Outline

1. Column Generation & Cut Separation in the Dual
2. Stabilization techniques
3. Numerical Analysis
Assume a bounded integer integer problem:

\[
[F] \equiv \min \{ cx : \]
\[x \in Y \equiv Ax \geq a \]
\[x \in Z \equiv Bx \geq b \]
\[x \in \mathbb{N}^n \]

Let \( X := Y \cap Z \). Assume that subproblem

\[
[SP] \equiv \min \{ cx : x \in Z \} \quad \text{(1)}
\]

is “relatively easy” to solve compared to problem \([F]\). Then,

\[
Z = \{ Z^q \}_{q \in Q}
\]

\[
= \{ x \in \mathbb{N}^n : x = \sum_{q \in Q} z^q \lambda_q, \sum_{q \in Q} \lambda_q = 1; \lambda_q \geq 0 \ \forall q \in Q \}
\]

and

\[
\text{conv}(Z) = \{ x \in \mathbb{R}_+^n : \sum_{q \in Q} z^q \lambda_q, \sum_{q \in Q} \lambda_q = 1, \lambda_q \geq 0 \ q \in Q \}
\]
\[ L(\pi) := \min_{q \in Q} \{ cz^q + \pi(a - Az^q) \} \]

\[ [LD] := \max_{\pi \in R^m_+} \min_{q \in Q} \{ cz^q + \pi(a - Az^q) \} \]
\[ [LD] \equiv \max_{\pi \in R^m_+} \min_{q \in Q} \{ \pi \cdot a + (c - \pi A)z^q \}; \]

\[ \equiv \max \{ \eta, \] 
\[ \eta \leq cz^q + \pi(a - Az^q) \quad q \in Q, \] 
\[ \pi \in R^m_+, \eta \in R^1 \}; \]

\[ \equiv \min \{ \sum_{q \in Q} (cz^q)\lambda_q, \] 
\[ \sum_{q \in Q} (Az^q)\lambda_q \geq a, \] 
\[ \sum_{q \in Q} \lambda_q = 1, \] 
\[ \lambda_q \geq 0 \quad q \in Q \}; \]

\[ \equiv \min \{ cx : Ax \geq a, x \in \text{conv}(Z) \}. \]
\[
\min \sum_{q \in Q} \lambda_q (cx)
\]
\[
\sum_{q \in Q} (Az^q) \lambda_q \geq a
\]
\[
\sum_{q \in Q} \lambda_q = 1
\]
\[
\lambda_q \in \{0, 1\} \quad \forall q \in Q.
\]

\[
[M^t] \equiv \min \left\{ \sum_{q \in Q^t} c \lambda_q : \sum_{q \in Q^t} A z^q \lambda_q \geq a; \sum_{q \in Q^t} \lambda_q = 1; \lambda_q \geq 0, \; q \in Q^t \right\}
\]

\[
[DM^t] \equiv \max \left\{ \eta : \pi (Az^q - a) + \eta \leq c z^q, \; q \in Q^t; \pi \in R^m_+; \eta \in R^1 \right\}
\]
\[ [M^t] \equiv \min \{ cx : Ax \geq a, x \in \text{conv}(\{z^q\}_{q \in Q^t}) \}. \]

\[ L^t() : \pi \rightarrow L^t(\pi) = \min_{q \in Q^t} \{ \pi a + (c - \pi A)z^q \}; \]

**Solving** \([\text{LSP}(\pi^t)]\) **yields:**
1. most neg. red. cost col. for \([M^t]\)
2. most violated constr. for \([DM^t]\)
3. correct value \(L()\) at point \(\pi^t\)
Column Generation & Cut Separation in the Dual Restricted Master, Dual Polyhedra, & Pricing Oracle

\[ \text{[M}^t\text{]} \equiv \min \{cx : Ax \geq a, x \in \text{conv}(\{z^q\}_{q \in Q^t})\}. \]

\[ L^t() : \pi \rightarrow L^t(\pi) = \min_{q \in Q^t} \{\pi a + (c - \pi A)z^q\}; \]

\textbf{Solving [LSP(}\pi^t\text{)] yields:}

1. most neg. red. cost col. for [M$^t$]
2. most violated constr. for [DM$^t$]
3. correct value $L()$ at point $\pi^t$
Dual Polyhedra: Outer and Inner approximations

\[ \pi \]

\[ \eta \]

\[ L(\pi) \]

\[ (\pi^t, \eta^t) \]

\[ (\hat{\pi}, \hat{L}) \]
Column generation for a master program
≡ cut generation for the dual master

⇓

Cutting plane “strategies” translate into in col. gen. “stabilization”

- In-Out separation [BenAmeurNeto07, FischettiSalvagnin10]
- Central point cutting strategy [GoffinVial, LeePark11]
- Lexicographic Simplex [ZanetteFischettiBalas11]
- ...

Colgen Workshop, June 2012
Column generation for a master program
≡ cut generation for the dual master

↓

Cutting plane “strategies” translate into in col. gen. “stabilization”

- In-Out separation [BenAmeurNeto07, FischettiSalvagnin10]
- Central point cutting strategy [GoffinVial, LeePark11]
- Lexicographic Simplex [ZanetteFischettiBalas11]
- ...

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
A sequence of candidate **dual solutions**

\[
\{\pi^t\}_t \rightarrow \pi^* \in \Pi^*
\]

A sequence of candidate **primal solutions** (a by-product to prove optimality)

\[
\{x^t\}_t \rightarrow x^* \in X^*
\]

**Oracle:**

\[
z^t \leftarrow \arg\min_{x \in Z} \{(c - \pi^{t-1} A) x\}.
\]

1. **Ascent methods:**
   - Subgradient
   - Volume
   - Conjugate Sub-gradient

2. **Polyhedral methods:**
   - Kelley
   - Bundle
   - ACCPM
A sequence of candidate dual solutions

\[ \{\pi^t\}_t \to \pi^* \in \Pi^* \]

A sequence of candidate primal solutions (a by-product to prove optimality)

\[ \{x^t\}_t \to x^* \in X^* \]

Oracle: \( z^t \leftarrow \text{argmin}_{x \in Z} \{(c - \pi^{t-1} A)x\} \).

1. Ascent methods:
   - Subgradient
   - Volume
   - Conjugate Sub-gradient

2. Polyhedral methods:
   - Kelley
   - Bundle
   - ACCPM
Convergence
\[ \{\pi^t\}_t \rightarrow \pi^* \in \Pi^* \]

- **Dual oscillations:** $\pi^t$ jump erratically (*bang-bang*), $\{L(\pi^t)\}$ non monotonic (*yo-yo*) and possibly

\[ ||\pi^t - \pi^*|| > ||\pi^{t-1} - \pi^*|| \]

- **Tailing-off effect:** towards the end, added inequalities yield marginal improvements / step sizes are very small.

- **Primal degeneracy / multiple dual optima:**
  - fewer non zero $\lambda_q$ than master constraints;
  - dual system with fewer constraints than variables;
  - cuts co-linear with objective (inherent to cutting plane procedures).
Penalty functions to drive the optimization towards $\hat{\pi}$:

$$\pi^t := \arg\max_{\pi \in R^m_+} \{L^t(\pi) - \hat{S}(\pi)\}$$

Dual price smoothing (In-Out separation)

[Wentges97] $\tilde{\pi}^t = \alpha \hat{\pi} + (1 - \alpha) \pi^t$

[Neame99] $\tilde{\pi}^t = \alpha \tilde{\pi}^{t-1} + (1 - \alpha) \pi^t$

Dual price centralization (central point separation)

- Analytic center (of trust polyhedra)
- Chebyshev center
- Optimal face center
**Stabilization Techniques**

- **Penalty functions** to drive the optimization towards $\hat{\pi}$:

  $$\pi^t := \operatorname{argmax}_{\pi \in \mathbb{R}_+^m} \{ L^t(\pi) - \hat{S}(\pi) \}$$

- **Dual price smoothing** *(In-Out separation)*

  - [Wentges97] $\tilde{\pi}^t = \alpha \hat{\pi} + (1 - \alpha) \pi^t$
  - [Neame99] $\tilde{\pi}^t = \alpha \tilde{\pi}^{t-1} + (1 - \alpha) \pi^t$

- **Dual price centralization** *(central point separation)*

  - Analytic center (of trust polyhedra)
  - Chebyshev center
  - Optimal face center
\[ \tilde{\pi}^t = \alpha \hat{\pi} + (1 - \alpha) \pi^t \leftrightarrow \text{(Wentges' rule)} \]

\( \pi^{\text{in}}, \eta^{\text{in}} := (\hat{\pi}, \hat{L}) \) \hspace{1cm} (2)

\( \pi^{\text{out}}, \eta^{\text{out}} := (\pi^t, \eta^t) \).

\( \pi^{\text{sep}}, \eta^{\text{sep}} := \alpha (\pi^{\text{in}}, \eta^{\text{in}}) + (1 - \alpha) (\pi^{\text{out}}, \eta^{\text{out}}). \) \hspace{1cm} (4)
\[
\begin{align*}
(\pi^{\text{in}}, \eta^{\text{in}}) := & \begin{cases} 
(\hat{\pi}, \hat{L}) & \text{under Wentges's rule.} \\
(\tilde{\pi}^{t-1}, L(\tilde{\pi}^{t-1})) & \text{under Neame's rule,}
\end{cases} \\
(\pi^{\text{out}}, \eta^{\text{out}}) := & (\pi^t, \eta^t). \\
(\pi^{\text{sep}}, \eta^{\text{sep}}) := & \alpha (\pi^{\text{in}}, \eta^{\text{in}}) + (1 - \alpha) (\pi^{\text{out}}, \eta^{\text{out}}).
\end{align*}
\]

**Oracle:** \[z^t \leftarrow \text{argmin}_{x \in Z}\{(c - \pi^{\text{sep}}A)x\} \]
In-Out Separation

Case A: SEP is cut, so is OUT
Case B: SEP is not cut, but OUT is cut
Case C: neither SEP nor OUT is cut → “mis-price”
1. If SEP is cut by $z^t$, OUT is cut.

2. Otherwise,
   a) SEP defines the next IN point.
   b) \[ L(\pi^{sep}) - \eta^{in} \geq (1 - \alpha)(\eta^{out} - \eta^{in}) \], i.e.,
      \[ \eta^{out} - L(\pi^{sep}) \leq \alpha(\eta^{out} - \eta^{in}) \]
   c) The OUT point might be cut.
   d) Otherwise, $\pi^t = \pi^{t-1} = \pi^{out}$ and
      \[ \|\tilde{\pi}^t - \pi^{out}\| = \alpha\|\tilde{\pi}^{t-1} - \pi^{out}\| \]
The number of columns that can be generated is finite.

A column, once added, cannot be regenerated: the associated constraint is satisfied by both IN and OUT points (hence by SEP).

However, there are iterations where no columns are generated, and the OUT point remains unchanged → mis-price

Thus, on needs to prove convergence of a subsequence of mis-price iterations, starting with $\pi^{in} = \tilde{\pi}^0$. Then,

$$\tilde{\pi}^{t+1} = \alpha_t \tilde{\pi}^t + (1 - \alpha_t) \pi^{out}.$$ 

Hence,

$$||\tilde{\pi}^{t+1} - \pi^{out}|| = \alpha_t ||\tilde{\pi}^t - \pi^{out}|| = \ldots = \prod_{\tau=0}^{t} \alpha_{\tau} ||\tilde{\pi}^0 - \pi^{out}||$$

If $\alpha = 0.8$,

At iteration $t$, $||\tilde{\pi}^0 - \pi^{out}||$ is cut by a factor $(1 - \prod_{\tau=0}^{t} \alpha_{\tau})$.

**BEWARE:** convergence is only asymptotic for a constant $0 < \alpha < 1$. 

Colgen Workshop, June 2012  Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
The number of columns that can be generated is finite.

A column, once added, cannot be regenerated: the associated constraint is satisfied by both IN and OUT points (hence by SEP).

However, there are iterations where no columns are generated, and the OUT point remains unchanged → mis-price

Thus, one needs to prove convergence of a subsequence of mis-price iterations, starting with $\pi^\text{in} = \tilde{\pi}^0$. Then,

$$
\tilde{\pi}^{t+1} = \alpha_t \tilde{\pi}^t + (1 - \alpha_t) \pi^\text{out}.
$$

Hence,

$$
||\tilde{\pi}^{t+1} - \pi^\text{out}|| = \alpha_t ||\tilde{\pi}^t - \pi^\text{out}|| = \ldots = \prod_{\tau=0}^t \alpha_\tau ||\tilde{\pi}^0 - \pi^\text{out}||
$$

If $\alpha = 0.8$,

At iteration $t$, $||\tilde{\pi}^0 - \pi^\text{out}||$ is cut by a factor $(1 - \prod_{\tau=0}^t \alpha_\tau)$.

**BEWARE:** convergence is only asymptotic for a constant $0 < \alpha < 1$. 

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
The number of columns that can be generated is finite.

A column, once added, cannot be regenerated: the associated constraint is satisfied by both IN and OUT points (hence by SEP).

However, there are iterations where no columns are generated, and the OUT point remains unchanged → mis-price

Thus, on needs to prove convergence of a subsequence of mis-price iterations, starting with $\pi^{in} = \tilde{\pi}^0$. Then,

$$\tilde{\pi}^{t+1} = \alpha_t \tilde{\pi}^t + (1 - \alpha_t) \pi^{out}.$$

Hence,

$$||\tilde{\pi}^{t+1} - \pi^{out}|| = \alpha_t ||\tilde{\pi}^t - \pi^{out}|| = \ldots = \prod_{\tau=0}^{t} \alpha_\tau ||\tilde{\pi}^0 - \pi^{out}||$$

If $\alpha = 0.8$

At iteration $t$, $||\tilde{\pi}^0 - \pi^{out}||$ is cut by a factor $(1 - \prod_{\tau=0}^{t} \alpha_\tau)$.

BEWARE: convergence is only asymptotic for a constant $0 < \alpha < 1$. 
The number of columns that can be generated is finite.

A column, once added, cannot be regenerated: the associated constraint is satisfied by both IN and OUT points (hence by SEP).

However, there are iterations where no columns are generated, and the OUT point remains unchanged → mis-price.

Thus, on needs to prove convergence of a subsequence of mis-price iterations, starting with $\pi^{in} = \tilde{\pi}^0$. Then,

$$\tilde{\pi}^{t+1} = \alpha_t \tilde{\pi}^t + (1 - \alpha_t) \pi^{out}.$$  

Hence,

$$||\tilde{\pi}^{t+1} - \pi^{out}|| = \alpha_t ||\tilde{\pi}^t - \pi^{out}|| = \ldots = \prod_{\tau=0}^{t} \alpha_\tau ||\tilde{\pi}^0 - \pi^{out}||$$

**if $\alpha = 0.8$**

At iteration $t$, $||\tilde{\pi}^0 - \pi^{out}||$ is cut by a factor $(1 - \prod_{\tau=0}^{t} \alpha_\tau)$.

**BEWARE:** convergence is only asymptotic for a constant $0 < \alpha < 1$. 

Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
The number of columns that can be generated is finite.

A column, once added, cannot be regenerated: the associated constraint is satisfied by both IN and OUT points (hence by SEP).

However, there are iterations where no columns are generated, and the OUT point remains unchanged → mis-price

Thus, one needs to prove convergence of a subsequence of mis-price iterations, starting with $\pi^{in} = \tilde{\pi}^0$. Then,

$$\tilde{\pi}^{t+1} = \alpha_t \tilde{\pi}^t + (1 - \alpha_t)\pi^{out}. \text{ Hence,}$$

$$||\tilde{\pi}^{t+1} - \pi^{out}|| = \alpha_t ||\tilde{\pi}^t - \pi^{out}|| = \ldots = \prod_{\tau=0}^{t} \alpha_{\tau} ||\tilde{\pi}^0 - \pi^{out}||$$

If $\alpha = 0.8$

At iteration $t$, $||\tilde{\pi}^0 - \pi^{out}||$ is cut by a factor $(1 - \prod_{\tau=0}^{t} \alpha_{\tau})$.

**BEWARE:** convergence is only asymptotic for a constant $0 < \alpha < 1$. 
In case of mis-pricing, one observes a sequence of iterations in which:

\[ \|\tilde{\pi}_{t+1} - \pi^{out}\| = \alpha_t \|\tilde{\pi}_t - \pi^{out}\| \]
\[ \|\eta^{t+1} - L(\tilde{\pi}^{t+1})\| \leq \alpha_t \|\eta^t - L(\tilde{\pi}^t)\| \]

Hence, as \( (1 - \Pi_{\tau=0}^t \alpha_{\tau}) \to 1 \),

- \( \tilde{\pi}^t \to \pi^{out} \)
- \( L(\tilde{\pi}^t) \to \eta^{out} \)

After a finite number of iterations, \( \|\eta^t - L(\tilde{\pi}^t)\| \leq \epsilon \) and the master is considered as optimized. As there is a finite number of possible values for \( L(\tilde{\pi}^t) \) and \( \eta^t \),

\[ \exists \epsilon : \|\eta^t - L(\tilde{\pi}^t)\| \leq \epsilon \Rightarrow \eta^t = L(\tilde{\pi}^t) \]

which proves finite convergence in theory [BenAmeurNeto07, Wentges97].

In practice, for finite convergence, one better choose \( \alpha_{\tau} : \)

\[ (1 - \Pi_{\tau=0}^t \alpha_{\tau}) \geq 1 \quad \text{for a finite} \ t \]
Algorithm for handling a *mis-pricing* sequence

Step 0 $k \leftarrow 1$

Step 1 $\beta \leftarrow [1 - k \ast (1 - \alpha)]^+$

Step 2 $\pi_{\text{sep}} = \beta \pi^0 + (1 - \beta) \pi^\text{out}$

Step 3 $k \leftarrow k + 1$

Step 4 call the oracle on $\pi_{\text{sep}}$

Step 5 if mis-pricing occurs, goto Step 1 else, goto Step 0.

\[ \text{if } \alpha = 0.8 \]

I.e; we chose $\alpha_\tau : (1 - \Pi_{\tau=0}^k \alpha_\tau) = k \ast (1 - \alpha)$. Hence,

\[ (1 - \Pi_{\tau=0}^k \alpha_\tau) \geq 1 \text{ after } k = \left\lceil \frac{1}{(1 - \alpha)} \right\rceil \text{ iterations, and smoothing stops} \]

with $\beta = 0$. 
- Turn-off smoothing during **HEADING-IN** Phase.
- Trust $\pi^{\text{out}}$ during **TAILING-OFF** Phase: small $\alpha$.
- **Decreasing** $\alpha$ as the algorithm converges: f.i., record decreased $\alpha$-value in mis-pricing sequence.
Sensitivity to static $\alpha$: Machine Scheduling

12 jobs per machines
(25 jobs / 2 machines ; 50 jobs / 4 machines)

- Master iterations
- Subproblem calls

- Time (sec.)

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Machine Scheduling

25 jobs per machines
25 jobs / 1 machine; 50 jobs / 2 machines; 100 jobs / 4 machines

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Machine Scheduling

50 jobs per machines

50 jobs / 1 machine; 100 jobs / 2 machines; 200 jobs / 4 machines
Sensitivity to static $\alpha$: Machine Scheduling

100 jobs per machines
100 jobs / 1 machine

---

- **Colgen Workshop, June 2012**
- **Col. Gen. Stabilization using Dual Smoothing: Theory & Practice**
Sensitivity to static $\alpha$: Generalized Assignment

OR-Library C, D, E instances: 5 jobs per machine

- Master iterations
- Subproblem calls

Time (sec.)
Sensitivity to static $\alpha$: Generalized Assignment

OR-Library C, D, E instances: 10 jobs per machine

- Master iterations
- Subproblem calls

- Time (sec.)

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Generalized Assignment

OR-Library C, D, E instances: 20 jobs per machine

- Master iterations
- Subproblem calls

- Time (sec.)

Colgen Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Generalized Assignment

OR-Library C, D, E instances: 40 jobs per machine

- Master iterations
- Subproblem calls

Time (sec.)

0 0.2 0.4 0.6 0.8 0.95
0 $\cdot 10^4$
0 1 1.5 2

Colgen Workshop, June 2012 | Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Vertex Coloring

**Number of calls to the oracle and Master CPU time**

![Graph showing the number of calls to the oracle and Master CPU time vs. smoothing parameter ($\alpha$)](image)

- **Legend**:
  - queen
  - myciel
  - DSJ-1
  - DSJ-2
  - miles
  - mugg
  - mulsol
  - zeroin
  - fpsol2
  - school
  - average

**Axes**:
- **x-axis**: Smoothing parameter ($\alpha$)
- **y-axis** left: Number of subproblem calls
- **y-axis** right: Time spent at solving master problems

Colgen Workshop, June 2012

Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Vertex Coloring

Subproblem CPU time and Overall CPU time

- Queen
- Myciel
- DSJ-1
- DSJ-2
- Miles
- Mugg
- Mulsol
- Zeroin
- Fpsol2
- School
- Average

Smithson Workshop, June 2012
Col. Gen. Stabilization using Dual Smoothing: Theory & Practice
Sensitivity to static $\alpha$: Vertex Coloring

Dual price vector density & subproblem CPU time on instance queen12_12, using the oracle of [Ostergard, 2001]
Dual smoothing is the equivalent of IN-OUT separation ↔ cross-fertilization with the **cutting-plane community**

Dual smoothing can have a **large impact on overall performance**.

Dual smoothing requires a **single parameter**; but it needs to be **carefully set**.

**Hard instances** for Kelley seems to be those with many columns (f.i. more jobs per machine). Then, it seems best to apply heavy stabilization f.i. $\alpha = 0.85$.

**Dynamic and auto-adaptative $\alpha$-schedules** are to be derived.

**Stabilization** can make the pricing **subproblem harder** and eliminates iterations where pricing is easy.