Tongyang	MB162
Kolcheva, Laura	TB946	Laradji, Issam	ThA128	Li, Wenjing	TA65
Kollias, Kostas	MB254	Larreal, F.	FB154	Li, Xiang	WA905
Kolmogorov, Vladimir	MB921 ThB910	Larson, Jeffrey	TA306 ThB300	Li, Xiao Li	FB154
Komann, Theodor	ThB318	Larsson, Torbjörn	TB915	Li, Xiaopeng	TA138
Kong, Siyu	TB136	Lasserre, Jean B.	TC23 WC909	Li, Xudong	ThB50
Koning, Nick	WB232	Lassota, Alexandra Anna	FA228	Li, Yan	WC328
Koppel, Alec	MB167	Latafat, Puya	ThB137 FA25	Li, Yingru	ThA246
Korsbo, Niklas	MC283	Latouche, Pierre	ThB918	Li, Zhuorui	WB222
Koster, Arie M.C.A.	MB921	Lau, Hoong Chuin	WC203	Li, Zihao	FA288
Kotary, James	WC942	Laude, Emanuel	ThB226	Li, Zukui	TA938
Kothari, Harshit	WB178	Lauinger, Dirk	ThA204	Liang, Jiaming	TA115
Koukouvinos, Thororis	MB6 FB184	Lavastida, Thomas	FB295	Liang, Jiaqi	FB172
Kouri, Drew	TB91 TC320	Lavington, Jonathan	TB67	Liang, Jingwei	TB136
Krause, Andreas	WB176	Lazaar, Nadjib	FB242	Liang, Ling	TA65 WA208
Kreacic, Eleonora	TA99	Le, Quoc Tung	TB915	Liang, Ya Chun	WA931
Krejic, Natasa	TA59 TB60	Lea, Rehlich	WA957	Liao, Feng Yi	TB136
Krishan Lal, Manish	ThA214 FB275	Lebbe, Nicolas	ThB318	Liao, Hongzhi	WB906
Krislock, Nathan	TB66 ThA214	Lebeuf, Xavier	WA309 ThC261	Liaw, Christopher	WB68
Krklec Jerinkic, Natasa	TB60	Leclère, Vincent	MC934 TA938 TB22 TB952	Liberti, Leo	WA334
Kroner, Frederic	MB932	Le Digabel, Sébastien	TB322 T2 TC305 TC953	Liebchen, Christian	FB268
Krulikowski, Evelin	WC323		WA309 ThC303	Liers, Frauke	TA297 WB939 WB958 WC258
Kseniia, Kurishchenko	MC298	Lee, Jongheon	TA956	Lighthart, Laurens	TC23
Kucukyavuz, Simge	MC57 TA31 WA8 ThC159	Lee, Ching Pei	WA208 FA25	Lim, Andrew	FA247
Kuen, Tobias	FA348	Lee, Dabeen	ThB26	Lima, Leonardo	TB86
Kuhn, Daniel	TA134 TB100 WB176	Lee, Euiwoong	WA931	Lima, Stanley	ThB194
Kukhareenko, Kirill	ThA237	Lee, Eunou	WA97	Lin, Bo	FB104
Kulkarni, Onkar	TA276	Lee, Hyunwoo	FA228	Lin, Meixia	ThB26
Kulkarni, Pooja	TC358	Lee, Jason	MB209	Lin, Qihang	MC44
Kulkarni, Shubhang	FA150	Lee, Jon	WB15	Lin, Tianyi	WB120
Kumar, Akshit	FA288	Lee, Kyungsik	TA956 WA916 ThB950	Lin, Wayne	ThB123
Kungurtsev, Vyacheslav	TA306 WA309 WB329	Lee, Thomas	ThC255	Lin, Wu	TB67
	ThC38	Lee, Yu Ching	FA330	Lin, Xingyu	TB187
Kunstner, Frederik	TB67 WB68	Lefebvre, Henri	TC920	Lin, Ying	MB92 WA285
Kuroiwa, Ryo	WB267 FB242	Lefgoum, Raian Noufel	MB107	Lin, Yueh	ThA256
Kurtz, Jannis	ThA211 FA133	Le Franc, Adrien	WC909	Lin, Zhenwei	TC940
Kurz, Sascha	FA224	Legat, Benoît	MC908	Linderth, Jeff	WA58
		Legat, Benoît	MB344		

346

Mouhoub, Malek	TA922	Nosaka, Keiyu	MB29	Paronuzzi, Paolo	WA916
Moursi, Walaa	TB314 WA338	Nourry, Charles	TB943	Parragh, Sophie	ThB123
Mousseau, Normand	MC337	Nunes Vicente, Luis	FB327	Parrilo, Pablo	TB87
Mueller, Juliane	WA309			Pascal, Timothée	TA297
Mühlenthaler, Moritz	MB223			Passacantando, Mauro	FA193
Müller, Dirk	FA245			Pasupathy, Raghu	ThA325 ThB35
Müller, Peter Marvin	TB91			Pataki, Gabor	FB275
Munari, Pedro	MB157	Ochs, Peter	ThB907	Patel, Krunal Kishor	WB341
Münch, Sabine	FA150	O'donoghue, Brendan	TB207	Patel, Neel	MB219
Münnich, Ralf	ThA924	Oertel, Andy	WC266	Patel, Rahul	WA243
Muñoz, Carlos	WB232	Oggioni, Giorgia	ThA256	Patel, Vivak	ThB35 FB70
Muñoz, Gonzalo	WB232 ThB161 ThC159	Oh, Seyoung	ThB950	Pathapati, Abhishek Pathapati	FA348
Munoz Zuniga, Miguel	MB302	Oikonomidis, Konstantinos	ThB226	Patience, Paul	ThA954
Muramatsu, Masakazu	TB87	Okabe, Kosuke	MC908	Patil, Gandharv	MB62
Murthy, Karthyek	WA106 WC258	Okuno, Takayuki	TB87	Patria, Daniele	ThB34
Musat, Andreea	FA94	Oliveira, Aurelio	MB962	Paules, Granville	FA288
Muszalski, Rafal	TA922	Oliveira, Fabricio	WB939	Pauphilet, Jean	TA31
Mwebembezi, Joachim	MB219	O'Neill, Mike	ThA42	Pauwels, Edouard	TB915
		Ono, Hirotaka	ThB352 FB926	Pavlovic, Viktor	TB314
		Ono, Kenji	ThA277	Pavone, Marco	ThA256
		Onwunta, Akwum	WB321	Payami Shabestari, Bit	TB166
		Orban, Dominique	TP WA333	Pedersen, Jaap	ThB342
		Ordóñez, Fernando	WA231	Pedrielli, Giulia	TC940
		Orfanides, George	TB199	Pedroso, João Pedro	FA224
		Orlitzky, Michael	MB312	Pedroso, Lucas	ThC904
		Ortmann, Janosch	TB166 WB178	Peiró, Juanjo	WB917
		Oskoorouchi, Mohammad	WC81	Pena, Javier	ThB54
		Osorio, Carolina	MB119	Peng, Bo	FB929
		Ostermann, Birte	MB932	Peng, Jiming	WA316
		Ostrovsky, Michael	ThA256	Peng, Renfeng	MB62
		Ota, Matheus Jun	MB157	Peng, Shen	TB946
		Oteiza, Pablo Felipe	WC160	Peng, Zedong	TC282 FB151
		Otto, Alena	TA930	Peper, Ashley	TC953
		Oudaoud, El Mehdi	FB926	Perboli, Guido	FA193
		Ouyang, Wenqing	TB69 TB311 ThA135 FB257	Pérez Aros, Pedro	TB111
		Ouyang, Yuyuan	MB6	Perivier, Noemie	FB295
		Ouzia, Hacène	FP	Perreault, Vincent	WA927
		Oyallon, Edouard	TC911	Pesenti, Silvana	ThA101
		Ozdaglar, Asuman	ThA174	Petridis, Periklis	TB168
		özener, Okan örsan	TB946	Petris, Matteo	TA938
				Peura, Heikki	TA134
				Pfetsch, Marc	TA263 WA957 WB917 ThB161
					ThB118 FA262
				Pham, Quang Hung	ThC261
				Philpott, Andy	TB22 FB109
				Phosavanh, Johnson	TA324
				Piccialli, Veronica	W1
				Pichler, Alois	ThB50
				Pierre Olivier, Pineau	TC142 WA957
				Pina, Jasson	TA117
				Pinçon, Camille	MB219
				Pineda, Salvador	TA117
				Pinheiro, Maria Eduarda	MC298
				Pinto, Renan	FP
				Pinto De Pina Ferreira, Vitor Luiz	TA938
				Pirnia, Mehrdad	WC160
				Plaquet Jourdain, Baptiste	ThA354
				Plate, Christoph	ThA361
				Poggi, Marcus	WA273
				Poirion, Pierre Louis	WA334 WB222
				Pokutta, Sebastian	WA284 WA97 WB939
					ThB161 FA27
				Polik, Imre	WB7
				Pollner, Tristan	FA259
				Pong, Ting Kei	MB92 TB311 FA25 FA17
				Ponte, Gabriel	WB15
				Pool Marquez, Sylvério	ThC303
				Porcelli, Margherita	MB39 MB302
				Poremba, Joseph	FB926
				Porumbel, Daniel	FA348
				Poss, Michael	ThB296

N

Nabli, Adel TC911
Nadarajah, Selvaprabu TA53 WA148
Naderi, Siamak ThB963
Nagarajan, Harsha ThA77 FB1
Nagarajan, Viswanath WB146
Naldi, Simone ThB910
Namchaisiri, Charles MB29
Namkoong, Hongseok MB167 ThB186
Nannicini, Giacomo MB162
Narayanan, Vishnu FA348
Nasini, Stefano ThA256
Nasiry, Javad ThA954
Nasserar, Shahla TB86
Natarajan, Karthik FB122
Nathwani, Jatin WC160
Natura, Bento ThB352
Nava, Rodrigo Garcia TA956
Nawel, Younes TA188
Necoara, Ion ThA42
Nedich, Angelia ThA42
Nemoto, Toshio FA925
Neogi, Rian TC358
Neshat, Najmeh MC337
Neto, José WB236 WC346
Neuwohner, Meike ThA237 FA245
Neville, Jennifer MC177
Newman, Alantha WA931
Ng, Galvin FA247
Nguyen, Alain TA188
Nguyen, Edward Duc Hien TA922 TB943
Nguyen, Hoang Nam WA102 ThB296
Nguyen, Lam WC903
Nguyen, Viet Hung ThA924
Niaparasat, Aidin FB295
Nicolae, Adriana MB312
Nicoletta, Vittorio FB164
Nie, Jiawang TA138 TB87 ThA139
Niemann, Tim WB939 ThA277
Nikbakhsh, Ehsan MB190
Nikola, Jedlicková WA169
Nikos, Zarifis WA196
Nikzad, Erfaneh WA58
Nilforoshan, Hamed FA234
Nishijima, Mitsuhiro FA17
Nishioka, Akatsuki WA333
Nocedal, Jorge WA196
Nomoto, Tazu FA955
Nordström, Jakob WC266 FB919
Norouzi Fard, Ashkan ThA19

O

Ochs, Peter ThB907
O'donoghue, Brendan TB207
Oertel, Andy WC266
Oggioni, Giorgia ThA256
Oh, Seyoung ThB950
Oikonomidis, Konstantinos ThB226
Okabe, Kosuke MC908
Okuno, Takayuki TB87
Oliveira, Aurelio MB962
Oliveira, Fabricio WB939
O'Neill, Mike ThA42
Ono, Hirotaka ThB352 FB926
Ono, Kenji ThA277
Onwunta, Akwum WB321
Orban, Dominique TP WA333
Ordóñez, Fernando WA231
Orfanides, George TB199
Orlitzky, Michael MB312
Ortmann, Janosch TB166 WB178
Oskoorouchi, Mohammad WC81
Osorio, Carolina MB119
Ostermann, Birte MB932
Ostrovsky, Michael ThA256
Ota, Matheus Jun MB157
Oteiza, Pablo Felipe WC160
Otto, Alena TA930
Oudaoud, El Mehdi FB926
Ouyang, Wenqing TB69 TB311 ThA135 FB257
Ouyang, Yuyuan MB6
Ouzia, Hacène FP
Oyallon, Edouard TC911
Ozdaglar, Asuman ThA174
özener, Okan örsan TB946

P

Paat, Joseph ThB161
Pacaud, Francois FA260 FB114
Padmanabhan, Divya WA106 FB122
Pagnoncelli, Bernardo WP FB109
Pai, Mallesh ThA256
Palagi, Laura TA130 FB929
Palitta, Davide MB39
Pallage, Julien WA927
Palmer, David TC299
Pamplona, João Vitor MB960
Pan, Bailin WA285
Pan, Jianting ThB137
Pan, Kai FA140
Panagiotis, Patrinos ThB137 ThB226 FA25
Panciatichi, Patrick WC909 FB929
Pang, Jong Shi TA112 TB195
Papa, Lorenzo TA130
Papadimitriou, Dimitri WC212
Papp, David WC909
Paquet, Marc ThB950
Paquette, Courtney W2 ThA128
Pardalos, Miitiades ThB26
Parekh, Ojas WA97
Parhi, Rahul FB336
Parjadis, Augustin WA243
Park, Hynk TB100 TB22
Parmentier, Axel WB230 ThA937

Potluru, Vamsi	TA99	Rei, Walter	MB190 MC202 TA192 TB166	Sæther, Frida	FB959
Potluru, Vamsi	WB125		WA947 WB178 FA193 F3	Safarina, Sena	MB29
Potschka, Andreas	WA914	Reimann, Adrian Manuel	MC283	Sager, Sebastian	MC283 TC920 ThA361
Poulidis, Stefanos	TB199	Reis, Victor	ThC180	Sahinidis, Nikolaos V.	TA324 FB1
Poursoltani, Mehran	TA170	Rey, David	MC286 ThB950	Sahinyazan, Feyza	FB164
Powell, Warren	FA182	Reza, Babanezhad	TB67	Sainvitu, Caroline	ThC303
Prabhu, Prakash	MB209	Ribeiro, Rafaela	ThB33	Salas, David	ThC159
Prato, Marco	ThB907	Ricard, Léa	TB952 FB172	Saller, Sophia	TA930
Precup, Doina	ThA246	Ricardez-Sandoval, Luis	WB15 FA182	Salomon, Ludovic	TC305 WC323
Prokopyev, Oleg	TA324	Riccietti, Elisa	MB302	Sambharya, Rajiv	TC198
Puerto, Justo	MC286	Richard, Jean Philippe P.	FB1	Samuel, Mccauley	FB295
Puhlmann, Luise	FA245	Richard C., Brewster	WA169	Sanches Portella, Victor	WB68
Pulaj, Jonad	WC266	Richtárik, Peter	TA130	Sandfort, Frederik	WC346
Pulyassary, Haripriya	MB254	Riekert, Adrian	WA941	Sanità, Laura	ThA237
		Rieussec, Jeremy	TA59	Santa, Rosemarie	WA947
		Riley, Benjamin P.	ThA24	Santos, Luiz Rafael	MB92 FB141
		Rinaldi, Francesco	MB12 WA309	Santos Xavier, Alinson	ThB161
		Rios, Ignacio	FB936	Sarabi, Ebrahim	ThA354
		Ripper, Pedro	MC339	Sartenaer, Annick	ThC303
		Ripsman, Danielle	TB168	Sato, Cristiane M.	FB919
		Röber, Tabea	ThA211	Saud, Alghumayjan	ThA204
		Robert, Anna	TA938	Saunders, David	FA247
		Roberts, Lindon	FA326 FB327	Saunders, Michael	FB114
		Robinson, Daniel	TA40 ThB200	Saunderson, James	FA83
		Rocha, Ana	ThA24	Sawkmie, Vanessa Ann Beddoe	WA916
		Rodrigues, Bárbara	TB915	Scaglione, Anna	ThB33
		Rodríguez Acevedo, Iria	WB7	Scavuzzo, Lara	Th1
		Rodríguez Fernández, Pablo	ThA77	Schaefer, Andrew	WB178 ThA256
		Roesel, Florian	WA916	Schaipp, Fabian	WA334
		Roh, Youngjoo	WA916	Schechtman, Sholom	ThC38
		Roland, Marius	ThB935	Scheinberg, Katya	WB41 WC903 ThA325 Panel
		Romeijnders, Ward	WB178		W Panel T
		Romero Morales, Dolores	MB248 MC298	Schell, Kristen	WA957 WC171
		Ronconi, Debora Pretti	FA955	Schettini, Tommaso	MB119
		Rönnberg, Elina	ThA343	Schewe, Lars	WB222 ThB118
		Ropke, Stefan	ThC951	Schewior, Kevin	MB153 WB146
		Roshchina, Vera	MB92 WB906	Schichl, Hermann	WA905
		Rostami, Borzou	MB12	Schiela, Anton	MB88
		Roth, Sarah	ThA343	Schiemann, Annika	TA90
		Rothvoss, Thomas	ThC180	Schiffer, Maximilian	WB147
		Rouhani, Ali	TB166	Schild, Aaron	MB254
		Rousseau, Louis Martin	TB952 WA243 WC241	Schindl, David	WA947 ThA343
			ThB342	Schlomberg, Niklas	FA351 FA245
		Royer, Clement	FB327	Schlöter, Jens	MB153
		Royer, Clément	TA59 TA306 FB327	Schmidt, Johannes	FA224
		Royset, Johannes	TA112 WB176 FA113	Schmidt, Mark	TB67 WB68 ThA128
		R. Penas, David	ThA77	Schmidt, Martin	MC298 TC920 WA231 WB236
		Ruan, Feng	FB75		WB958 ThA924 ThB118 FA133
		Rudi, Alessandro	WB230	Schmidt, Stephan	WA914
		Ruggiero, Valeria	TA59	Schneppe, Felix	ThA912
		Ruiz, Manuel	WC909 FB929	Schröder, Jesper	TC320
		Rung, Thomas	TB91	Schroeder, Louisa	ThB918
		Rush, Cynthia	TA158	Schrotenboer, Albert	MB157
		Ruszczyński, Andrzej	FA182	Schudy, Warren	TB207
		Rutesic, Luka	TB60	Schürks, Jan Malte	ThB244
		Rutkowski, Krzysztof	ThA912	Schürmann, Lukas	WB341
		Rychener, Yves	TA134	Schutte, Emily	FA348
		Ryu, Minseok	TA956	Schütz, Peter	MC202
				Schwartz, Alexandra	ThA256
				Schwiddessen, Jan	ThA214
				Scieur, Damien	FA94
				Scott, Jennifer	WA931
		Sabach, Shoham	WB331	Scutella, Maria Grazia	FA193
		Sabach, Shoham	ThA246	Seattle, Madeleine	TC142
		Sabaei, Soroosh	TC142	Segev, Danny	TA229 FB225
		Saberi, Amin	FA259	Segschneider, Jenny	MB921
		Sadana, Utsav	ThC252	Selvi, Aras	TA99 WA106
		Sadat Asl, Ali Akbar	ThC951	Semenova, Lesia	FA234
		Saddoune, Mohammed	TA274	Semet, Frederic	WB232
		Sadok, Jerad	ThB35 FB70	Sen, Suvrajeet	MC934 MC57
		Sadykov, Ruslan	TA188	Sendov, Hristo	MB312

Q

Qi, Meng	TB73 WB125
Qin, Zhiwei (Tony)	MB167
Qin, Ziqi	TB136
Qiu, Feng	ThB161
Qiu, Junwen	MB71 TB136 ThB137 FB154
Qiu, Yuyang	ThB226
Qu, Siqi	WB227
Qu, Xiaoyi	TA40
Qu, Yushan	WB15
Quanrud, Kent	FA150
Quanz, Brian	TB73
Quesnel, Frédéric	TA274
Quttineh, Nils Hassan	TB915

R

Raach, Stephen	FA150
Racette, Philippe	TA274
Raghunathan, Arvind	WC258 FB1
Rahimi, Fahimeh	TB168
Rahimian, Hamed	FB109
Rajabalizadeh, Atefeh	WC241
Rakshith, Sai	WC258
Ralphs, Ted	WA231
Ramachandra, Arjun	FB122
Ramachandran, Akshay	FA94
Rambau, Jörg	FA224
Rame, Martin	WB147
Ramirez, Hector	WB329
Ramirez Ayerbe, Jasone	MB248
Ran, Lun	FB141
Rancourt, Marie ève	WA947 FB164
Rangel Martinez, Daniel	FA182
Ranjan, Vinit	TC198
Rao, Sujit	WA97
Rasouli, Nadia	TA274
Rastgar Amini, Farin	FA955
Rathi, Tushar	ThA24
Rauhut, Holger	WB68
Ravi, R	WA273
Raydan, Marcos	WC323
Raymond, Annie	ThA214
Raymond, Jean Florent	MB223
Raymond, Vincent	ThB918
Razaviyayn, Meisam	FA113
R. Banga, Julio	ThA77
Rebegoldi, Simone	ThB907
Rebennack, Steffen	MC14 TB5
Rebjock, Quentin	ThA30 ThB93
Redlich, Amanda	ThA214
Regan, Amelia	ThB342
Rehfeldt, Daniel	WB267

S

Sabach, Shoham	WB331
Sabach, Shoham	ThA246
Sabaei, Soroosh	TC142
Saberi, Amin	FA259
Sadana, Utsav	ThC252
Sadat Asl, Ali Akbar	ThC951
Saddoune, Mohammed	TA274
Sadok, Jerad	ThB35 FB70
Sadykov, Ruslan	TA188

Sæther, Frida	FB959
Safarina, Sena	MB29
Sager, Sebastian	MC283 TC920 ThA361
Sahinidis, Nikolaos V.	TA324 FB1
Sahinyazan, Feyza	FB164
Sainvitu, Caroline	ThC303
Salas, David	ThC159
Saller, Sophia	TA930
Salomon, Ludovic	TC305 WC323
Sambharya, Rajiv	TC198
Samuel, Mccauley	FB295
Sanches Portella, Victor	WB68
Sandfort, Frederik	WC346
Sanità, Laura	ThA237
Santa, Rosemarie	WA947
Santos, Luiz Rafael	MB92 FB141
Santos Xavier, Alinson	ThB161
Sarabi, Ebrahim	ThA354
Sartenaer, Annick	ThC303
Sato, Cristiane M.	FB919
Saud, Alghumayjan	ThA204
Saunders, David	FA247
Saunders, Michael	FB114
Saunderson, James	FA83
Sawkmie, Vanessa Ann Beddoe	WA916
Scaglione, Anna	ThB33
Scavuzzo, Lara	Th1
Schaefer, Andrew	WB178 ThA256
Schaipp, Fabian	WA334
Schechtman, Sholom	ThC38
Scheinberg, Katya	WB41 WC903 ThA325 Panel
	W Panel T
Schell, Kristen	WA957 WC171
Schettini, Tommaso	MB119
Schewe, Lars	WB222 ThB118
Schewior, Kevin	MB153 WB146
Schichl, Hermann	WA905
Schiela, Anton	MB88
Schiemann, Annika	TA90
Schiffer, Maximilian	WB147
Schild, Aaron	MB254
Schindl, David	WA947 ThA343
Schlomberg, Niklas	FA351 FA245
Schlöter, Jens	MB153
Schmidt, Johannes	FA224
Schmidt, Mark	TB67 WB68 ThA128
Schmidt, Martin	MC298 TC920 WA231 WB236
	WB958 ThA924 ThB118 FA133
Schmidt, Stephan	WA914
Schneppe, Felix	ThA912
Schröder, Jesper	TC320
Schroeder, Louisa	ThB918
Schrotenboer, Albert	MB157
Schudy, Warren	TB207
Schürks, Jan Malte	ThB244
Schürmann, Lukas	WB341
Schutte, Emily	FA348
Schütz, Peter	MC202
Schwartz, Alexandra	ThA256
Schwiddessen, Jan	ThA214
Scieur, Damien	FA94
Scott, Jennifer	WA931
Scutella, Maria Grazia	FA193
Seattle, Madeleine	TC142
Segev, Danny	TA229 FB225
Segschneider, Jenny	MB921
Selvi, Aras	TA99 WA106
Semenova, Lesia	FA234
Semet, Frederic	WB232
Sen, Suvrajeet	MC934 MC57
Sendov, Hristo	MB312

Sepho, Loreley	WC171	Sinha, Deeksha	ThA277	Szechtman, Roberto	TA170
Seraji, Reza Roohi	ThC38	Sinha, Saumya	ThA237		
Seri, Matan	TA170	Sinn, Rainer	TB195		
Seroussi, Inbar	TB63	Sinnl, Markus	ThB123		
Seto, Akane	ThB155	Sinoquet, Delphine	MB302		
Severitt, Marvin	TA90	Slot, Lucas	ThB37		
Sevim, Ismail	MB119	Smith, Cole	WA58		
Seweryn, Michal	ThC180 FA228	Smith, Michael	TB66		
Shafiee, Soroosh	TA158	So, Man Cho	TA158 WB350 FB257		
Shafieezadeh Abadeh, Soroosh	WA8	Soares Goncalves, Douglas	ThA30		
Shah, Suhail	MB11	Sohab, Oumaima	FB327		
Shalaby, Amer	TB945	Soheili, Negar	WA148		
Shames, Iman	ThA46	Solà Vilalta, Albert	WC171		
Shanbhag, Uday	WA284 WB227 ThB226	Soma, Tasuku	MB921		
Shao, Zhen	FB70	Song, Yongjia	TC96 FB109		
Shapiro, Alexander	WB176	Song, Yongzhong	FB52		
Shapley, Richard	ThA293	Sophia, Calderón	WB232		
Sharan, Vaswani	MB62 TB67	Sorgento, Carmine	FA925		
Sharbaf, Maedeh	FB164	Sossa, David	WB906		
Sharma, Kartikey	WB939 FA27	Soumis, Francois	TA274 WA281		
Sharrock, Louis	ThC129	Soytürk, Muhammet	MB287		
Shashaani, Sara	ThA325	Spieksma, Frits	FA262		
Shea, Betty	WB68	Spitzley, Yannik K.D.	ThB244		
Shehadeh, Karmel	WB82	Spracklen, Patrick	FB242		
Shen, Haoming	TA170 ThB124	Spyridon, Pougakiotis	WB906 ThB935		
Shen, Houcai	WB939	Sra, Suvrit	W2		
Shen, Jinglai	ThA56	Srinivas, Vaidehi	FB75		
Shen, Yifan	MB254	Srinivasan, Aravind	FB225		
Shen, Zuo Jun	TB73	Sritharan, R	WB185		
Shen, Zuo Jun Max	WB125 Th3 FA140	Stadlerova, Sarka	MC202		
Shepherd, Bruce	FB926	Starreveld, Justin	TB168		
Sherkat Masoumi, Zoha	FB109	Staudigl, Mathias	WB227 ThB226		
Shi, Chenbo	WA243	Steffen, Eckhard	TA263		
Shi, Jiahao	WB41 ThB300	Steffy, Dan	MB344		
Shi, Qingjiang	FA173	Stein, Clifford	FB295		
Shimizu, Shoji	WA961	Steinhardt, Marcel Maximilian	WB321		
Shin, Sungho	FA260 FB114	Stella, Lorenzo	FA25		
Shinano, Yuji	WB267	Stellato, Bartolomeo	TC198 WC203		
Shioura, Akiyoshi	FA925	Stephen, Tamon	TC359		
Shirai, Yuki	WC258	Stiller, Sebastian	MB932 WB939 ThA277		
Shmoys, David	MB254 ThA293 FB225	Stingl, Michael	WC258		
Shojaie, Ali	MC57	Strahl, William	FB1		
Shroff, Ravi	FA234	Street, Alexandre	ThB33 FA259		
Shtern, Shimrit	WB125 ThB296	Stuber, Matthew	TA324		
Shu, Kevin	WB18 FA89	Sturt, Bradley	TC96		
Si, Fan	MB108	Subramanyam, Anirudh	WC258 FB936		
Si, Nian	WA106	Sudoso, Antonio M.	WA281		
Siddharth, Gollapudi	FB225	Sugishita, Nagisa	MB119		
Sidford, Aaron	MC163	Sugrue, Ian	TB66		
Siebenhofer, Melanie	TB66	Sujanani, Arnesh	WB222		
Sikora, Jamie	MC163 TC282	Sumida, Mika	TA229		
Silina, Olha	MC286	Sun, Andy	TA152 ThA204 ThC255		
Silva, Everton	TC305 WC323	Sun, Chunlin	ThB194		
Silva, Gilson N.	WA905	Sun, Defeng	WA285		
Silva, Paulo J. S.	MB39 MB92 WA338	Sun, Hailin	TB63		
Silveti Falls, Antonio	ThC38	Sun, Hailong	FB122		
Silvia, Bonettini	ThB907	Sun, Huai Jun (Norina)	ThB155		
Sim, Chee Khian	ThB910	Sun, Jennifer	MB254		
Sim, Melvyn	MB108 TB73 ThB194 FA140	Sun, Jianyong	FA173		
Simai, He	FB122	Sun, Ju	FA94		
Simon, Nicolai	TB91	Sun, Shengding	WA251		
Simoncini, Valeria	MB39	Sun, Shuo	WB125		
Simons, Stephen	TB311	Sun, Yifan	WB146		
Simpson Porco, John W.	ThA46	Sun, Zhenyu	MB62		
Sinclair, Sean	MC177	Suonperä, Ensio	MB88		
Singer, Amit	TC299	Suriyanarayana, Varun	FB225		
Singh, Gaurav	ThB963	Sutter, Tobias	TB100		
Singh, Shambhavi	TA313	Svensson, Anton	ThC159		
Singh, Shikha	FB295	Swaminathan, Adith	MC177		
Singh, Vikras Vikram	WA102	Swamy, Chaitanya	MB153 TC358		
Singla, Sahil	MB153	Syga, Monika	TC911		

T

Tadaros, Marduch	TB915
Taeb, Armeen	MC57
Tahmasbi, Nadia	TB166
Takahashi, Shota	FA83
Takeda, Akiko	WA334 WB222 ThA277 ThB137 FA83 FA103
Tanaka, Mirai	ThA277
Tanaka, Takashi	TB100
Tang, Tianyun	TB207 WA208
Tang, Xindong	TA138 ThA139
Tanji, Sofiane	TB340
Tanner, Jared	ThC928
Tansley, Edward	FB70
Tao, Min	WA251
Tappenden, Rachael	MC902
Tardella, Fabio	WA905
Tarnawski, Jakub	ThA19
Taskesen, Bahar	TA134
Tavasoli, Ahmadreza	MB962
Tawarmalani, Mohit	FB1
Teboulle, Marc	WB331
Temprano, Francisco	MC286
Teo, Chung Piau	FB122
Terlaky, Tamás	MP MB162 WB317
Tesch, Alexander	FB242
Tetraushvili, Lubov	WB125
Themelis, Andreas	ThB137 FA25
Theobald, Thorsten	ThB123
Thiele, Hanjo	FA351
Thomä, Simon	WB147
Thürauf, Johannes	WA231 WB236 ThB118 FA133
Tian, Lai	FB257
Tillmann, Andreas	MB932 FA933
Tjandraatmadja, Christian	ThA181
Toh, Kim Chuan	MB71 TA65 TB207 WA208 ThA135 FA83
Toint, Philippe	ThB35 FB70
Tomás, Feder	WA169
Tomasgard, Asgeir	TA956
Tong, Shanyin	WC258
Toriello, Alejandro	ThB950 FA89 FA351
Torrico, Alfredo	ThB124 FB936
Toussaint, Hélène	ThA19
Tran, Hien	WC266
Tran, Huyen Trang	WC903
Tran, Nghia	FA113
Tran Dinh, Quoc	TA40
Traub, Vera	ThA293 FA245
Traversi, Emiliano	MB12 WC346
Trepanier, Martin	MC337
Tribes, Christophe	TC305 WC323
Trivella, Alessio	WA148
Tröbst, Thorben	FA245
Trombettoni, Gilles	FA27
Trombini, Ilaria	TA59
Tsaknakis, Ioannis	MC902
Tsang, Man Yiu	WB82
Tsay, Calvin	WC346 WC191
Tsigaridas, Elias	ThB123
Tsuchiya, Takashi	TB87
Tulabandhula, Theja	WA941
Tuncel, Levent	WB317 FA17 FB919
Turner, Mark	MC339
Turner, Richard E.	TB67
Turner, Rowan	ThB118

U

Udell, Madeleine FB141
 Udwani, Rajan TA229
 Uetz, Marc MB153
 Ulbrich, Michael WA334
 Ulbrich, Stefan TC320 WA957 WB321 WB958
 ThB318 ThB118 FB52

Ulch, Daniel WA273
 Umetani, Shunji FA933
 Uno, Takeaki FB926
 Uribe, Cesar A TA922 TB943
 Uru, Cagin FA259
 Uyama, Kazuya ThB155

V

Vaiter, Samuel TB915
 Valencia, Juan WB232
 Valimoradi, Reza ThA101
 Valkonen, Tuomo MB88
 Valladão, Davi ThB33 FA259 FB109
 Vanaret, Charlie FA260
 Van Beesten, Ruben WB232
 Vanberkel, Peter MB219
 Van Dellen, Austin TA297
 Van Dyk, Madison TA276 WA273
 Van Heeswijk, Wouter WA148
 Van Hentenryck, Pascal TA152 T1
 Van Hoeve, Willem Jan WC241 FA933
 Van Leeuwaarden, Johan FA133
 Van Lent, Freija WA923
 Van Osselaer, Willem ThB963
 Van Roy, Benjamin W3
 Van Roy, Wim TC920
 Van Ryzin, Garrett TB67
 Van Woensel, Tom TA192
 Vargas Koch, Laura MB239
 Varol, Taha TB946
 Varun, Sivashankar FB225
 Varvitsiotis, Antonios ThB123
 Vasconcelos, Francisca ThB123
 Vaseileios, Koutsoukis WC212
 Vassilvitskii, Sergei FB295
 Vavasis, Stephen TB314 ThB54 FA25
 Vayanos, Phebe ThA211 FB141
 Velasco, Mauricio TB195
 Veloso, Manuela TA99
 Vera Lizcano, Juan WB18
 Verschae, José FA262
 Vértesi, Tamás WA97
 Vetta, Adrian FA262
 Vidal, Thibaut TA297 WA243 ThA937 ThB935
 ThC928 FB926

Vidarthi, Navneet WC160
 Vieira, Bruno WC191
 Vielma, Juan Pablo MB344 ThA181 FB165
 Villagra, Matias TA152
 Vinagre, Cybele TB86
 Vitercik, Ellen WB230
 Vlatakis Gkaragkounis, Emmanouil
 Vaseileios ThB123
 Vogl, Lukas WA931
 Volkwein, Stefan ThB34
 Vossler, Patrick ThA211
 Vuchkov, Radoslav ThC319
 Vuong, Thuy Duong ThA19
 Vygen, Jens ThA293 FA351 FA245

W

Wachsmuth, Daniel MB88
 Waddell, Lucas MB12
 Waechter, Andreas WB41
 Waga, Mateus FA259 FB109
 Wajc, David FA259
 Wallace, Stein W. TA192
 Walter, Daniel TB111
 Walter, Matthias FA348
 Walteros, Jose TA922
 Walwil, Iyad ThC904
 Walz, David WC346
 Wang, Akang FA173
 Wang, Alex MB6 WB18 FB75
 Wang, Chun TB945
 Wang, Chun Han FA330
 Wang, Guanyi ThB200
 Wang, Hao MB162 TA65 TB311
 Wang, Irina FB104
 Wang, Jiayi TB187 ThA325
 Wang, Jie WA106 WB176
 Wang, Jimmy MB167 ThB186
 Wang, Kai WB147
 Wang, Liping WB49
 Wang, Mengdi ThA246
 Wang, Puqian WA196
 Wang, Qi MB11
 Wang, Qingsong ThB907
 Wang, Shengbo WA106
 Wang, Shuming MB108
 Wang, Tianxiao WA106 WB939
 Wang, Tianyu ThB137 ThB194 FA326
 Wang, Xiaoyu ThA128
 Wang, Xiaozhou FA17
 Wang, Xin ThB155
 Wang, Yicheng WC171
 Wang, Yifan MB153
 Wang, Yijia TA53
 Wang, Yuhao TB100
 Wang, Zhengchao TA134
 Wang, Zhiyuan FB141
 Wang, Zhongxiang FA288
 Wang, Zicheng FA326
 Wang, Zitong MC783
 Warme, David WB917
 Warnow, Leo ThB34
 Warwicker, John MC14
 Wasilkoff, Alex ThC255
 Watson, Jean Paul MC783
 Weames, Tyler ThB910
 Webber, Robert FB70
 Wegener, Raimund WB321
 Wehbeh, Jad WA914
 Wei, Ermin MB62 WB41
 Wei, Hao Ting FB295
 Wei, Ningji WB82
 Wei, Zhou MC14
 Weider, Steffen FB268
 Weldeyesus, Alemseged TB292
 Welker, Kathrin ThB318
 Weller, Paula WB939
 Weltge, Stefan ThC180 FA228 FA150
 Welzel, Karl ThB35
 Wen, Zaiwen WA208 FA173
 Weninger, Noah WB236
 Weo, Linchuan WA8
 Wessel, Sten FA262
 Weston, Jorge WB958

Wiebe, Amy TC359 ThB300
 Wiecek, Margaret ThB34
 Wiedmer, Manuel ThB37
 Wiegele, Angelika TB66
 Wiertz, Ann Kathrin ThA256
 Wiese, Sven TB340
 Wiesemann, Wolfram TA99 TA134 TB100
 Wild, Stefan M. ThA325 FB327
 Willemssen, Rick TC289
 Williamson, David ThA293
 Winkler, Michael MC339 TA269
 Wirth, Elias ThB54
 Wodecki, Ales WB329
 Woerner, Stefan MB162
 Wolf, Isaak H. TA263
 Wolf, Nathalia TB292
 Wolfler Calvo, Roberto ThA937
 Wolkowicz, Henry FB275
 Wollner, Winnifried TB91
 Wood, Killian ThA325
 Woodruff, David MC783 ThB186
 Woodstock, Zev WA284
 Wright, Stephen TA40 TB69
 Wu, Haoze WC266
 Wu, Jian ThA101
 Wu, Jiayuan TB314
 Wu, Manxi MB254
 Wu, Qinyu FA247
 Wu, Tong ThB33
 Wunderling, Roland WA927
 Wyschka, Henrik TB91

X

Xia, Yong ThB907 FB36
 Xiao, Nachuan MB71 ThA135
 Xiao, Zhuoyu WB227
 Xiaobo, Li TA170 FB122
 Xie, Jiaxin TA130
 Xie, Miaolan WA196
 Xie, Pengcheng ThB300
 Xie, Weijun TA31 TC96 ThB124 ThB296
 Xie, Yaqi MC177
 Xin, Linwei MC177
 Xiong, Peng MB108
 Xiong, Zikai TB207 WC45
 Xu, Bolun ThA204
 Xu, Chao ThA924
 Xu, Huifu ThA101
 Xu, Jingye WA8
 Xu, Lingling ThA912
 Xu, Luze TB195 WB15
 Xu, Pan ThA246
 Xu, Tong MC57
 Xu, Yangyang WA284
 Xu, Zi TB314
 Xue, Chenyu ThB194

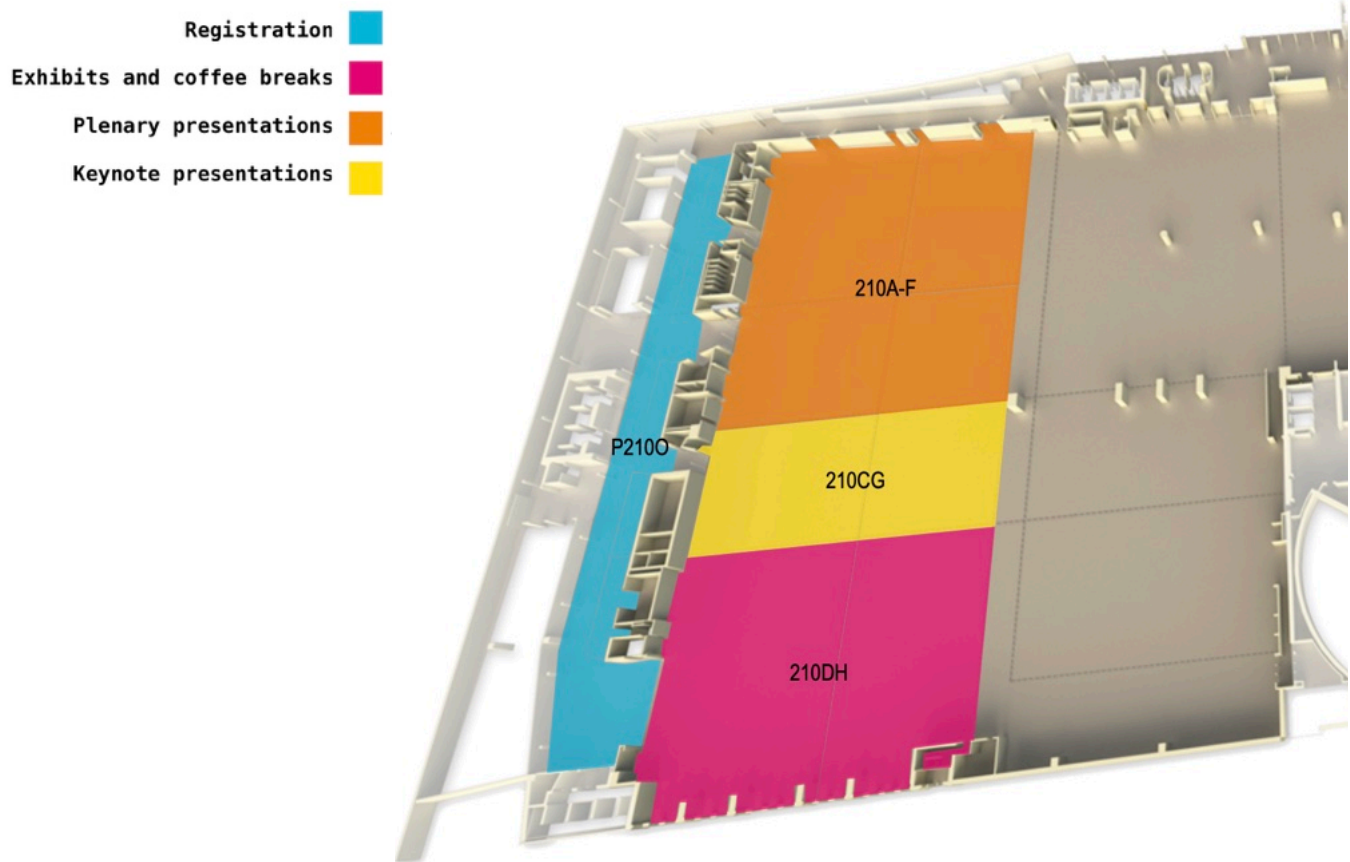
Y

Yaakoubi, Yassine WA927
 Yadegari, Ehsan TC920
 Yamada, Syuui ThB26
 Yaman, Hande WB917
 Yamashita, Makoto MB29 FB36
 Yan, Chiwei MB254
 Yan, Julia MB254
 Yan, Junchi FA173
 Yan, Ming TA130 ThB137

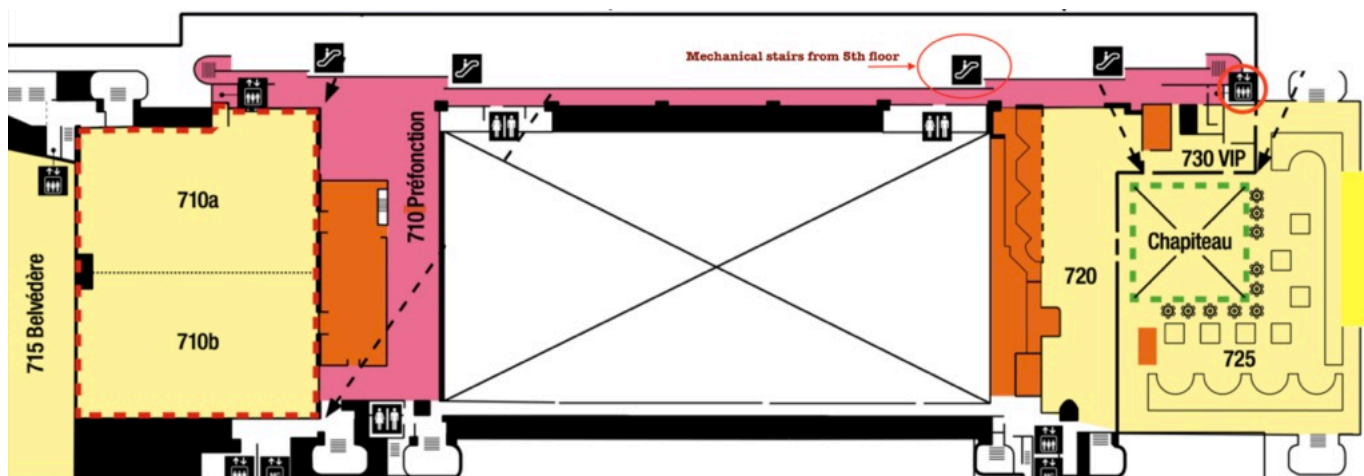
Yang, Boshi	MB6	Yu, Jiajia	FA326	Zhang, Wenjie	FA94
Yang, Dominic	ThC270	Yu, Kim	ThC159	Zhang, Wenxing	FA51
Yang, Haeseong	TA324	Yu, Shijie	ThA56	Zhang, Xun	TB73
Yang, Hanbin	ThC116	Yu, Xian	WB82 ThA174	Zhang, Yi	MB162
Yang, Haoxiang	TB187 ThC116	Yu, Yaoliang	FA25	Zhang, Yihang	MC934 MC57
Yang, Jincheng	TB187	Yuan, Chenyang	FA89	Zhang, Yiling	WB82
Yang, Jinwen	ThA74	Yuan, Yancheng	WA285	Zhang, Yulan	FA27
Yang, Junfeng	FB75	Yuan, Ya Xiang	MB62 ThB300	Zhang, Zaikun	TA306 TB322 FB327
Yang, Lei	TA65	Yuditsky, Yelena	ThC180 FA228	Zhang, Zhi Hai	ThB296
Yang, Lixing	FB172	Yue, Man Chung	TA99	Zhang, Zhuo	TB187
Yang, Shangzhe	FA182	Yuni, Lai	WA285	Zhang, Zijin	MC177
Yang, Shuoguang	ThB50			Zhao, Jinye	WA206
Yang, Tianbao	MC44			Zhao, Long	TB73
Yang, Wenguo	FA173			Zhao, Renbo	ThB54
Yang, Xiaoqi	ThA56			Zhao, Yonggan	MB107
Yang, Zhen	MB108			Zhao, Yue	ThB194
Yang, Zi	ThA139			Zhao, Yue	FA140
Yanikoglu, Ihsan	ThA174			Zhen, Jianzhe	MB6 FB184
Yao, Wei	ThB913			Zheng, Alice	MC163
Yao, Wenfang	ThA56			Zheng, Taoli	TA158
Yao, Xie	WA106 WB176			Zheng, Tongxin	WA206
Yao, Yao	MC44			Zheng, Yang	TB136
Yarkony, Julian	ThB342			Zhong, Suhan	TA138 TB87
Yasuda, Taisuke	MB209			Zhou, Angela	TA229
Yazdandoost Hamedani, Erfan	MC55			Zhou, Baoyu	WB41
Ye, Jane	ThB913			Zhou, Daniel	WC266
Ye, Ke	TA138			Zhou, Danqing	FB75
Ye, Yinyu	TC940 WA106 ThA135 ThB194			Zhou, Dongruo	ThA246
Ye, Zeqi	WA941			Zhou, Duo	TB100
Yemshanov, Denys	FB164			Zhou, Enlu	TB100
Yetis Kara, Bahar	ThA949			Zhou, Minglong	TB73 FB141
Yi, Ming	ThA204			Zhou, Yi	TB131
Yi, Tianyang	FB936			Zhou, Yuan	WB15
Yifan, Jiang	ThB137			Zhou, Zhengyuan	MC177 WA106 FB154
Yildirim, Sinan	WC942			Zhou, Zirui	ThB963
Yin, Jiateng	FB172			Zhu, Ian Yihang	MB108
Yin, Ming	ThA246			Zhu, Jiamin	TA112
Ying, Bicheng	TA922 TB943			Zhu, Kate Wenqi	TB69
Ying, Zhiliang	ThB137			Zhu, Linglingzhi	FB257
Yokoi, Yu	WA923			Zhu, Quanyan	ThC252
Yorke Smith, Neil	WC203 Th1			Zhu, Zheqing	MB167
Yoshida, Takufumi	ThA277			Zhu, Zoey	WB82
Yoshise, Akiko	MB29			Zinflou, Arnaud	ThC261
Yousefian, Farzad	WA284 ThB226			Zorgati, Riadh	TB292
Yu, Annan	WB185			Zöttl, Gregor	ThA256
Yu, Chang	WA338			Zoumpoulis, Spyros	TB199
Yu, Di	ThB35			Zufferey, Nicolas	WC160
Yu, Dongmei	FA51			Zuluaga, Luis	MB162 WA316

Floor plan and map to the banquet

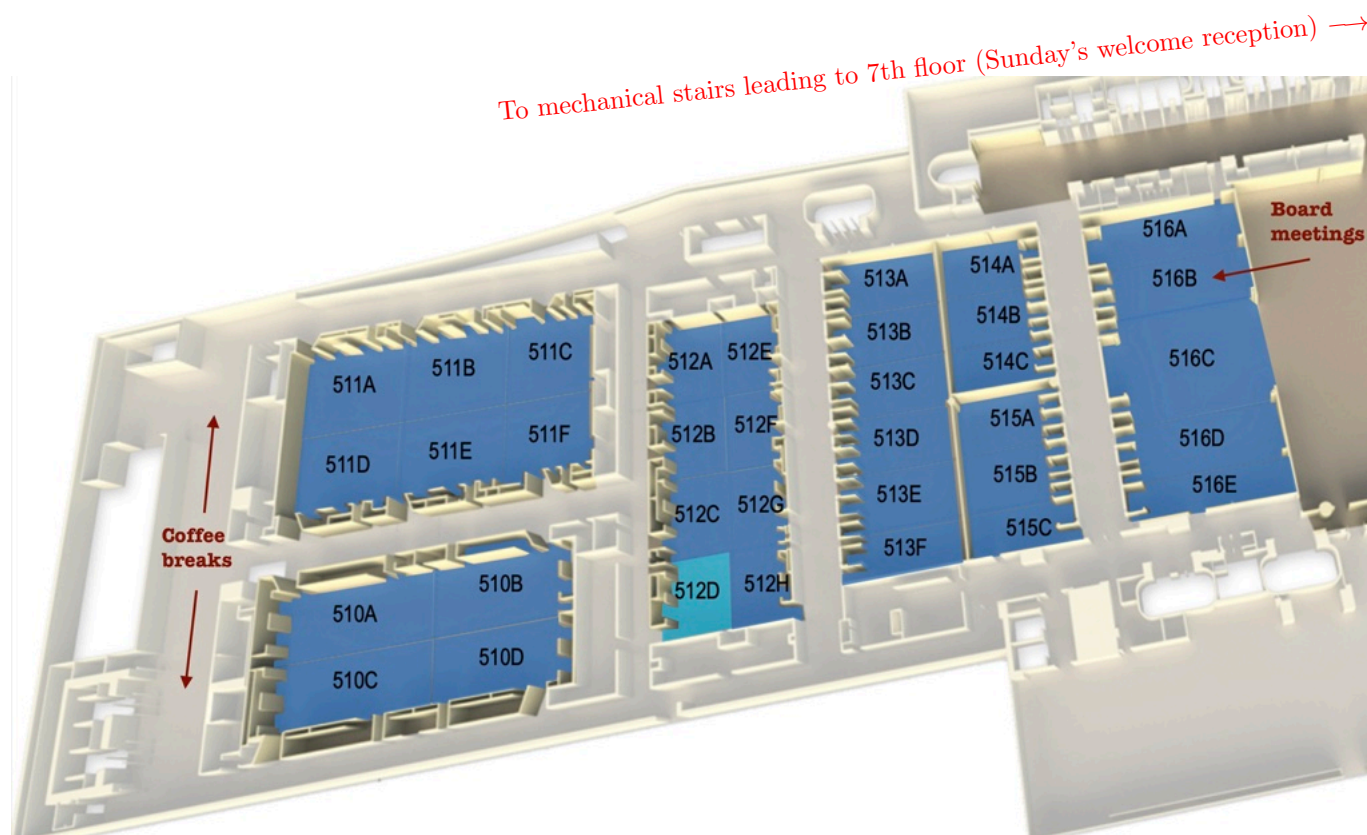
Second floor Registration, plenary and keynote sessions, exhibits and AM & PM coffee breaks.
Above the Jean-Paul Riopelle entrance.



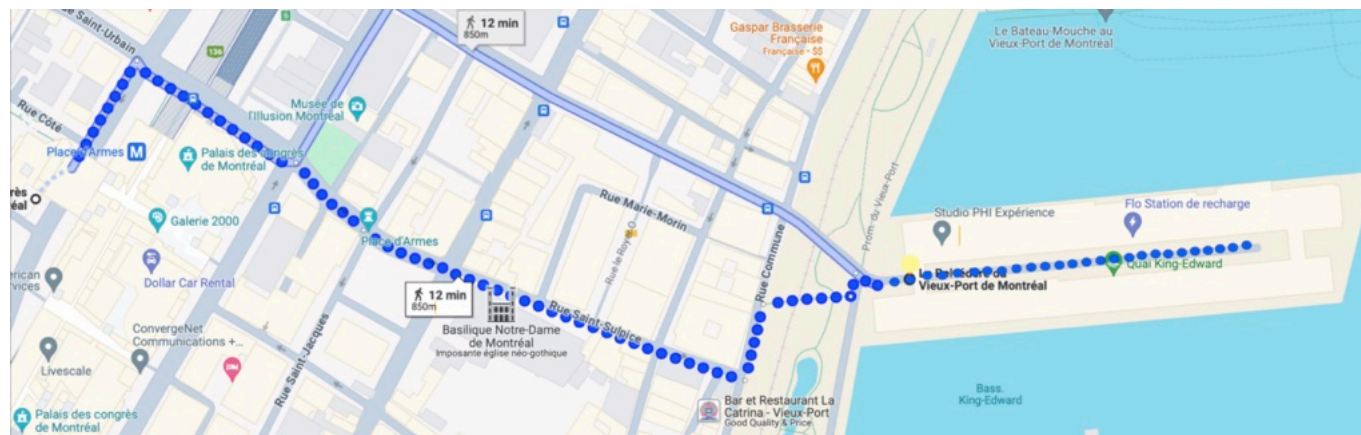
Seventh floor Sunday's cocktail in Le Chapiteau (room 720).



Fifth floor Parallel sessions, board meetings, panels and afternoon coffee breaks.



Thursday's banquet at the Belvédère in the Old Port 15-minute walk.
2 Rue de la Commune Ouest, Montréal, QC H2Y 4B2



Remember to bring your badge !