Nonconvex Combinatorial Nonlinear Optimization: New Methodologies and Critical Applications

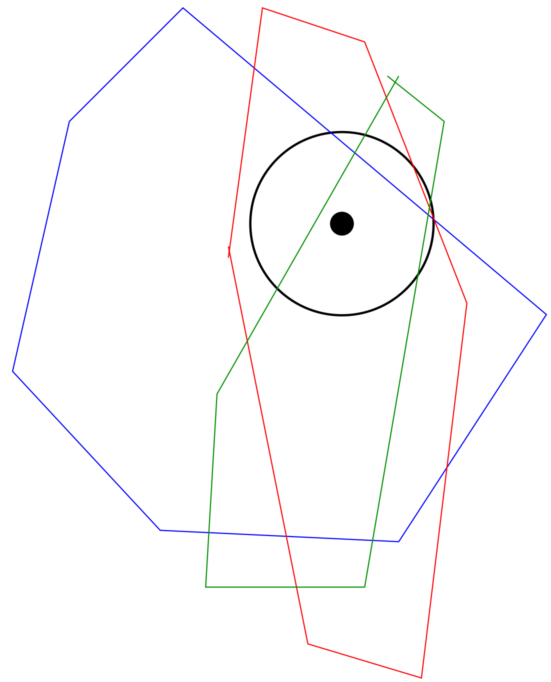
Daniel Bienstock (Columbia), Ismael deFarias (Texas Tech)

1. The "SUV" problem

- given full-dimensional polyhedra $P^1, \ldots, P^K$ in $\mathbb{R}^d$,
- find a point closest to the origin *not* contained inside any of the P^h.

$$\begin{aligned} &\min \ \|x\|^2 \\ &s.t. \ \ x \in \mathbb{R}^d - \bigcup_{h=1}^{K} \operatorname{int}(P^h), \end{aligned}$$

(application: X-ray lythography)



- Typical values for d (dimension): less than 20; usually even smaller
- Typical values for K (number of polyhedra): possibly hundreds, but often less than 50
- Very hard problem

Formulation as mixed-integer quadratic program

$$(\text{Polyhedron } \boldsymbol{P^h} \;:\; \{x \in \mathbb{R}^d : a_{h,i}^T x \le b_{h,i}, \quad 1 \le i \le m_h\})$$

$$\begin{aligned}
\min \quad & \sum_{j=1}^{d} x_j^2 \\
& \boldsymbol{y_{h,i} \in \{0,1\}}, \quad 1 \le i \le m_h,\ 1 \le h \le K \\
s.t. \quad & a_{h,i}^T x \ge b_{h,i}\ \boldsymbol{y_{h,i}}, \quad 1 \le i \le m_h,\ 1 \le h \le K \\
& \sum_{i=1}^{m_h} y_{h,i} \ge 1, \quad 1 \le h \le K \\
& x \in \mathbb{R}^d, \quad 1 \le i \le m_h,\ 1 \le h \le K.
\end{aligned}$$

A hard instance: 33 symmetric polyhedra in $\mathbb{R}^8$

$\rightarrow$ 561 constraints, 536 variables (264 binaries)

A hard instance: 33 symmetric polyhedra in $\mathbb{R}^8$

$\rightarrow$ 561 constraints, 536 variables (264 binaries)

A hard instance: 33 symmetric polyhedra in $\mathbb{R}^8$

$\rightarrow$ 561 constraints, 536 variables (264 binaries)

$\rightarrow$ Experiments with Cplex 12.6, current 8-core 48 GB machine

A hard instance: 33 symmetric polyhedra in $\mathbb{R}^8$

→ 561 constraints, 536 variables (264 binaries)

→ Experiments with Cplex 12.6, current 8-core 48 GB machine

Time (sec.)	**Lower Bound**	**Upper Bound**	**Nodes**
500	0.00	0.2645	6×10^6
1000	0.00	0.2257	1.1×10^7
1500	0.00	0.2257	1.6×10^7
2500	0.00	0.2257	2.7×10^7
3000	0.00	0.2257	3.2×10^7
3600	0.00	0.2257	3.8×10^7
7200	0.00	0.2257	7.9×10^7

A hard instance: 33 symmetric polyhedra in $\mathbb{R}^8$

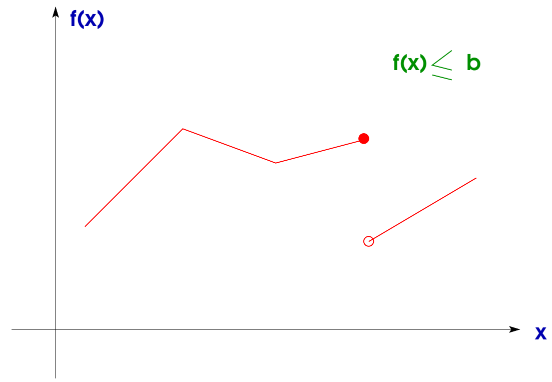
$\longrightarrow$ 561 constraints, 536 variables (264 binaries)

$\longrightarrow$ Experiments with Cplex 12.6, current 8-core 48 GB machine

Time (sec.)	**Lower Bound**	**Upper Bound**	**Nodes**
500	0.00	0.2645	6×10^6
1000	0.00	0.2257	1.1×10^7
1500	0.00	0.2257	1.6×10^7
2500	0.00	0.2257	2.7×10^7
3000	0.00	0.2257	3.2×10^7
3600	0.00	0.2257	3.8×10^7
7200	0.00	0.2257	7.9×10^7

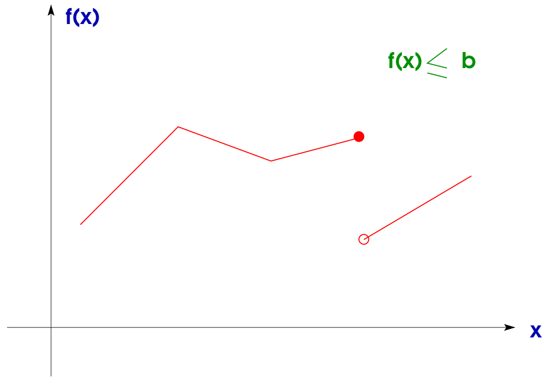
$\longrightarrow$ (Other techniques:) upper bound **0.0977**

2. Optimization with nonstandard constraints



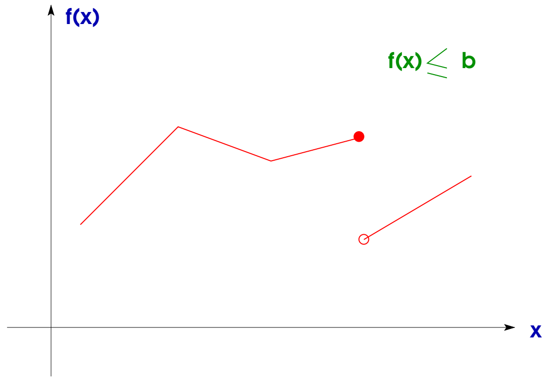
- Piecewise-linear functions
- Discontinuous variables
- Disjunctions on complex conditions: "either $x^T y = 0$ or $\sum x_i \geq 5$"

2. Optimization with nonstandard constraints



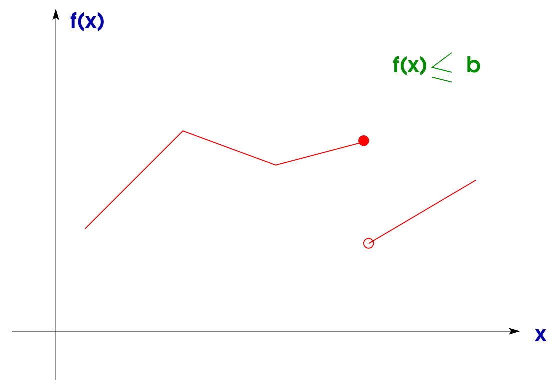
- Piecewise-linear functions

2. Optimization with nonstandard constraints



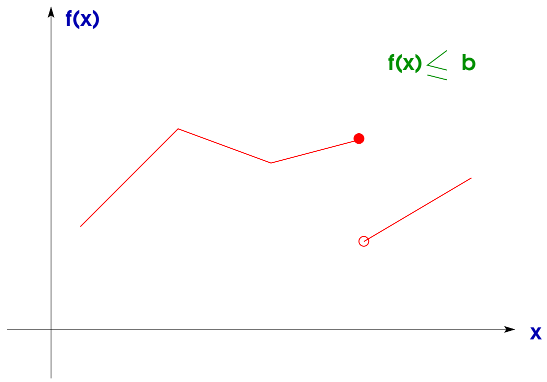
- Piecewise-linear functions
- Discontinuous or *semi-continous* variables, e.g. "$x \in [0,1] \cup [2,3]$"

2. Optimization with nonstandard constraints



- Piecewise-linear functions
- Discontinuous or *semi-continous* variables, e.g. "$x \in [0, 1] \cup [2, 3]$"
- Disjunctions on complex conditions: "either $x^T y = 0$ or $\sum x_i \geq 5$"

2. Optimization with nonstandard constraints



- Piecewise-linear functions
- Discontinuous or *semi-continous* variables, e.g. "$x \in [0,1] \cup [2,3]$"
- Disjunctions on complex conditions: "either $x^T y = 0$ or $\sum x_i \geq 5$"

These arise in:

- Pricing problems
- Applications in physical sciences
- Approximations of nonlinear functions

A hard piecewise-linear optimization instance

→ Experiments with Cplex 12.6, current 8-core 48 GB machine

Time (sec.)	**Lower Bound**	**Upper Bound**	**Gap**	**Nodes**
50	559687.7609			5700
500	560556.7700	613016.6495	8.56 %	22411
1500	561991.0724	608745.6914	7.68 %	56041
3000	566861.1899	608745.6914	6.88 %	150056
5000	567845.8559	607282.2571	6.49 %	279090
22759	571076.4105	606578.6048	5.85 %	1.3×10^6

→ Our techniques: **optimal value**, in **338 seconds**

2. Cardinality constrained, convex quadratic programming

$$\begin{aligned} \min \ & x^T Q x + c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq 0, \quad \|x\|_0 \leq k \end{aligned}$$

$\|x\|_0$ = number of nonzero entries in x.

- $Q \succeq 0$
- $x \in \mathbb{R}^n$ for n possibly large
- k relatively small, e.g. $k = 100$ for $n = 10000$
- VERY hard problem – just getting good bounds is tough

3. AC-OPF problem in rectangular coordinates

Given a power grid, determine voltages at every node so as to minimize a convex objective

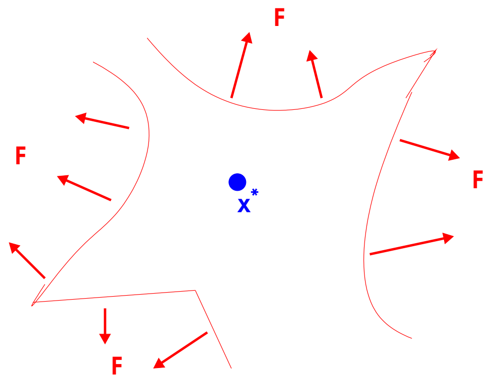
$$\begin{aligned} &\min \ v^T A v \\ &\text{s.t.} \quad L_k \le v^T F_k v \le U_k, \quad k = 1, \ldots, K \\ &\quad v \in \mathbb{R}^{2n}, \quad (n = \text{number of nodes}) \end{aligned}$$

- voltages are complex numbers; v is the vector of voltages in rectangular coordinates (real and imaginary parts)
- $A \succeq 0$
- n could be in the tens of thousands, or more
- the F_k are very sparse (neighborhood structure for every node)
- Problem HARD when grid under distress and $L_k \approx U_k$.

Why are these problems so hard

$$\text{Generic problem:} \quad \min \; Q(x), \quad s.t. \quad x \in F,$$

- $Q(x)$ (strongly) convex, especially: positive-definite quadratic
- F nonconvex

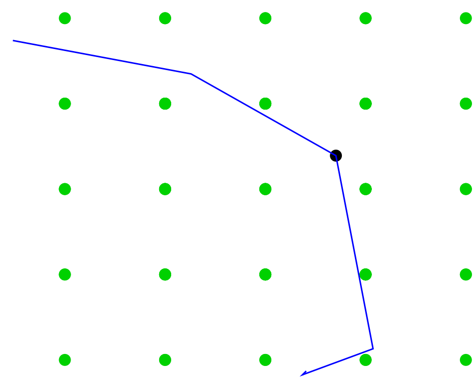


x^* solves $\min \left\{ Q(x), \;\; : \;\; x \in \hat{F} \right\}$ where $F \subset \hat{F}$ and $\hat{F}$ convex

$\longrightarrow$ straightforward relaxations are weak

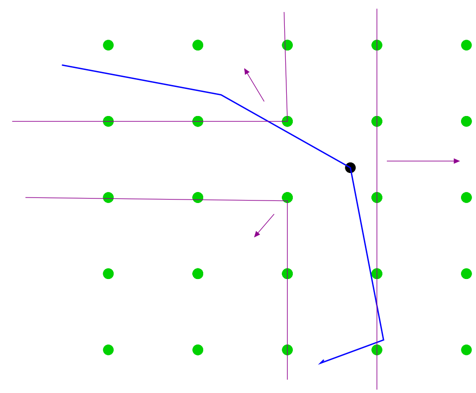
Lattice-free cuts for **linear** integer programming

Generic problem: $\min \; c^T x, \quad s.t. \quad Ax \leq b, \;\; z \in Z^n$



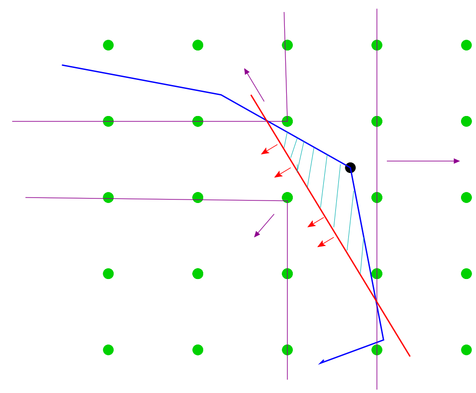
Lattice-free cuts for **linear** integer programming

Generic problem: $\min \; c^T x, \quad s.t. \quad Ax \leq b, \;\; z \in Z^n$



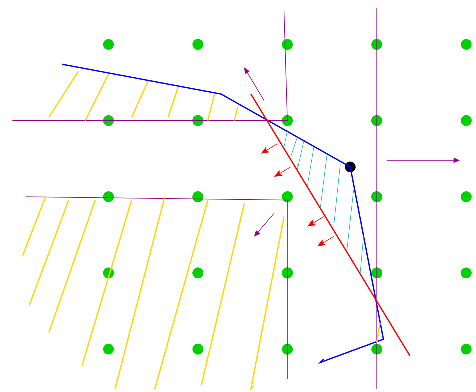
Lattice-free cuts for **linear** integer programming

Generic problem: $\min \; c^T x, \quad s.t. \quad Ax \leq b, \;\; z \in Z^n$



Lattice-free cuts for **linear** integer programming

Generic problem: $\min \; c^T x, \quad s.t. \quad Ax \leq b, \;\; z \in Z^n$



Special case: standard **disjunctions**

How to apply in a continuous, nonconvex setting?

Technique 1: Exclude-and-cut

$$\begin{aligned} &\min \ Q(x) \\ \text{s.t.} \quad & x \in F \end{aligned}$$

Technique 1: Exclude-and-cut

$$\begin{aligned} & \min \; z \\ \text{s.t.} \quad & z \geq Q(x), \\ & x \in F \end{aligned}$$

Technique 1: Exclude-and-cut

$$\begin{array}{rl} \min & z \\ \text{s.t.} & z \ \geq \ Q(x), \\ & x \in F \end{array}$$

0. $\hat{F}$: a **convex relaxation** of $\text{conv}\,\{(x, z) \,:\, z \ \geq \ Q(x), \ x \in F\}$

1. Let $(x^*, z^*) \ = \ \text{argmin}\{\, z \,:\, (x, z) \in \hat{F}\}$

Technique 1: Exclude-and-cut

$$\begin{aligned} \min\ & z \\ \text{s.t.}\quad & z \geq Q(x), \\ & x \in F \end{aligned}$$

0. $\hat{F}$: a **convex relaxation** of conv $\{(x,z) : z \geq Q(x),\ x \in F\}$

1. Let $(x^*, z^*) = \operatorname{argmin}\{ z : (x,z) \in \hat{F}\}$

2. Find an **open set** S s.t. $x^* \in S$ and $S \cap F = \emptyset$.
Examples: lattice-free sets, geometry

Technique 1: Exclude-and-cut

$$\begin{aligned} \min \;& z \\ \text{s.t.} \;& z \geq Q(x), \\ & x \in F \end{aligned}$$

0. $\hat{\boldsymbol{F}}$: a **convex relaxation** of conv $\{(x, z) : z \geq Q(x),\ x \in F\}$

1. Let $(x^*, z^*) = \text{argmin}\{ z : (x, z) \in \hat{F}\}$

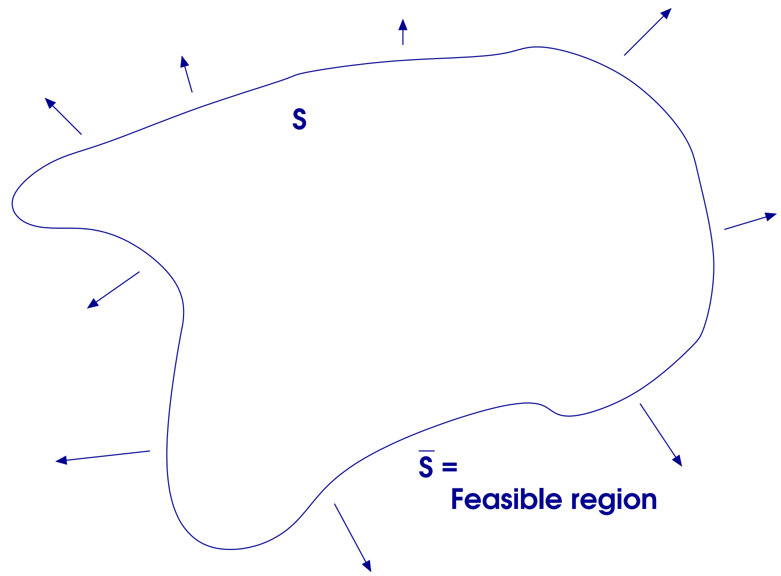
2. Find an **open set** $\boldsymbol{S}$ s.t. $x^* \in S$ and $S \cap F = \emptyset$.
Examples: lattice-free sets, geometry

3. Add to the formulation an inequality $\boldsymbol{az + \alpha^T x \geq \alpha_0}$ valid for

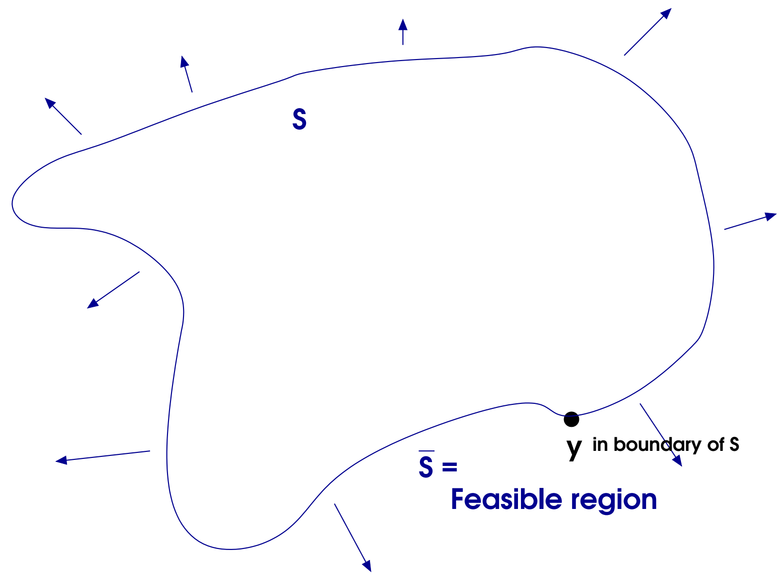
$$\{ (x, z) : x \in \overline{S},\ z \geq Q(x) \}$$

but violated by (x^*, z^*).

Valid **linear** inequalities for $\{\,(x,z)\,:\,x \in \overline{S},\;\; z \;\geq\; Q(x)\,\}$.



Valid **linear** inequalities for $\{\,(x,z)\,:\,x \in \overline{S},\ z \ \geq\ Q(x)\,\}$.

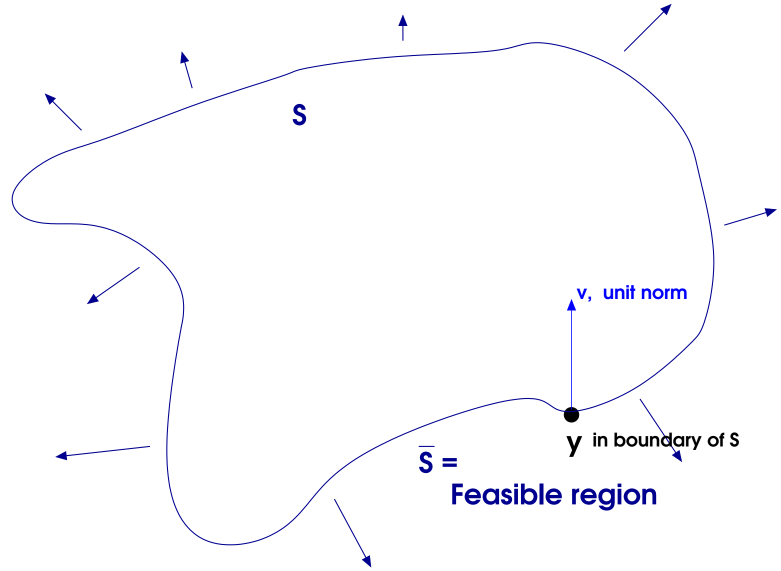


First order inequality:

$$z \ \geq\ [\nabla Q(y)]^T (x-y) + Q(y)$$

is valid EVERYWHERE – does not cut-off any points

Valid **linear** inequalities for $\{ (x, z) \,:\, x \in \overline{S},\ z \ \geq\ Q(x) \}$.



First order inequality:

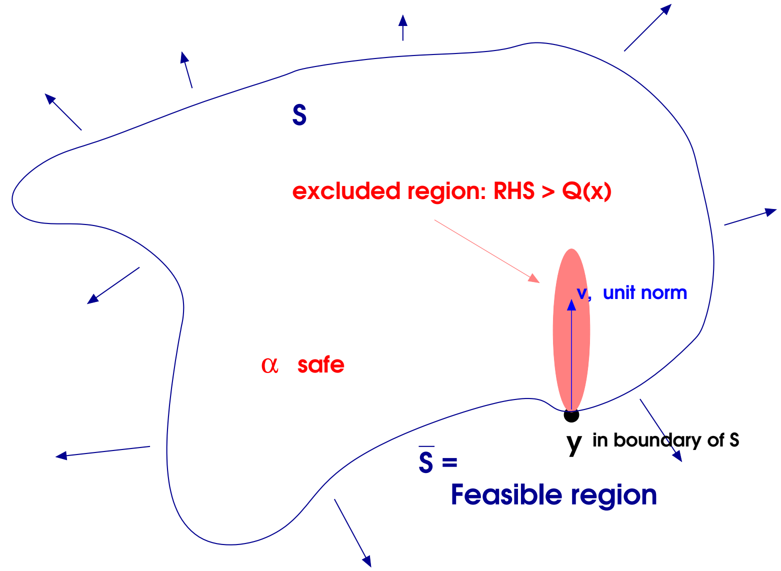
$$z \ \geq\ [\nabla Q(y)]^T (x - y) + Q(y)$$

is valid EVERYWHERE – does not cut-off any points **Lifted** first order inequality, for $\alpha \geq 0$:

$$z \ \geq\ \underbrace{[\nabla Q(y)]^T (x - y) + Q(y)}_{\text{first-order term} \approx Q(x)} + \underbrace{\alpha v^T (x - y)}_{\text{lifting}}$$

NOT valid EVERYWHERE: RHS $> Q(x)$ for $\alpha > 0$, $v^T(x - y) > 0$ and $x \approx y$.
– want $RHS \leq Q(x)$ in $\bar{S}$ ($\alpha = 0$ always OK)

Valid **linear** inequalities for $\{(x, z) : x \in \overline{S},\ z \geq Q(x)\}$.



First order inequality:

$$z \geq [\nabla Q(y)]^T (x - y) + Q(y)$$

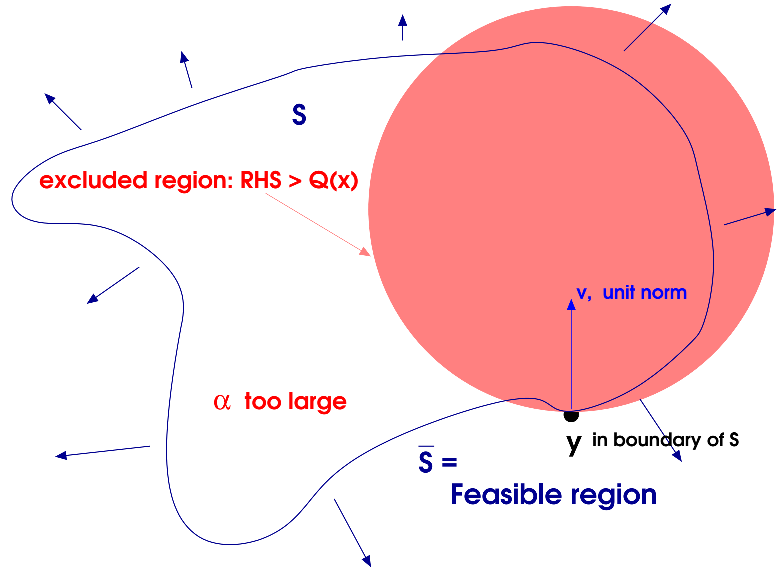
is valid EVERYWHERE – does not cut-off any points **Lifted** first order inequality, for $\alpha \geq 0$:

$$z \geq \underbrace{[\nabla Q(y)]^T (x - y) + Q(y)}_{\text{first-order term} \approx Q(x)} + \underbrace{\alpha v^T (x - y)}_{\text{lifting}}$$

NOT valid EVERYWHERE: RHS $> Q(x)$ for $\alpha > 0$, $v^T (x - y) > 0$ and $x \approx y$.

– want $RHS \leq Q(x)$ in $\bar{S}$ ($\alpha = 0$ always OK)

Valid **linear** inequalities for $\{ (x, z) : x \in \overline{S},\ z \ \geq \ Q(x) \}$.



First order inequality:

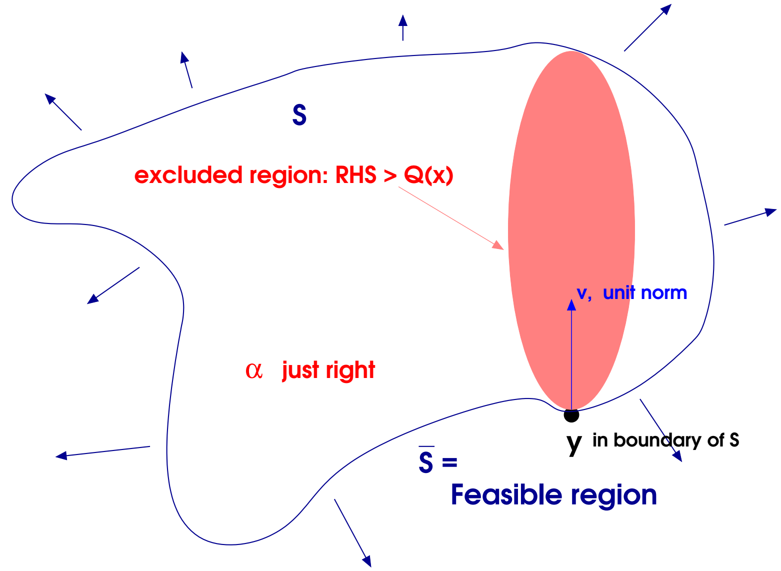
$$z \ \geq \ [\nabla Q(y)]^T (x-y) + Q(y)$$

is valid EVERYWHERE – does not cut-off any points **Lifted** first order inequality, for $\alpha \geq 0$:

$$z \ \geq \ \underbrace{[\nabla Q(y)]^T (x-y) + Q(y)}_{\text{first-order term} \approx Q(x)} + \underbrace{\alpha v^T (x-y)}_{\text{lifting}}$$

NOT valid EVERYWHERE: RHS $> Q(x)$ for $\alpha > 0$, $v^T(x-y) > 0$ and $x \approx y$.
– want $RHS \leq Q(x)$ in $\bar{S}$ ($\alpha = 0$ always OK)

Valid **linear** inequalities for $\{ (x, z) \, : \, x \in \overline{S}, \; z \; \geq \; Q(x) \}$.



First order inequality:

$$z \; \geq \; [\nabla Q(y)]^T (x - y) + Q(y)$$

is valid EVERYWHERE – does not cut-off any points **Lifted** first order inequality, for $\alpha \geq 0$:

$$z \; \geq \; \underbrace{[\nabla Q(y)]^T (x - y) + Q(y)}_{\text{first-order term} \approx Q(x)} \; + \; \underbrace{\alpha v^T (x - y)}_{\text{lifting}}$$

NOT valid EVERYWHERE: RHS $> Q(x)$ for $\alpha > 0$, $v^T(x - y) > 0$ and $x \approx y$.

– want $RHS \leq Q(x)$ in $\bar{S}$ ($\alpha = 0$ always OK)

Valid **linear** inequalities for $\mathcal{F} \doteq \{\,(x,z) \in \mathbb{R}^n \times \mathbb{R} \;:\; x \in \overline{S},\;\; z \;\geq\; Q(x)\,\}$.

Given $y \in \partial S$, let

$\alpha^* \doteq \;\; \boldsymbol{sup}\,\{\,\boldsymbol{\alpha \geq 0} \;:\; Q(x) \;\geq\; [\nabla Q(y)]^T(x-y) + Q(y) + \alpha v^T(x-y) \;\;\}$

valid for $\mathcal{F}$. Note: $\alpha^* = \alpha^*(v,y)$

Theorem. If Q is convex and differentiable, then $conv(\mathcal{F})$ is given by

$$Q(x) \;\geq\; [\nabla Q(y)]^T(x-y) + Q(y) \qquad \forall y$$
$$Q(x) \;\geq\; [\nabla Q(y)]^T(x-y) + Q(y) + \alpha^* v^T(x-y)$$
$$\forall v \text{ and } y \in \partial S.$$

(abridged)

Quadratics in action

Lifted first-order inequalities for $\mathcal{F} = \{ (x, z) : x \in \overline{S},\ z \geq Q(x) \}$.

$$(Q(x) \succeq 0)$$

Separation problem

Given $(x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}$, find a lifted-first order inequality maximally violated by (x^*, z^*) (if any)

Theorem: We can separate in polynomial time when:

- $\bar{S}$ (or S) is a union of polyhedra
- S is an ellipsoid or paraboloid (many cases)

Quadratics in action

Lifted first-order inequalities for $\mathcal{F} \ = \ \{ \, (x, z) \ : \ x \in \overline{S}, \ z \ \geq \ Q(x) \, \}$.

$$(\, Q(x) \succeq 0 \,)$$

Separation problem

Given $(x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}$, find a lifted-first order inequality maximally violated by (x^*, z^*) (if any)

Theorem: We can separate in polynomial time when:

- $\bar{S}$ (or S) is a union of polyhedra
- S is an ellipsoid or paraboloid (many cases)

Quadratics in action

Lifted first-order inequalities for $\mathcal{F} \;=\; \{\,(x,z)\;:\;x \in \overline{S},\;\; z \;\geq\; Q(x)\,\}$.

$$(\,Q(x) \succeq 0\,)$$

Separation problem

Given $(x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}$, find a lifted-first order inequality maximally violated by (x^*, z^*) (if any)

Theorem: We can separate in polynomial time when:

- $\bar{S}$ (or S) is a union of polyhedra
- S is an ellipsoid or paraboloid (many cases)

Key proof technique: S-Lemma

$$\begin{array}{lll} \min & Q_1(x) & \\ \text{s.t.} & Q_2(x) & \leq \;\; 0 \\ & & x \in \mathbb{R}^n \end{array}$$

($Q_i(x)$ **arbitrary** quadratics) is poly-time solvable

Ongoing work: S-Lemma

$$\begin{aligned} \min \quad & Q_1(x) \\ \text{s.t.} \quad & Q_2(x) \;\leq\; 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

($Q_i(x)$ **arbitrary** quadratics) is poly-time solvable

Trust-region subproblem:

$$\begin{array}{lrl} \min & Q_1(x) & \\ \text{s.t.} & \|x\| & \leq 1 \\ & x \in \mathbb{R}^n & \end{array}$$

Extension

$$
\begin{array}{llrcll}
\textbf{(TGEN):} & \min & x^T A x + b^T x + c & & & \\
& \text{s.t.} & \|x - x^k\|^2 & \leq & f_k & k = 1, \ldots, L_k \\
& & \|x - y^k\|^2 & \geq & g_k & k = 1, \ldots, M_k \\
& & \|x - z^k\|^2 & = & h_k & k = 1, \ldots, E_k \\
& & a_i^T x & \leq & b_i & i = 1, \ldots, m \\
& & x \in \mathbb{R}^n. & & &
\end{array}
$$

Extension

$$
\textbf{(TGEN):} \quad \begin{array}{lrcll}
\min & x^T A x + b^T x + c & & & \\
\text{s.t.} & \|x - x^k\|^2 & \leq & f_k & k = 1, \ldots, L_k \\
& \|x - y^k\|^2 & \geq & g_k & k = 1, \ldots, M_k \\
& \|x - z^k\|^2 & = & h_k & k = 1, \ldots, E_k \\
& a_i^T x & \leq & b_i & i = 1, \ldots, m \\
& x \in \mathbb{R}^n. & & &
\end{array}
$$

- $P = \{x \,:\, a_i^T x \;\leq\; b_i \quad i = 1, \ldots, m\}$
- $F^* =$ the number of **faces** of P that intersect $\bigcap_k \{x \,:\, \|x - x^k\| \leq f_k\}$.

Extension

$$
\textbf{(TGEN):}\quad \begin{array}{lllll}
\min & x^T A x + b^T x + c & & & \\
\text{s.t.} & \|x - x^k\|^2 & \leq & f_k & k = 1, \ldots, L_k \\
& \|x - y^k\|^2 & \geq & g_k & k = 1, \ldots, M_k \\
& \|x - z^k\|^2 & = & h_k & k = 1, \ldots, E_k \\
& a_i^T x & \leq & b_i & i = 1, \ldots, m \\
& x \in \mathbb{R}^n. & & &
\end{array}
$$

- $P = \{x \,:\, a_i^T x \;\leq\; b_i \quad i = 1, \ldots, m\}$
- $F^* =$ the number of **faces** of P that intersect $\bigcap_k \{x \,:\, \|x - x^k\| \leq f_k\}$.

Theorem: For every **fixed** $L_k \geq 1$, $M_k \geq 0$, $E_k \geq 0$, problem **TGEN** can be solved in time polynomial in the problem size and F^*.

(SODA 2014)

Extends results by Ye, Ye-Zhang, Burer-Anstreicher, Burer-Yang

Even more general

Barvinok (STOC 1992):

For each fixed $\boldsymbol{p \geq 1}$, there is a polynomial-time algorithm for deciding feasibility of a system

$$\begin{aligned} x^T M_i\, x &= 0, \quad 1 \leq i \leq p, \\ \|x\| &= 1, \end{aligned}$$

where the M_i are general matrices.

Even more general

Barvinok (STOC 1992):

For each fixed $\boldsymbol{p \geq 1}$, there is a polynomial-time algorithm for deciding feasibility of a system

$$\begin{aligned} x^T M_i x &= 0, \quad 1 \leq i \leq p, \\ \|x\| &= 1, \end{aligned}$$

where the M_i are general matrices.

- **Non-constructive.** Algorithm says "yes" or "no."
- **Computational model?**

Theorem.

For each fixed $m \geq 1$ there is a polynomial-time algorithm that, given an optimization problem

$$\begin{aligned} \min \quad & f_0(x) \doteq x^T Q_0 x + c_0^T x \\ \text{s.t.} \quad & x^T Q_i x + c_i^T x + d_i \leq 0 \quad 1 \leq i \leq m, \end{aligned}$$

where $Q_1 \succ 0$, and $0 < \epsilon < 1$, either

(1) proves that the problem is infeasible,

or

(2) computes an ϵ-feasible vector $\hat{x}$ such that there exists no feasible $x \in \mathbb{R}^n$ with $f_0(x) < f(\hat{x}) - \epsilon$.

The complexity of the algorithm is polynomial in the number of bits in the data and in $\log \epsilon^{-1}$

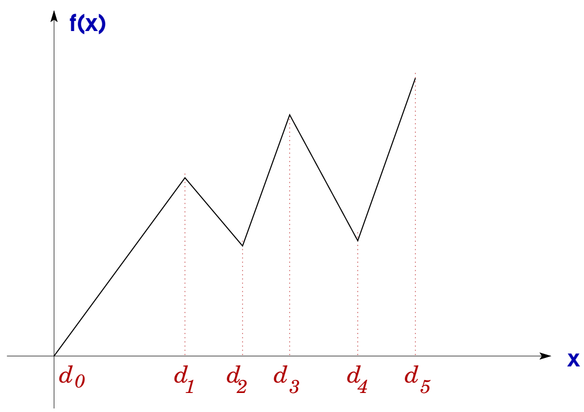
Technique 2: Methodologies for piecewise-linear optimization

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} f_i(x_i) \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

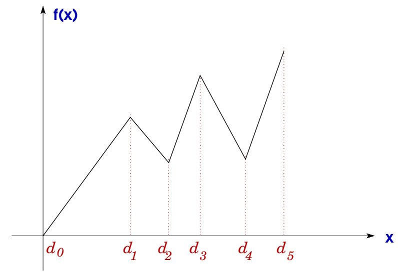
where for $1 \leq i \leq n$,

- $f_i \; : \; \mathbb{R} \to \mathbb{R}$
- f_i is continuous, piecewise linear.

We assume that some of the $\boldsymbol{f_i}$ are nonconcave.



Representing piecewise-linear functions

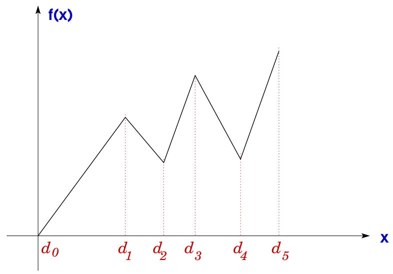


$$x = \sum_{k \in K} d_k \lambda_k$$

$$f(x) = \sum_{k \in K} f(d_k) \lambda_k$$

$$\sum_{k \in K} \lambda_k = 1, \quad \lambda \geq 0$$

Representing piecewise-linear functions



$$x = \sum_{k=1}^{T} d_k \lambda_k$$

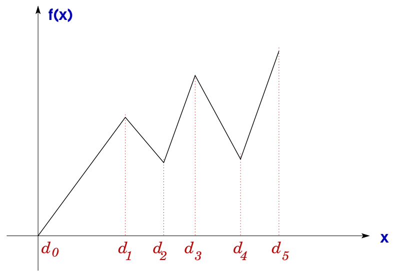
$$f(x) = \sum_{k=1}^{T} f(d_k) \lambda_k$$

$$\sum_{k=1}^{T} \lambda_k = 1, \quad \lambda \geq 0$$

And:

- At most 2 of the λ_k are nonzero
- if 2 of the λ_k are nonzero they have consecutive indices

Representing piecewise-linear functions



$$x = \sum_{k=1}^{T} d_k \lambda_k$$

$$f(x) = \sum_{k=1}^{T} f(d_k)\lambda_k$$

$$\sum_{k=1}^{T} \lambda_k = 1, \quad \lambda \geq 0$$

And:

- At most 2 of the λ_k are nonzero
- if 2 of the λ_k are nonzero they have consecutive indices

In other words $\{\lambda_1, \ldots, \lambda_T\}$ is a *special ordered set of type 2*, or **SOS2** set.

SOS2

Note that the **SOS2** method is more general than might seem. For example, it can be used to enforce:

- multiple-choices
- semi-continuous variables
- general integer variables

SOS2

Note that the **SOS2** method is more general than might seem. For example, it can be used to enforce:

- multiple-choices
- semi-continuous variables
- general integer variables

Example: enforcing semi-continuity. Suppose we want to model

$$\boldsymbol{x} \in \{\mathbf{0}\} \cup [\mathbf{1}, \mathbf{2}]$$

SOS2

Note that the **SOS2** method is more general than might seem. For example, it can be used to enforce:

- multiple-choice variables
- semi-continuous variables
- general integer variables

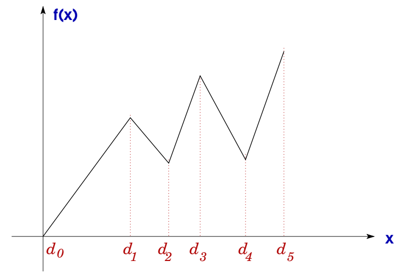
Example: enforcing semi-continuity. Suppose we want to model

$$\boldsymbol{x \in \{0\} \cup [1, 2]}$$

Then:

1. Set-up the SOS2 set $\{\lambda_0, \lambda, \lambda_1, \lambda_2\}$ (λ and all λ_i in $[0, 1]$).
2. Write $\boldsymbol{x = 0 \cdot \lambda_0 + \frac{1}{2} \cdot \lambda + 1 \cdot \lambda_1 + 2 \cdot \lambda_2}$.
3. Fix $\boldsymbol{\lambda = 0}$.

Back to epresenting piecewise-linear functions



$$f(x) = \sum_{k=1}^{T} f(d_k)\lambda_k$$

$$x = \sum_{k=1}^{T} d_k\lambda_k$$

$$\sum_{k=1}^{T} \lambda_k = 1, \quad \lambda \geq 0$$

$\{\boldsymbol{\lambda_1, \ldots, \lambda_T}\}$ is SOS2.

Putting it all together

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{n} f_i(x_i) \\
\text{s.t.} \quad & Ax \leq b \\
& x \in \mathbb{R}^n_+
\end{aligned}
$$

where each f_i is piecewise-linear

Putting it all together

$$\begin{aligned} \max \quad & \sum_{j=1}^{n} f_j(x_j) \\ \text{s.t.} \quad & Ax \le b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

where each f_j is piecewise-linear, with breakpoints d_j^k, $1 \le k \le T_j$

Putting it all together

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} f_j(x_j) \\
\text{s.t.} \quad & Ax \ \leq \ b \\
& x \in \mathbb{R}^n_+
\end{aligned}
$$

where each f_j is piecewise-linear, with breakpoints d_j^k, $1 \leq k \leq T_j$

→ use the SOS2 construction for each x_j, i.e.

$$
f_j(x_j) \ = \ \sum_{k=1}^{T_j} f_j(d_k)\lambda_k,
$$

$$
x_j \ = \ \sum_{k=1}^{T_j} d_k \lambda_j^k \tag{1}
$$

$$
\sum_{k=1}^{T_j} \lambda_j^k \ = \ 1, \quad \lambda \geq 0, \quad \{\lambda_1, \ldots, \lambda_{T_j}\} \quad \text{is SOS2} \tag{2}
$$

→ Derive cuts from (1)-(2) together with *each* constraint $\sum_j a_{ij} x_j \ \leq \ b_i$.

Underlying knapsack set

$$\sum_{j \in N} a_j x_j \leq b \qquad + \text{SOS2 construction for each } x_j\text{:}$$

Underlying knapsack set

$$\sum_{j \in N} a_j x_j \ \leq \ b \qquad + \text{SOS2 construction for each } x_j\text{:}$$

$$\sum_{j \in N^+} \sum_{k=0}^{T_j} a_j^k \lambda_j^k - \sum_{j \in N^-} \sum_{k=0}^{T_j} a_j^k \lambda_j^k \ \leq \ b \tag{3}$$

$$\sum_{k=1}^{T_j} \lambda_j^k \ \leq \ 1, \quad \forall j \in N \tag{4}$$

$$\lambda_j^k \geq 0,\ \forall j \in N,\ 0 \leq k \leq T_j \tag{5}$$

$$\left\{ \lambda_j^0, \ldots, \lambda_j^T \right\} \ \text{is SOS2}, \forall j \in N \tag{6}$$

(each $a_j^k \geq 0$)

Cuts:

- Lifted convexification constraints
- Cover, lifted cover inequalities

These are generalizations of clasical cut families but specific to our model.

Cuts:

- Lifted convexification constraints
- Cover, lifted cover inequalities

These are generalizations of clasical cut families but specific to our model.

But is this necessary?

After all, commercial solvers already have the generic versions of these cuts.

Cuts:

- Lifted convexification constraints
- Cover, lifted cover inequalities

These are generalizations of clasical cut families but specific to our model.

But is this necessary?

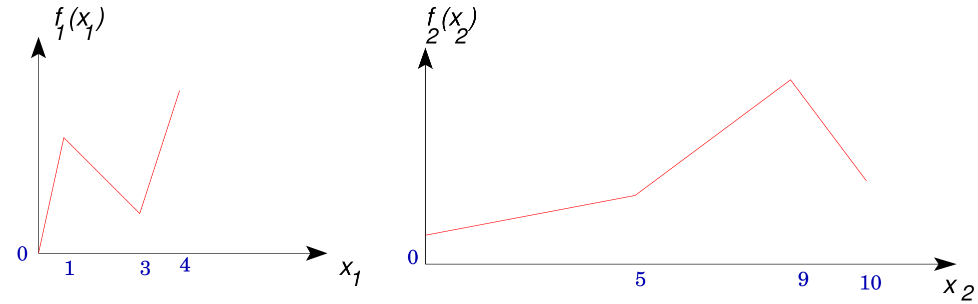
After all, commercial solvers already have the generic versions of these cuts.

Some computational tests:

- Minimum concave-cost transportation and transshipment problems
- Ranging from 25 supply and 50 demand nodes, 7 breakpoints, to 100 supply and 400 demand nodes, 22 breakpoints.
- Integrality gap is small – need a formulation to close it and prove a solution optimal.

# Nodes & part.	Time default	Time w/ cuts
25×50 & 5	936	18
25×100 & 5	971	34
25×200 & 5	2,578	101
25×300 & 5	3,600	103
25×400 & 5	3,600	479
50×100 & 5	171	37
50×200 & 5	272	43
50×300 & 5	617	99
50×400 & 5	1,754	139

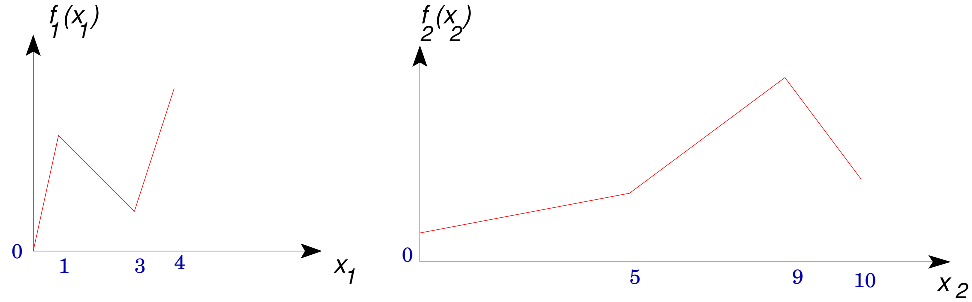
Example: lifted convexity constraint:



$$2x_1 + x_2 \ \le \ 10 \quad \rightarrow$$

$$(0 \cdot \lambda_1^0 + 2 \cdot \lambda_1^1 + 6\lambda_1^2 + 8\lambda_1^3) \ + \ (0 \cdot \lambda_2^0 + 5 \cdot \lambda_2^1 + 9\lambda_2^2 + 10\lambda_1^3) \ \le \ 10$$

Example: lifted convexity constraint:



$$2x_1 + x_2 \;\leq\; 10 \quad \rightarrow$$

$$(0 \cdot \lambda_1^0 + 2 \cdot \lambda_1^1 + 6\lambda_1^2 + 8\lambda_1^3) \;+\; (0 \cdot \lambda_2^0 + 5 \cdot \lambda_2^1 + 9\lambda_2^2 + 10\lambda_1^3) \;\leq\; 10$$

The point $\lambda_1^2 = 5/6, \lambda_2^1 = 1, \lambda_i^j = 0$ **otherwise** is an extreme point of the relaxation, but cut-off by the lifted convexity constraint:

$$-3\lambda_1^1 + \lambda_1^2 + 3\lambda_1^3 + 5\lambda_2^1 + 5\lambda_2^2 + 5\lambda_2^3 \;\leq\; 5$$

Theorem.

Let $N_1^- \subseteq N^-$ and $b' = b + \sum_{i \in N_1^-} a_i^{m_i}$, where $m_i \in K \; \forall i \in N_1^-$. Let $I = \left\{ i \in N^+ - \{j\} : a_j^s + a_i^T > b' \right\}$ and $k_i = \min \left\{ k \in K : a_j^s + a_i^k > b' \right\} \; \forall i \in I$. Suppose that $I \neq \emptyset$. Then,

$$\frac{1}{a_j^s} \sum_{k=1}^{s-1} a_j^k \lambda_j^k + \sum_{k=s}^{T} \lambda_j^k + \sum_{i \in I} \sum_{k=\max\{1, k_i - 1\}}^{T} \alpha_i^k \lambda_i^k - \sum_{i \in N_1^-} \sum_{k=m_i+1}^{T} \beta_i^k \lambda_i^k - \sum_{i \in N^- - N_1^-} \sum_{k \in K} \frac{a_i^k}{a_j^s} \lambda_i^k \leq 1$$

is valid for P, where

$$\left(\alpha_i^{k_i - 1}, \alpha_i^{k_i} \right) \in \left\{ (0,0) \, , \left(\frac{a_j^s + a_i^{k_i - 1} - b'}{a_j^s}, \frac{a_j^s + a_i^{k_i} - b'}{a_j^s} \right) \right\} \; \forall i \in I \text{ with } k_i > 1 \text{ and } a_j^s + a_i^{k_i - 1} < b',$$

$$\left(\alpha_i^{k_i - 1}, \alpha_i^{k_i} \right) = \left(0, \frac{a_j^s + a_i^{k_i} - b'}{a_j^s} \right) \; \forall i \in I \text{ with } k_i > 1 \text{ and } a_j^s + a_i^{k_i - 1} = b',$$

$$\alpha_i^{k_i} = 0 \; \forall i \in I \text{ with } k_i = 1,$$

$$\alpha_i^k = \frac{a_j^s + a_i^k - b'}{a_j^s} \; \forall i \in I \text{ with } k > k_i,$$

and

$$\beta_i^k = \frac{a_i^k - a_i^{m_i}}{a_j^s}.$$

(One of several such theorems)

Summary of results on cutting planes

- The vast majority of the instances of either transportation or transshipment could not be solved by GUROBI in default setting
- Virtually all instances are solved through proven optimality with the cuts
- For the instances GUROBI could solve without our cuts, the average reduction in computational time was of 92%, and in nodes 98%.

Numerical example – cardinality constrained convex QPs.

$$
\begin{aligned}
\min\ & x^T Q x + c^T x \\
\text{s.t.} \sum_j x_j\ &=\ 1 \\
x \geq 0,&\quad \|x\|_0 \leq k
\end{aligned}
$$

Numerical example – cardinality constrained convex QPs.

$$
\begin{aligned}
\min\ & x^T Q x + c^T x \\
\text{s.t.} \sum_{j=1}^{n} x_j \ &= \ 1 \\
& x \geq 0, \quad \|x\|_0 \leq k
\end{aligned}
$$

MIP Formulation

$$
\begin{aligned}
\min\ & x^T Q x + c^T x \\
\text{s.t.} \sum_{j=1}^{n} x_j \ &= \ 1 \\
x_j - y_j \ &\leq \ 0, \quad \forall\, j \\
\sum_{j=1}^{n} y_j \ &\leq \ k \\
& x \geq 0, \quad y \in \{0,1\}^n
\end{aligned}
$$

Applying exclude-and-cut.

$$\mathcal{F} \doteq \left\{ x \in \Delta^{n-1} \; : \; \|x\|_0 \leq k \right\}$$

Applying exclude-and-cut.

$$\mathcal{F} \doteq \left\{ x \in \Delta^{n-1} \ : \ \|x\|_0 \leq k \right\}$$

Lemma. Let $w \in \Delta^{n-1}$. Then $\min\{\|y - \omega\|^2 \ : \ y \in \mathcal{F}\} = \rho(\omega)$,

$$\rho(\omega) \ \doteq \ \frac{(1 - \sum_{j \notin X} \omega_j)^2}{K} \ + \ \sum_{j \in X} \omega_j^2,$$

$X \subseteq \{1, \ldots, n\}$ is the set of indices of the $n - K$ smallest ω_j.

Applying exclude-and-cut.

$$\mathcal{F} \doteq \left\{ x \in \Delta^{n-1} \,:\, \|x\|_0 \leq k \right\}$$

Lemma. Let $w \in \Delta^{n-1}$. Then $\min\{\|y-\omega\|^2 \,:\, y \in \mathcal{F}\} = \rho(\omega)$,

$$\rho(\omega) \doteq \frac{(1 - \sum_{j \notin X} \omega_j)^2}{K} + \sum_{j \in X} \omega_j^2,$$

$X \subseteq \{1, \ldots, n\}$ is the set of indices of the $n-K$ smallest ω_j.

$$\min \left\{ z \,:\, z \geq x^T Q x + c^T x,\ x \in \mathcal{F} \right\}$$

Applying exclude-and-cut.

$$\mathcal{F} \doteq \left\{ x \in \Delta^{n-1} \; : \; \|x\|_0 \leq k \right\}$$

Lemma. Let $w \in \Delta^{n-1}$. Then $\min\{\|y - \omega\|^2 \; : \; y \in \mathcal{F}\} = \rho(\omega)$,

$$\rho(\omega) \;\doteq\; \frac{(1 - \sum_{j \notin X} \omega_j)^2}{K} \;+\; \sum_{j \in X} \omega_j^2,$$

$X \subseteq \{1, \ldots, n\}$ is the set of indices of the $n - K$ smallest ω_j.

$$\min \; \{z \; : \; z \geq x^T Q x + c^T x, \; x \in \mathcal{F}\}$$

$$\rightarrow \min \; \{z \; : \; z \geq x^T Q x + c^T x, \; \|x - \omega\|^2 \geq \rho(\omega)\}$$

n	k	LFO-L	MIP-L	MIP-U	LFO-t (sec)	MIP-t (sec)	MIP nodes
100	20	0.0411	0.0005	0.0587	0.127	227	1011704
100	50	0.0108	0.0006	0.0314	0.102	222	1004975
100	20	0.0465	0.0009	0.1284	0.120	288	1008679
1000	100	9.1009	0.0010	18.2534	0.883	1012	246063
1000	100	10.0109	0.0048	87.8492	0.848	1004	208633
1000	70	13.5842	0.0011	32.0741	0.879	1000	176152
2000	100	9.5178	0.0003	26.8787	3.014	1086	34699
2000	90	10.6348	0.0003	32.2729	2.563	1019	14298
2000	80	12.0266	0.0003	33.8795	3.186	1015	152638