# Column Generation with the Primal-Dual Interior Point Method

**Jacek Gondzio**

joint work with

*Pablo González-Brevis and Pedro Munari*

# Outline

- IPMs for Optimization
  → central path, perturbed complementarity
- Column Generation/Cutting Planes *with* IPM
- Warmstarting IPMs
  → theory and practice
- Applications (numerous examples)
  → summary results
- Conclusions

## What is the best method to solve LPs/QPs:

- the Simplex Method (active-set method) or
- the Interior Point Method ?

## Maths question: How to cross a polytope?

... let us ask the **expert**

# Brazil view of LP/QP:


# How to solve LP/QP problems?

If we asked **Neymar Jr**, the likely answer would be:
**"go through the interior of the polytope"**.

# Interior Point Methods

# Primal-Dual Pair of Quadratic Programs

Primal | Dual

$$\begin{array}{rrcl} \min & c^T x + \frac{1}{2} x^T Q\, x & & \\ \text{s.t.} & Ax & = & b, \\ & x \geq 0; & & \end{array}$$

$$\begin{array}{rrcl} \max & b^T y - \frac{1}{2} x^T Q\, x & & \\ \text{s.t.} & A^T y + s & = & c, \\ & s \geq 0. & & \end{array}$$

# Lagrangian

$$L(x,y) = c^T x + \frac{1}{2} x^T Q\, x - y^T (Ax - b) - s^T x.$$

# Optimality Conditions

$$\begin{array}{rcl} Ax & = & b, \\ A^T y + s - Qx & = & c, \\ XSe & = & 0, \quad (\text{ i.e., } \; x_j \cdot s_j = 0 \quad \forall j), \\ (x, s) & \geq & 0, \end{array}$$

$X = diag\{x_1, \cdots, x_n\}$, $S = diag\{s_1, \cdots, s_n\}$, $e = (1, \cdots, 1) \in \mathcal{R}^n$.

# First Order Opt Conditions for QP

$$\begin{aligned} Ax &= b, \\ A^T y + s - Qx &= c, \\ XSe &= \mathbf{0}, \\ (x, s) &\geq 0, \end{aligned}$$

# First Order Opt Conditions for Barrier QP

$$\begin{aligned} Ax &= b, \\ A^T y + s - Qx &= c, \\ XSe &= \boldsymbol{\mu e}, \\ (x, s) &> 0, \end{aligned}$$

# Complementarity in the Interior Point Method

**The first order optimality conditions** (FOC)

$$\begin{aligned} Ax &= b, \\ A^T y + s - Qx &= c, \\ XSe &= \boldsymbol{\mu e}, \\ x, s &\geq 0, \end{aligned}$$

where $X = diag\{x_j\}, S = diag\{s_j\}$ and $e = (1, \cdots, 1) \in \mathcal{R}^n$.

**Analytic centre ($\mu$-centre):** a (unique) point $(x(\mu), y(\mu), s(\mu)),\ x(\mu) > 0,\ s(\mu) > 0$ that satisfies FOC.

The interior point method gradually reduces the complementarity products

$$x_j \cdot s_j \approx \mu \to 0 \quad \forall j = 1, 2, ..., n.$$

# Attractive features

## IPMs:

- excell on large-scale problems
- can easily control the distance to optimality
- are indifferent to degeneracy
- are able to warm start

**JG**, Interior Point Methods 25 Years Later,
*EJOR*, 218 (2012) 587–601.

## Decomposition Approaches

# What IPMs can offer?

- Use IPM to solve the Master
- Use IPM to solve the Subproblem
  early termination with $\varepsilon$-optimality
  $\rightarrow$ **$\varepsilon$-subgradients** ("on-demand accuracy")

**JG and Vial**, Warm start and $\varepsilon$-subgradients in cutting plane scheme for block-angular linear programs, *Comput Optimization and Applications* 14(1999) 17-36.

**JG and Kouwenberg**, High performance computing for ALM, *Operations Research* 49 (2001) 879–891.

# The First Order Optimality Conditions

$$
\begin{aligned}
Ax &= b, \\
-Qx + A^T y + s &= c, \\
XSe &= \boldsymbol{\mu} e, \\
(x, s) &> 0.
\end{aligned}
$$

Assume primal-dual feasibility:

$$Ax = b \qquad \text{and} \qquad -Qx + A^T y + s = c$$

Parameter $\boldsymbol{\mu}$ controls the distance to optimality.

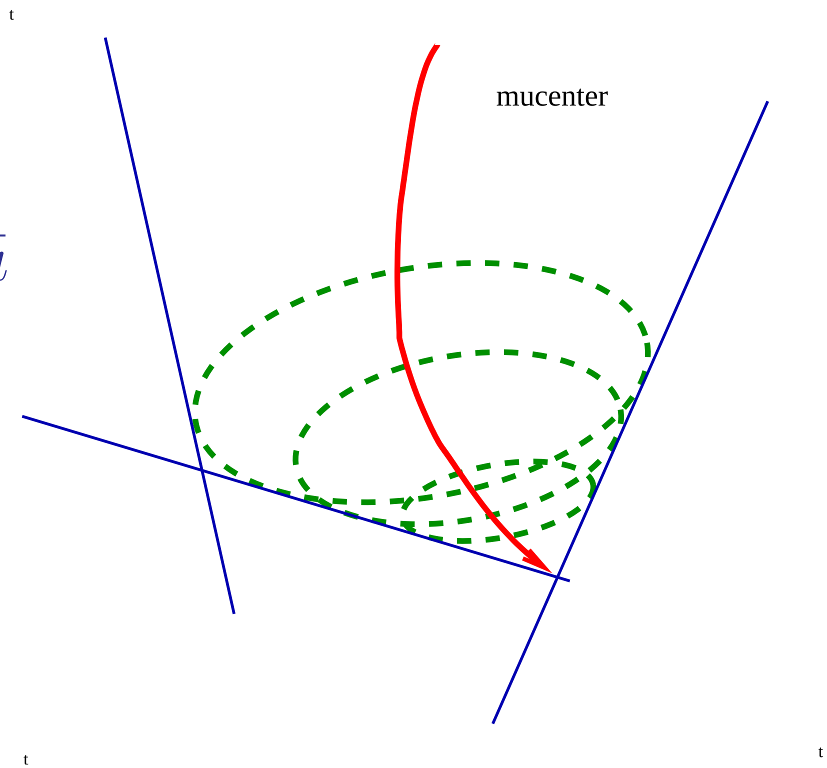
$$(c^T x + \frac{1}{2} x^T Q x) - (b^T y - \frac{1}{2} x^T Q x) = x^T s = n\boldsymbol{\mu}.$$

**On-demand accuracy** is readily available.

# **Primal-Dual Column Generation Method**:

Heading-in problems:
terminate RMP solution early:
→ **get stable dual solution** $\bar{u}$

Tailing-off problems:
use IPM to solve the RMP:
→ **no degeneracy issues**



# Primal-Dual
# Column Generation Method

# Column Generation (CG)

Consider an LP, called the master problem (MP):

$$
\begin{aligned}
z^\star := \min \quad & \sum_{j \in N} c_j \lambda_j, \\
\text{s.t.} \quad & \sum_{j \in N} a_j \lambda_j = b, \\
& \lambda_j \geq 0, \qquad \forall j \in N.
\end{aligned}
$$

- $N$ is too big;
- The columns $a_j$ are implicit elements of $\mathcal{A}$;
- We know how to generate them!

## CG: Restricted master problem (RMP): $\overline{N} \subset N$

$$
\begin{aligned}
z_{RMP} := \min \quad & \sum_{j \in \overline{N}} c_j \lambda_j, \\
\text{s.t.} \quad & \sum_{j \in \overline{N}} a_j \lambda_j = b, \\
& \lambda_j \geq 0, \qquad \forall j \in \overline{N}.
\end{aligned}
$$

- Optimal $\bar{\lambda}$ for the RMP $\Rightarrow$ feasible $\hat{\lambda}$ for the MP;
- $\hat{\lambda}_j = \bar{\lambda}_j$, $\forall j \in \overline{N}$, and $\hat{\lambda}_j = 0$ otherwise;
- Hence, $z^\star \leq z_{RMP} = UB$ (Upper Bound).
- How to know it is optimal?
  - Call the oracle!

# CG:

- Oracle: check the feasibility of the dual $\overline{u}$;
- Reduced costs: $s_j = c_j - \overline{u}^T a_j$, $\forall j \in N$;
- But the columns are not explicit and, hence,

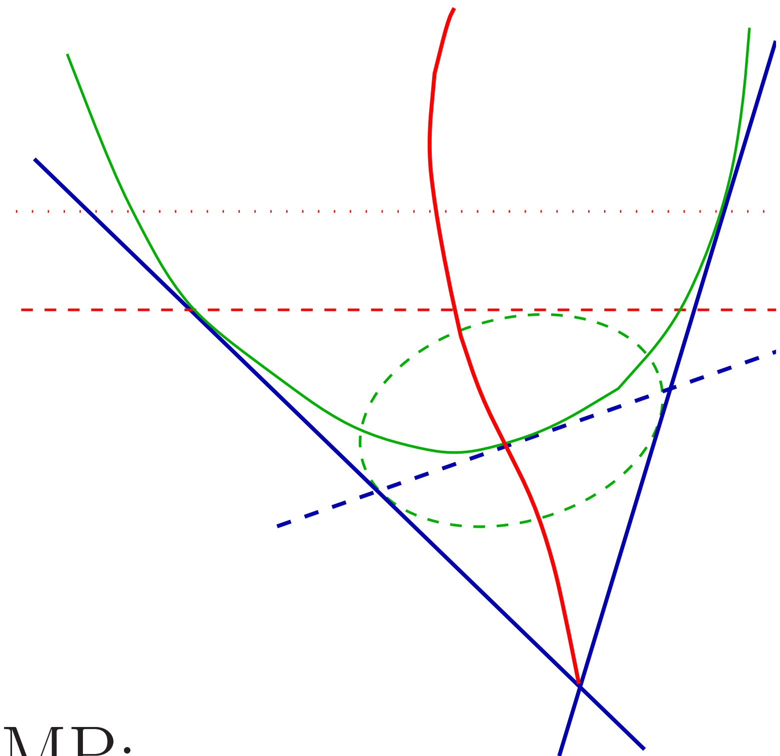
$$z_{SP} := \min\{c_j - \overline{u}^T a_j | a_j \in \mathcal{A}\}.$$

- (we reset $z_{SP} := 0$, if $z_{SP} > 0$);
- Lower Bound: $LB = z_{RMP} + \kappa z_{SP} \leq z^\star$, where

$$\kappa \geq \sum_{i \in N} \lambda_i^\star,$$

- If $z_{SP} < 0$, then new columns are generated;
- Otherwise, an optimal solution of the MP was found!

# Appealing features of IPMs:



- Use IPM to solve the RMP:
  → **no degeneracy issues**
- Terminate RMP solution early:
  → **use reliable estimate of duality gap**
  → **get stable dual solution $\bar{u}$**

## PDCGM Algorithm

Parameters: $\varepsilon_{\max}$, $D > 1$, $\delta$, $\kappa$

1. set $\mathrm{LB} = -\infty$, $\mathrm{UB} = \infty$, $\mathrm{gap} = \infty$, $\varepsilon = 0.5$;
2. while ($\mathrm{gap} > \delta$) do
3. find a well-centred $\varepsilon$-opt $(\tilde{\lambda}, \tilde{u})$ of the RMP;
4. $\mathrm{UB} = \min\{\mathrm{UB}, \tilde{z}_{RMP}\}$;
5. call the oracle with the query point $\tilde{u}$;
6. $\mathrm{LB} = \max\{\mathrm{LB}, \kappa \tilde{z}_{SP} + b^T \tilde{u}\}$;
7. $\mathrm{gap} = (\mathrm{UB} - \mathrm{LB})/(1 + |\mathrm{UB}|)$;
8. $\varepsilon = \min\{\varepsilon_{\max}, \ \mathrm{gap}/D\}$;
9. if ($\tilde{z}_{SP} < 0$) then add new columns to the RMP;
10. end (while)

## Theorem

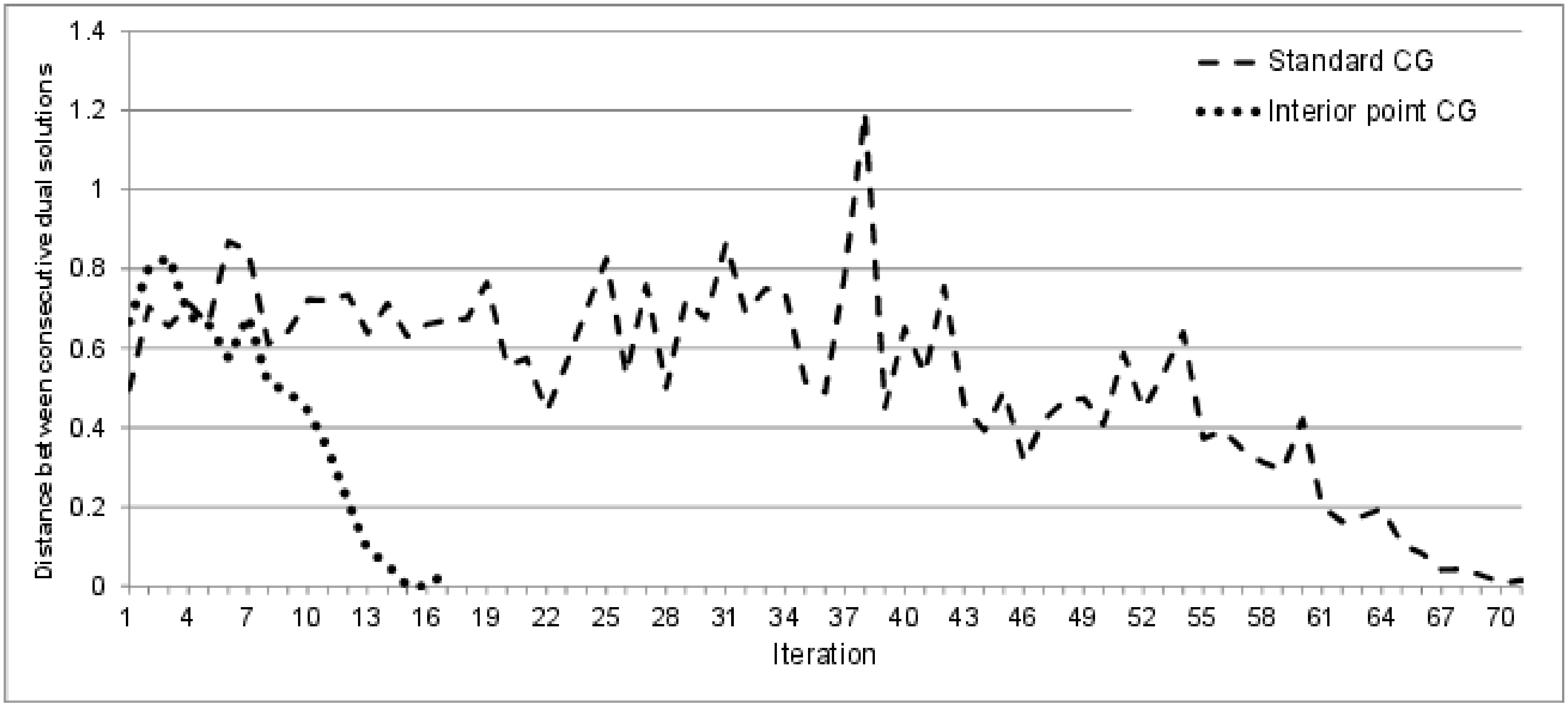
Let $z^*$ be the optimal solution of the (MP).
Given $\delta > 0$, the primal-dual column generation method converges in a finite number of steps to a feasible solution $\hat{\lambda}$ of the MP with objective value $\hat{z}$ that satisfies

$$\hat{z} - z^* \leq \delta(1 + |z^*|).$$

**JG, González-Brevis and Munari**,
New developments in the primal-dual column generation technique, *European J. of Oper Res* 224 (2013) 41–51.

# Oscillation in a real instance

Changes of dual variables between iterations $j$ and $j+1$, $\|u^j - u^{j+1}\|_2$ on VRPTW instance (Solomon C207):



# Warmstarting IPMs

# A need to solve a sequence of similar problems

- **column generation**
- **cutting plane methods**
- subproblems in the block-angular LPs
  (Dantzig-Wolfe decomp., Benders decomp.)
- **B&B, (and B&Cut, B&Cut&Price, etc)**
- SQP
- any sequence of similar problems
  example: computing efficient frontier in Markowitz portfolio optimization

**Warm Starts** Which method should be used?

- Simplex Method, or
- Interior Point Method.

$$\boldsymbol{x_j s_j = 0, \ \forall j}$$

**When is the Simplex Method better?**

→ **few** indices change optimal partition
B & B, adding *one* cut in CPM, etc.

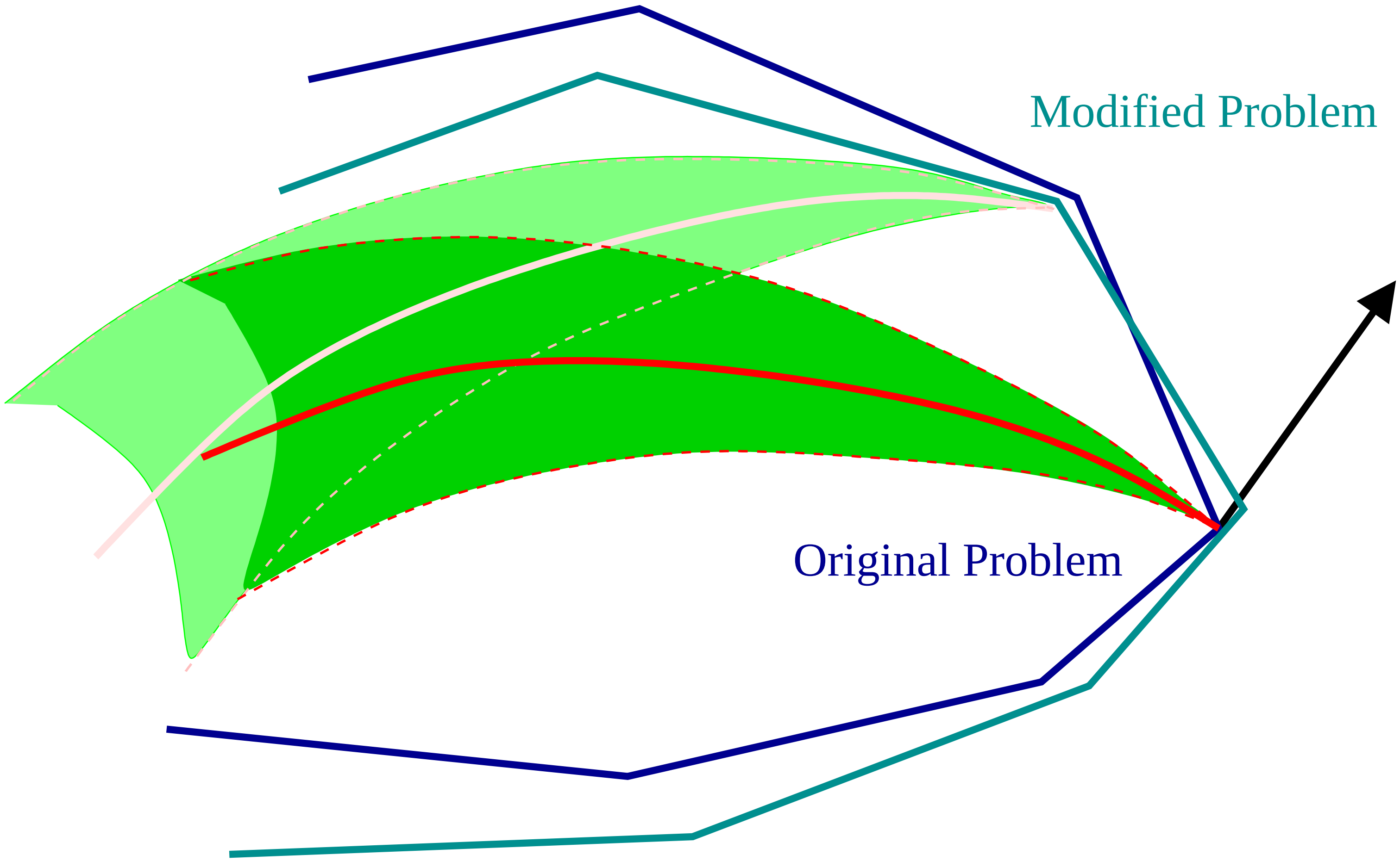
**When is the Interior Point Method better?**

→ **many** indices change optimal partition
adding *many* cuts in CPM,
dealing with a general change of problem data, etc

**Conjecture:**

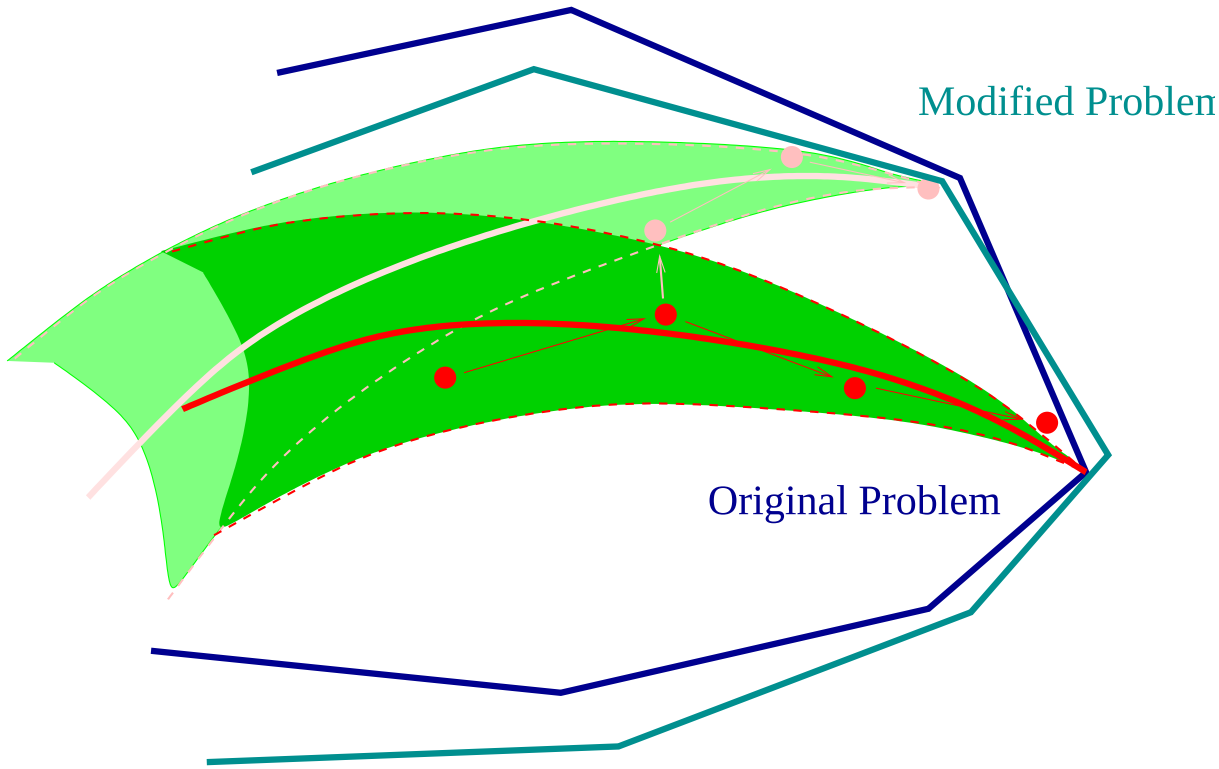
The more changes in the (large) problem
the more attractive IPM-based warm starts are.

# Difficulty of IPM Warm Starts



# Warmstarting Heuristic

Idea: Start close to the (new) central path, not close to the (old) solution




**JG**, Warm start of the primal-dual method applied in CPM, *Mathematical Programming* 83 (1998) 125–143

# Warm Start in PDCGM context: new results

## Theorem

Let $n$ and $n+k$ be the dims of RMP and newRMP, resp.
Let a $\mu^0$-centre $(x^0, y^0, z^0) \in N_S(\gamma)$ of RMP be given.
The warm start solution $(x^w, \bar{x}^w, y^w, z^w, \bar{z}^w) \in N_S(\beta\gamma)$ produced by the new warmstarting algorithm satisfies

$$(x^w)^T s^w + (\bar{x}^w)^T \bar{s}^w \leq \frac{6}{\gamma} \frac{n+k}{n} (x^0)^T s^0.$$

**JG and González-Brevis**,
A new warmstarting strategy for the primal-dual column generation method, *Mathematical Programming A* 152 (2015) 113–146

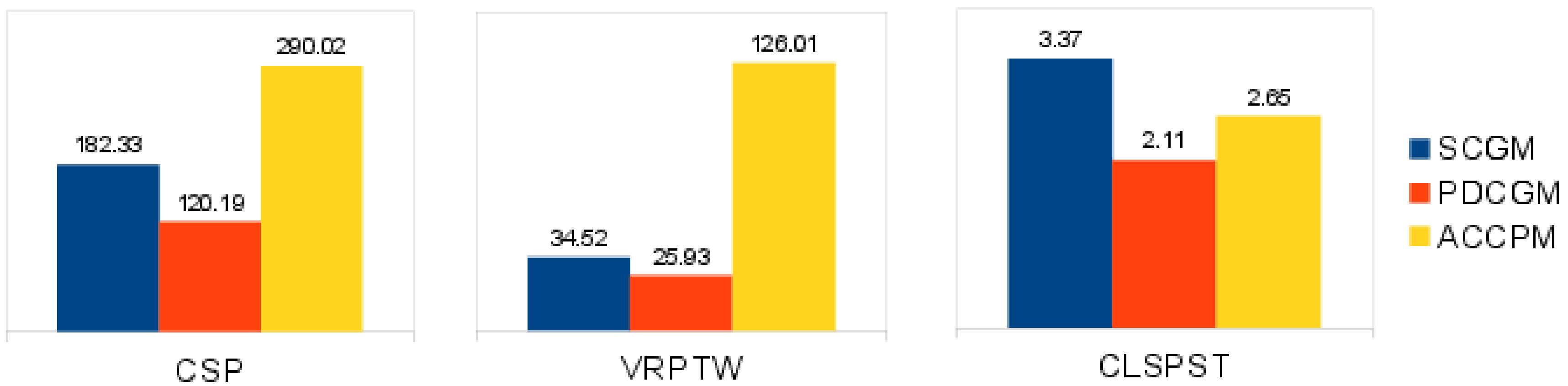
# Applications of PDCGM

- UFLP: Uncapacitated Facility Location Problem
- MKL: Multiple Kernel Learning problem
- TSSP: Two-Stage Stochastic Problem
- MCNF: MultiCommodity (MinCost) Network Flow problem
- CSP: Cutting Stock Problem
- CLSPST: Capacitated Lot-Sizing Problem with Setup Times
- VRPTW: Vehicle Routing Pb with Time Windows

**PDCGM Software** available at:

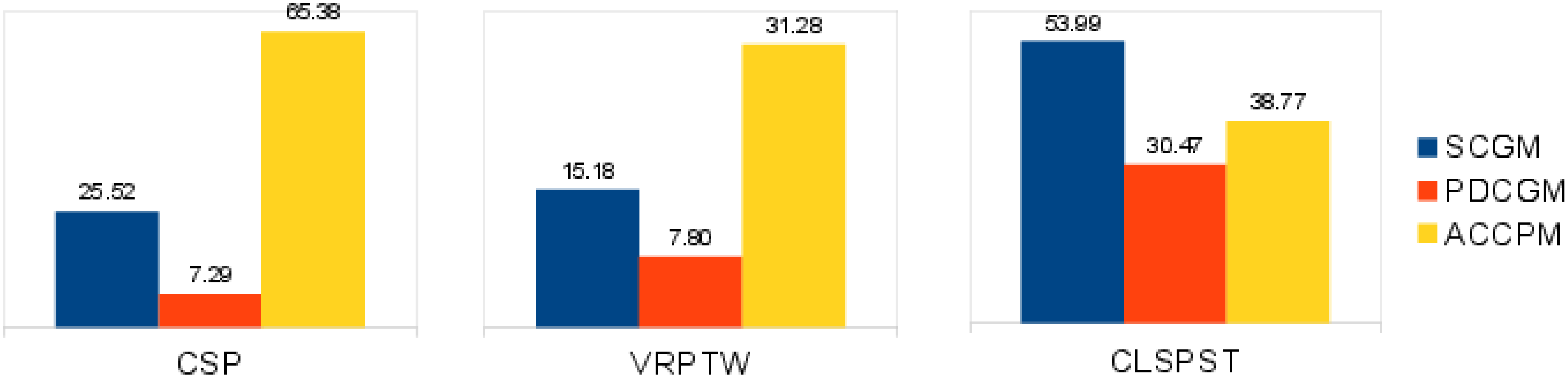
`http://www.maths.ed.ac.uk/~gondzio/software/pdcgm.html`

# **Number of iterations** (relative to PDCGM)




| | CSP | VRPTW | CLSPST |
|---|---|---|---|
| SCGM | 1.52 | 1.33 | 1.60 |
| ACCPM | 2.41 | 4.86 | 1.26 |

# CPU time [s] (relative to PDCGM)




| | CSP | VRPTW | CLSPST |
|---|---|---|---|
| SCGM | 3.50 | 1.95 | 1.26 |
| ACCPM | 8.97 | 4.01 | 1.27 |

# Conclusions

A completely new perspective is needed to exploit the insight offered by IPMs in a number of **combinatorial optimization** applications:

- column generation
- cutting plane methods
- B & B, (and B & Cut, B & Cut & Price, etc)

**Warmstarting works well** in the CG/CPM context:

problems are re-optimized in **3-5** IPM iterations

# References

**JG, Gonzalez-Brevis, Munari**, New developments in the primal-dual column generation technique, *European Journal of Oper Res*, 224 (2013) 41–51.

**Munari, JG**, Using the primal-dual interior point algorithm within the branch-price-and-cut method, *Computers and Oper Res*, 40 (2013) 2026–2036.

**JG, Gonzalez-Brevis**, A new warmstarting strategy for the primal-dual column generation method, *Mathematical Programming A*, 152 (2015) 113–146.

**JG, González-Brevis, Munari**, Large-scale optimization with the primal-dual column generation method, *Mathematical Programming Comput*, 8 (2016) 47–82.

http://www.maths.ed.ac.uk/~gondzio/software/pdcgm.html

# Vehicle Routing Problem with Time Windows

A company delivers goods to customers $i \in C$.
The company has vehicles $k \in V$ and each of them starts at a depot, travels to several customers and returns to the depot. The visit of vehicle $k$ to customer $i$ needs to take place in a specific time window: $a_i \leq s_{ik} \leq b_i$, where $s_{ik}$ is the time when vehicle $k$ reaches customer $i$.

Objective: Minimize the total cost of delivery.

Define binary variable $x_{ijk}$ which takes value 1 if vehicle $k$ travels from customer $i$ to customer $j$ $(k \in V,\ i, j \in C)$ and takes value zero otherwise.

# Vehicle Routing Problem with Time Windows

## Constraints:

Exactly one vehicle leaves customer $i$:

$$\sum_{k \in V} \sum_{j \in N} x_{ijk} = 1, \quad \forall i \in C$$

Vehicle capacity constraint:

$$\sum_{i \in C} d_i \sum_{j \in N} x_{ijk} \leq q, \quad \forall k \in V$$

Each vehicle leaves the depot and returns to it:

$$\sum_{j \in N} x_{0jk} = 1 \quad \text{and} \quad \sum_{j \in N} x_{i(n+1)k} = 1, \quad \forall k \in V$$

# VRPTW: Constraints (continued)

Time-window constraint

$$s_{ik} + t_{ij} - M(1 - x_{ijk}) \leq s_{jk}, \quad \forall i, j \in N, \forall k \in V.$$

Since $x_{ijk}$ is binary the above constraint has the following meaning: If $x_{ijk} = 1$ (vehicle $k$ travels from customer $i$ to customer $j$) then

$$s_{ik} + t_{ij} \leq s_{jk}$$

that is, the arrival time of vehicle $k$ to customer $j$ is greater than or equal the sum of time when vehicle $k$ arrives to customer $i$ and the time $t_{ij}$ it takes to travel from $i$ to $j$.

Otherwise (if $x_{ijk} = 0$) the presence of “big” $M$ guarantees that the constraint is always inactive.

# VRPTW

$$
\begin{aligned}
\min \quad & \sum_{k\in V}\sum_{i\in N}\sum_{j\in N} c_{ij}x_{ijk} & \\
\text{s.t.} \quad & \sum_{k\in V}\sum_{j\in N} x_{ijk} = 1, & \forall i \in C,\\
& \sum_{i\in C} d_i \sum_{j\in N} x_{ijk} \le q, & \forall k \in V,\\
& \sum_{j\in N} x_{0jk} = 1, \ \sum_{i\in N} x_{i(n+1)k} = 1, & \forall k \in V,\\
& \sum_{i\in N} x_{ihk} - \sum_{j\in N} x_{jhk} = 0, & \forall h \in C, \forall k \in V,\\
& s_{ik} + t_{ij} - M\big(1 - x_{ijk}\big) \le s_{jk}, & \forall i, j \in N, \forall k \in V,\\
& a_i \le s_{ik} \le b_i, & \forall i \in N, \forall k \in V,\\
& x_{ijk} \in \{0, 1\}, & \forall i, j \in N, \forall k \in V.
\end{aligned}
$$