# Advanced Mixed Integer Programming Formulation Techniques

# Nonlinear MIP Formulations

Joey Huchette and Juan Pablo Vielma

Massachusetts Institute of Technology

Spring School of the International Symposium on Combinatorial Optimization (ISCO 2018)

Marrakesh, Morocco, April, 2018

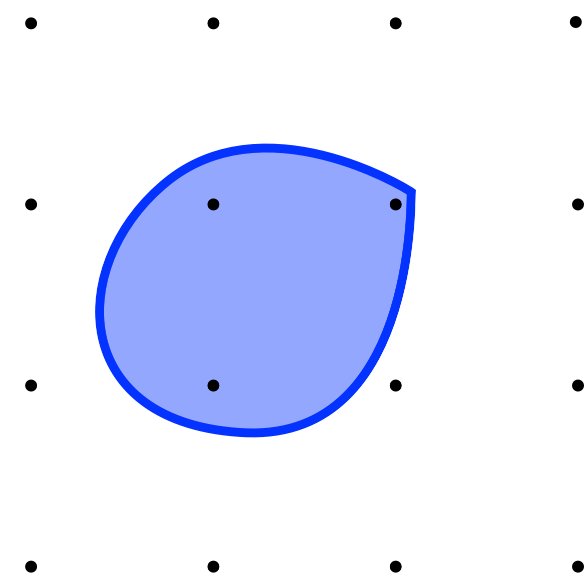
# (Nonlinear) Mixed Integer Programming (MIP)

$$\min \quad f(x)$$

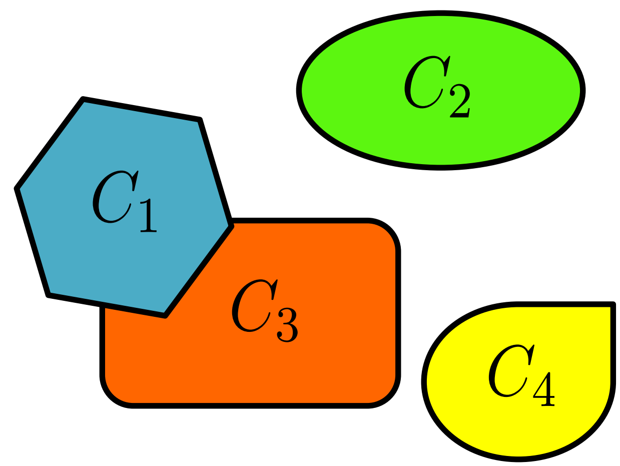
$$s.t.$$

$$x \in C$$

$$x_i \in \mathbb{Z} \quad i \in I$$



Mostly convex $f$ and $C$.



# “Convex” Nonlinear MIP Problem

$$
\begin{aligned}
\max \quad & g(x) + h(y) \\
s.t. & \\
& f_i(x, y) \leq b \qquad \forall i \in [k] \\
& x \in \mathbb{Z}^n \\
& y \in \mathbb{R}^m
\end{aligned}
$$

$$g,\ h \text{ and } f_i \text{ convex}$$

- Only non-convexities are due to integrality

# MI Second Order Cone Programming (MISOCP)

$$\max \quad c \cdot x + h \cdot y$$

$$s.t.$$

$$\left\| D^i x + E^i y \right\|_2 \leq f^i \cdot x + g^i \cdot y + c_0^i \quad \forall i \in \{1, \ldots, k\}$$

$$Ax + By \leq b$$

# Second Order Conic or Conic Quadratic Problems

- Problems using Euclidean norm:
  - e.g. Portfolio Optimization Problems

$$\max \quad \bar{a}x$$

$$s.t.$$

$$\left\|Q^{1/2}x\right\|_2 \leq \sigma$$

$$\sum_{j=1}^{n} x_j = 1, \quad x \in \mathbb{R}_+^n$$

$$x_j \leq z_j \quad \forall j \in [n]$$

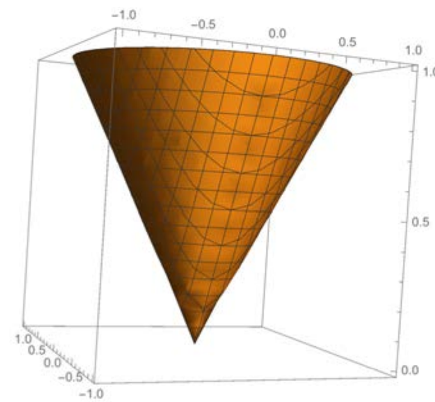
$$\sum_{j=1}^{n} z_j \leq K, \quad z \in \{0,1\}^n$$

- $\bar{a}$ expected returns.
- $Q^{1/2}$ square root of covariance matrix.
- $K$ maximum number of assets.
- $\sigma$ maximum risk.

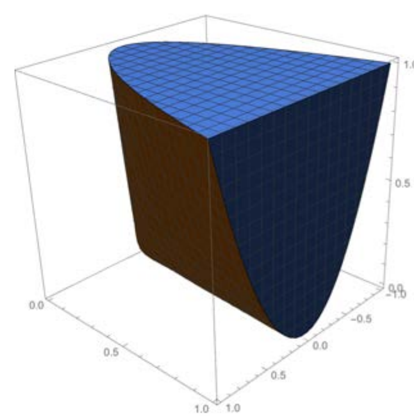
# Conic Quadratic or Second Order Cone

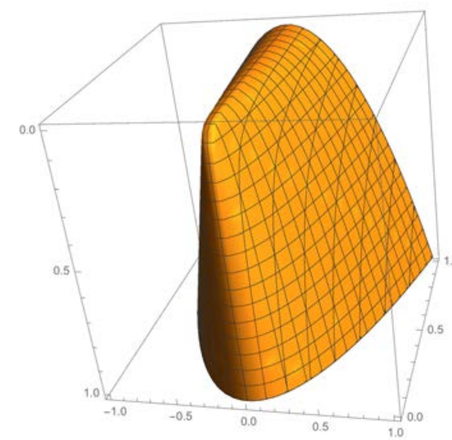
$$\left\|D^i x + E^i y\right\|_2 \leq f^i \cdot x + g^i \cdot y + c_0^i \quad \forall i \in \{1, \ldots, k\}$$

- Or linear inequalities +



  - $\displaystyle\sum_{i=1}^{m-1} y_i^2 \leq y_m^2, \quad y_m \geq 0$





  - $y_1^2 \leq y_2 y_3, \quad y_2,\ y_3 \geq 0$

# Mixed Integer Conic Programming

- $\mathcal{K}$
  - Convex cone = closed under sum and scaling
  - Closed (usually pointed and maybe more)

$$
\begin{aligned}
P: \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} \\
& \text{s.t.} \\
& \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \in \mathcal{K}^c \qquad \boldsymbol{x} \in \mathcal{K}^v \\
& \qquad x_i \in \mathbb{Z} \quad \forall x \in I
\end{aligned}
$$

# Conic Duality

- Cone dual to $\mathcal{K}$
  - $\mathcal{K}^* = \{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{x}'\boldsymbol{y} \geq 0, \ \forall \boldsymbol{x} \in \mathcal{K}\}$
  - $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_K$ $\qquad \mathcal{K}^* = \mathcal{K}_1^* \times \cdots \times \mathcal{K}_K^*$

$P:$ $\quad \min \quad \boldsymbol{c}'\boldsymbol{x}$

s.t.

$\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \in \mathcal{K}^c \qquad \boldsymbol{x} \in \mathcal{K}^v$

$D:$ $\quad \max \quad -\boldsymbol{b}'\boldsymbol{y}$

s.t.

$\boldsymbol{c} + \boldsymbol{A}'\boldsymbol{y} \in \mathcal{K}^{v*} \qquad \boldsymbol{y} \in \mathcal{K}^{c*}$

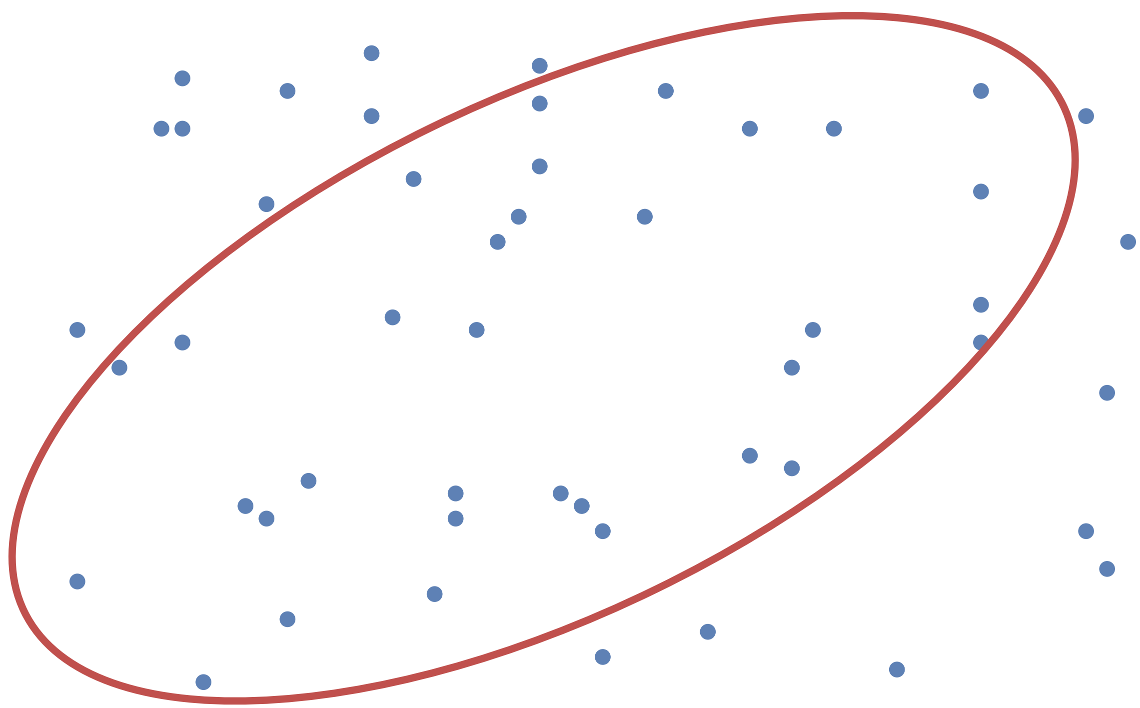
# Some Standard Cones

- Free and zero cones: `:Free` $= \mathbb{R}^n$, `:Zero` $= \{\mathbf{0}\}$
- Orthant cones: `:NonNeg` $= \mathbb{R}^n_+$, `:NonPos` $= \mathbb{R}^n_-$
- Second-order (Lorentz) cone `:SOC`
  $= \{(t, \boldsymbol{v}) \in \mathbb{R} \times \mathbb{R}^{n-1} : \|\boldsymbol{v}\|_2 \le t\}$
- Rotated second-order cone `:SOCRotated`
  $= \{(w, t, \boldsymbol{v}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-2} : \|\boldsymbol{v}\|_2^2 \le 2tw, t \ge 0, w \ge 0\}$
- Positive semidefinite cone `:SDP` $= \{\boldsymbol{V} \in \mathbb{S}^n : \boldsymbol{V} \succeq \mathbf{0}\}$
- Exponential cone `:ExpPrimal`
  $= \mathrm{cl}\{(r, s, t) \in \mathbb{R}^3 : s > 0, s \exp(r/s) \le t\}$
- Dual exponential cone `:ExpDual`
  $= \mathrm{cl}\{(u, v, w) \in \mathbb{R}^3 : u < 0, w \ge 0, v \ge -u \log(-u/w) + u\}$

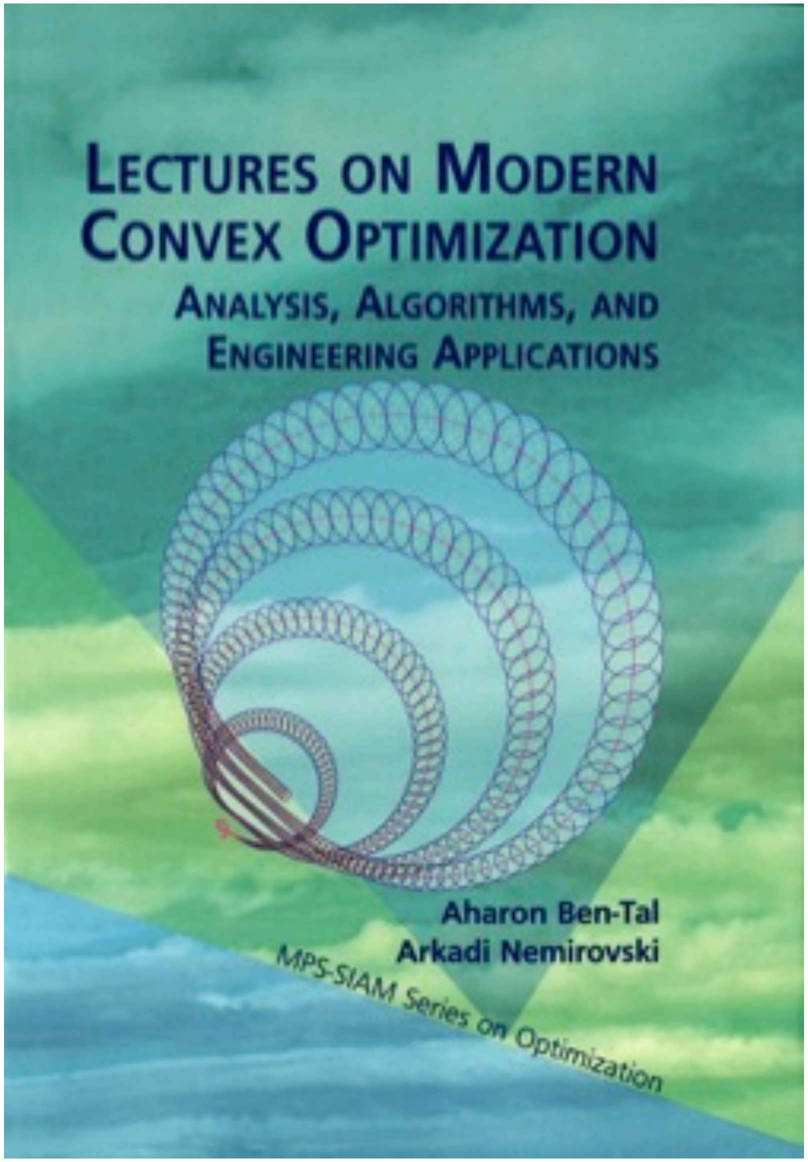
Cones not supported by conic solvers include the power cone for $\alpha \in (0, 1)$: $\{(r, s, t) \in \mathbb{R}^3 : |t| \le r^\alpha s^{1-\alpha}, r \ge 0, s \ge 0\}$

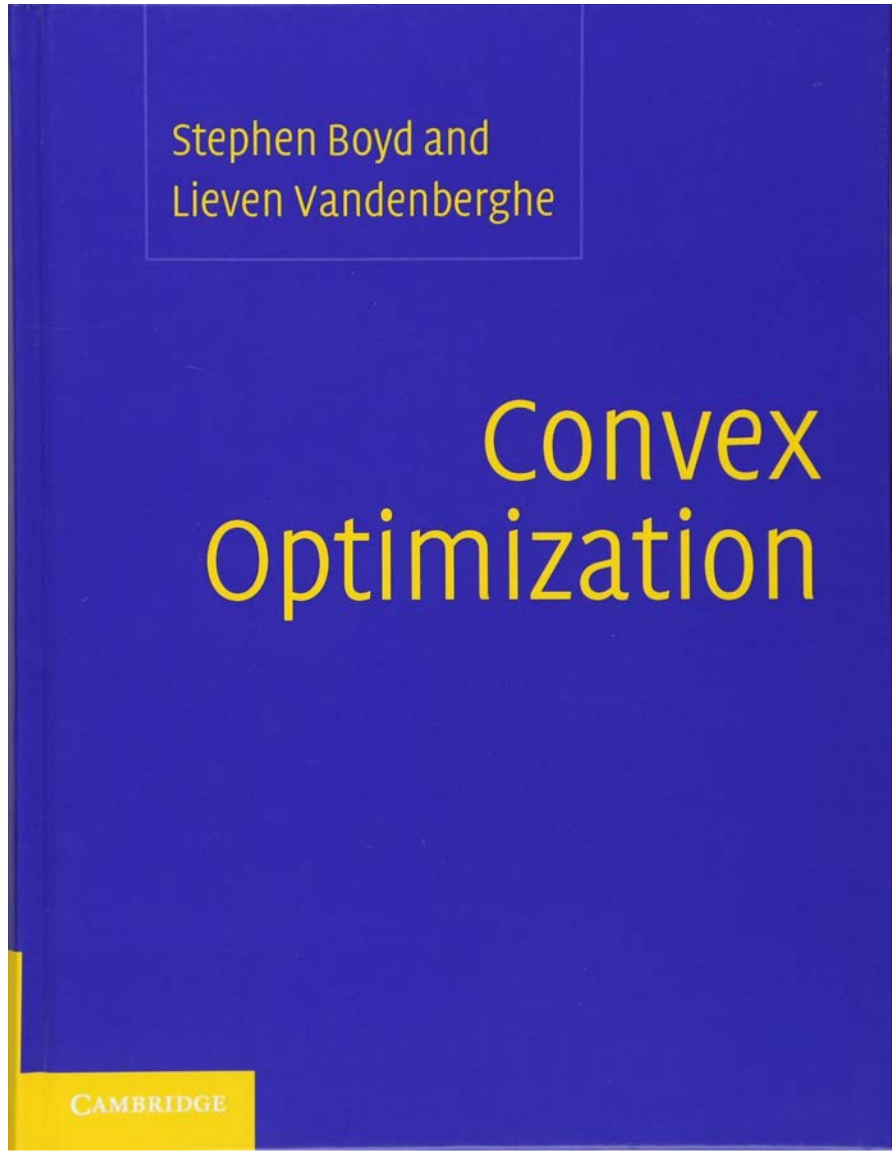
# Also Nonlinear Mixed Integer Programming (MIP)

- Example: Find minimum volume ellipsoid that contains 90% of data points



# Minimum Volume Ellipsoid = SDP



# Nonlinear MIP B&B Algorithms

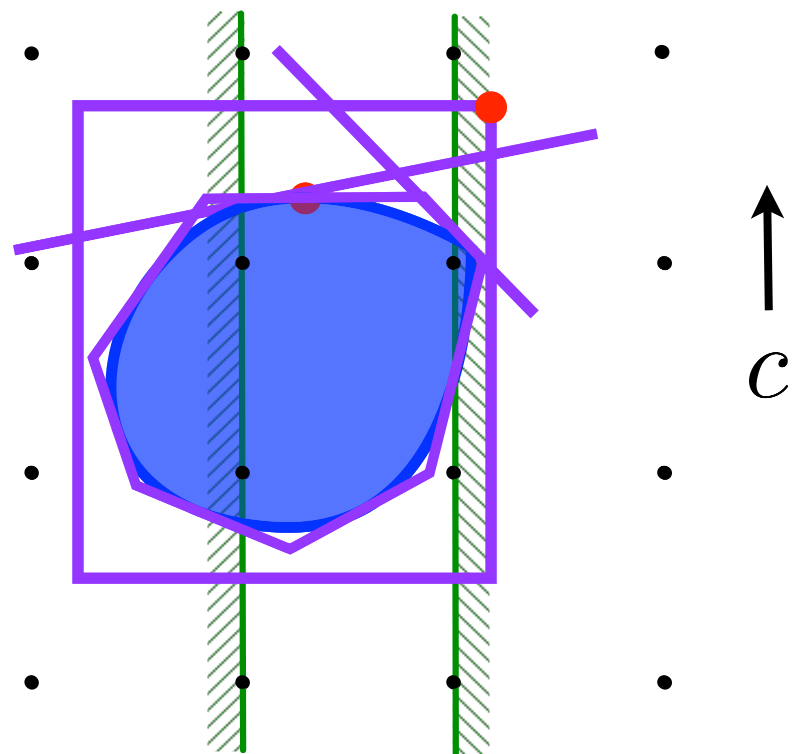
- NLP (QCP) Based B&B
- (Dynamic) LP Based B&B
  - Few cuts = high speed.
  - Possible slow convergence.
- Lifted LP B&B
  - Extended or Lifted relaxation.
  - Static relaxation
    - Mimic NLP B&B.
  - Dynamic relaxation
    - Standard LP B&B

$$\max \quad \sum_{i=1}^{n} c_i x_i$$

$$s.t. \quad Ax + Dz \leq b,$$

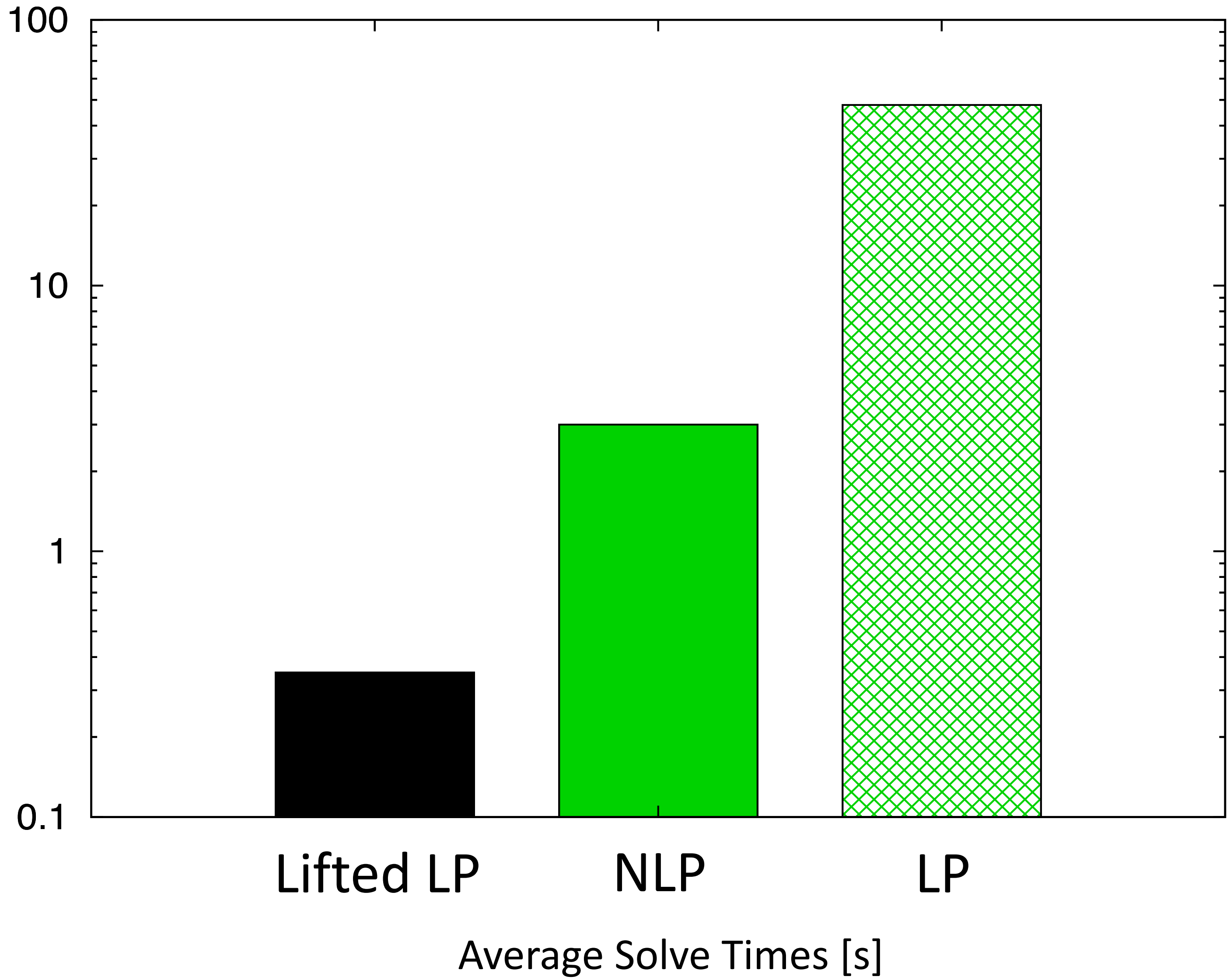
$$g_i(x) \leq 0,\ i \in I, \quad x \in \mathbb{Z}^n$$

$$x \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$$



# LP v/s NLP B&B for CPLEX v11 for n = 20 and 30

- Results from V., Ahmed and  Nemhauser 2008.



# Dynamic Lifted LP for Separable Problems

- Motivating example from Hijazi et al. ‘14

$$F^n := \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \left( x_i - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$

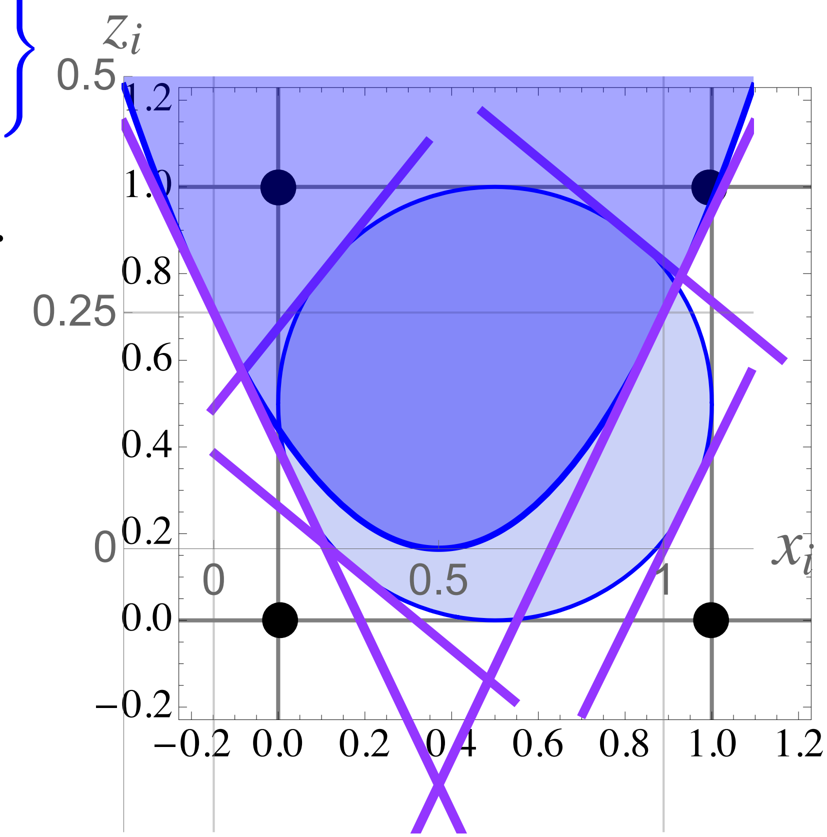
Showing $F^n \cap \mathbb{Z}^n = \emptyset$ requires $2^n$ cuts.

Extended formulation of $F^n$:

$$\left( x_i - \frac{1}{2} \right)^2 \leq z_i \qquad \forall i \in [n]$$

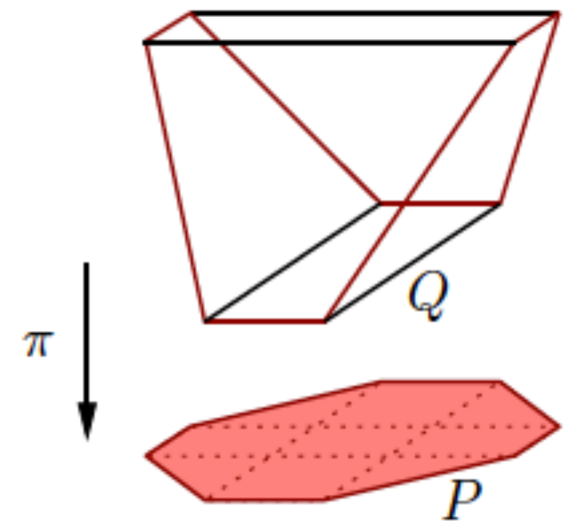
$$\sum_{i=1}^{n} z_i \leq \frac{n-1}{4}$$

$B^n \cap \mathbb{Z}^n = \emptyset$ with only $2n$ cuts on extended formulation.



# Lifted or Extended Approximations

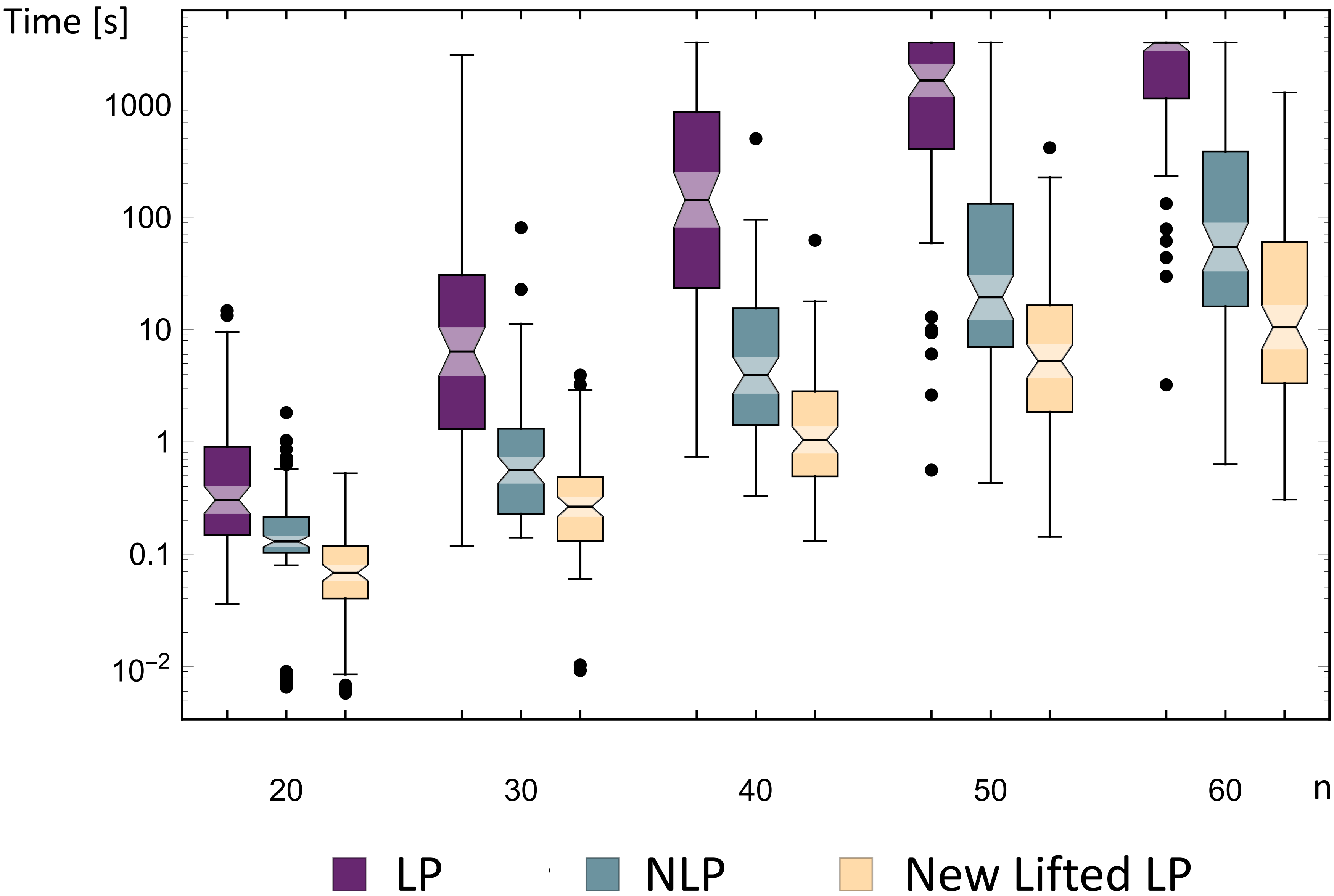
- Projection = multiply constraints.
- V., A. and N. 2008:
  - Extremely accurate, but static and complex approximation by Ben-Tal and Nemirovski
- V., Dunning, Huchette and Lubin 2016: Simple, dynamic and good approximation:



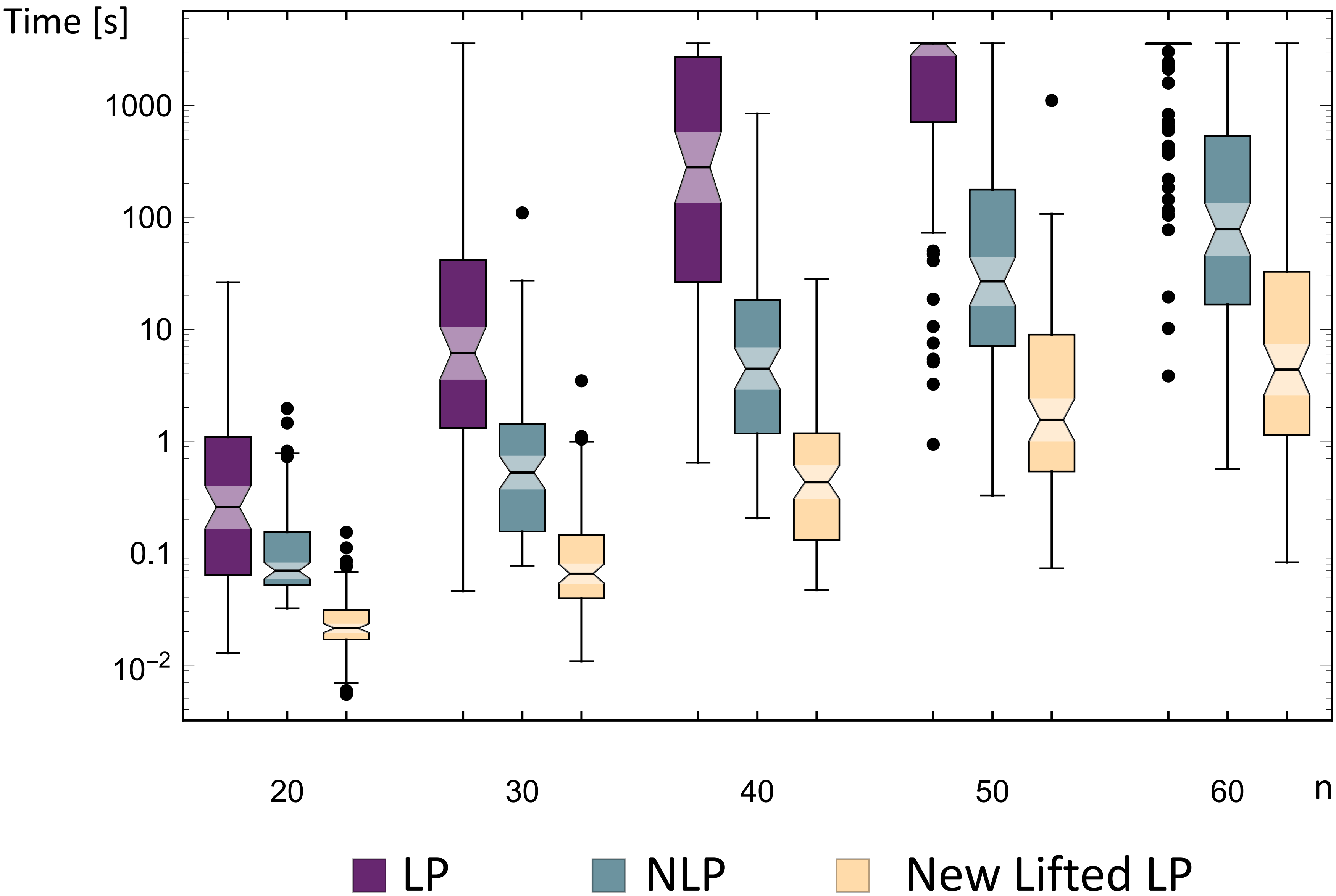

$$\|y\|_2 \leq y_0 \quad \longrightarrow \quad \begin{array}{l} y_i^2 \leq z_i \cdot y_0 \quad \forall i \in [n] \\ \sum_{i=1}^{n} z_i \leq y_0 \end{array}$$

Image from Lipton and Regan, https://rjlipton.wordpress.com

# CPLEX v12.6 for n = 20, 30, 40, 50 and 60



# Gurobi v5.6.3 for n = 20, 30, 40, 50 and 60



# All Major Solvers Now Implement Lifted LP

julia / JuMP

- First Talks:
  - SIAM Optimization (SIOPT), May 2014 ≈ two weeks coding.
  - IBM Thomas J. Watson Research Center, December 2014.
- Paper in arxive, May 28, 2015.

Two weeks!

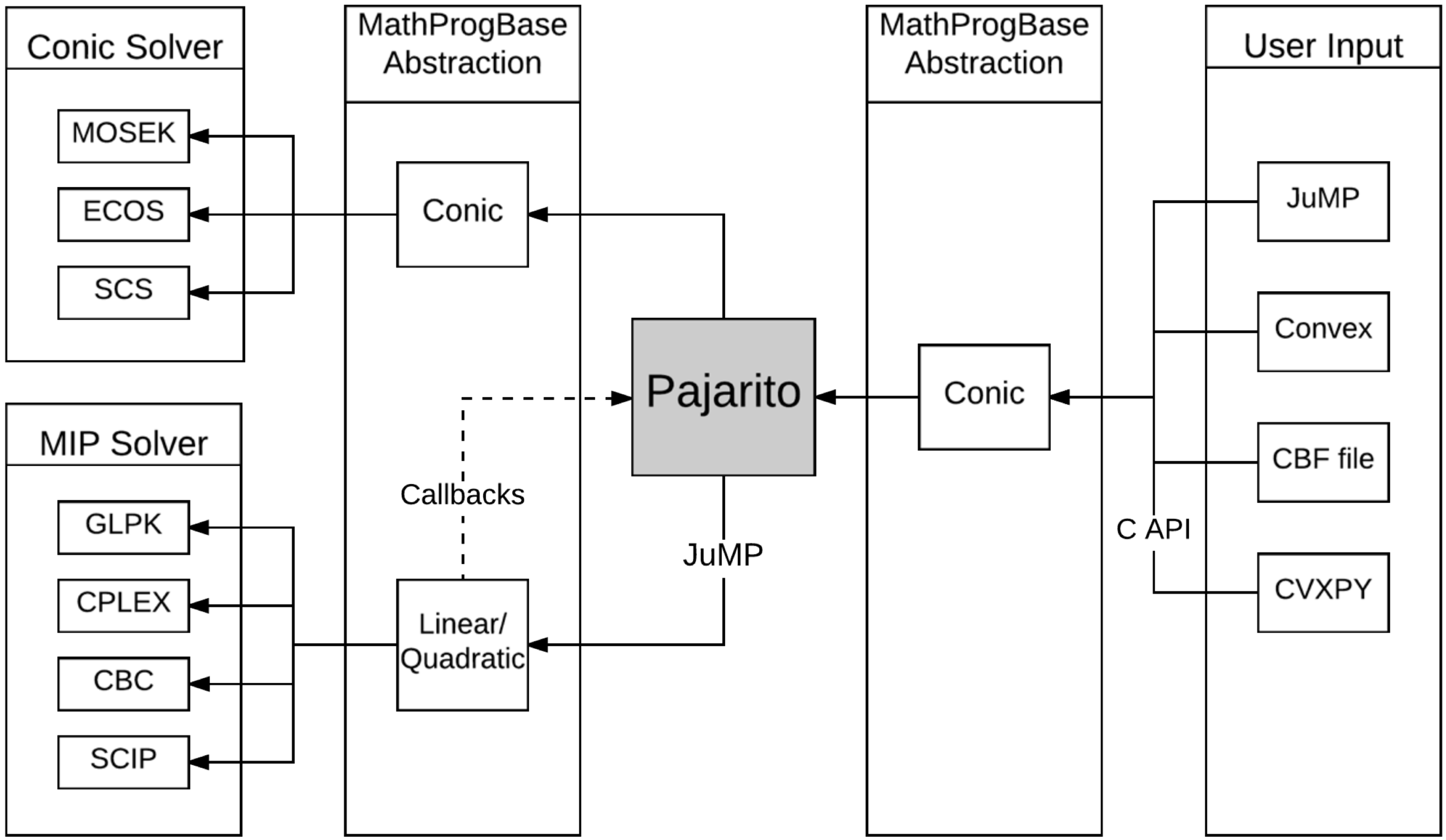
- IBM CPLEX v12.6.2, June 12, 2015.
- GUROBI OPTIMIZATION v6.5, October 2015.
- optimized by Xpress FICO v8.0, May 2016.
- SCIP v4.0, March 2017.

# However... We Can Sill Beat CPLEX!

- julia / JuMP-based solver Pajarito
  - Lubin, Yamangil, Bent and V. ’16 and Coey, Lubin and V. ’17.


| solver | termination status counts | | | | time(s) |
|---|---|---|---|---|---|
| | conv | wrong | not conv | limit | |
| SCIP | 78 | 1 | 0 | 41 | 43.36 |
| CPLEX | 96 | 3 | 5 | 16 | 14.30 |
| Paj-iter | 96 | 1 | 0 | 23 | 38.70 |
| Paj-MSD | 101 | 0 | 0 | 19 | 18.12 |

# Flexible Architecture Thanks to Julia-Opt Stack

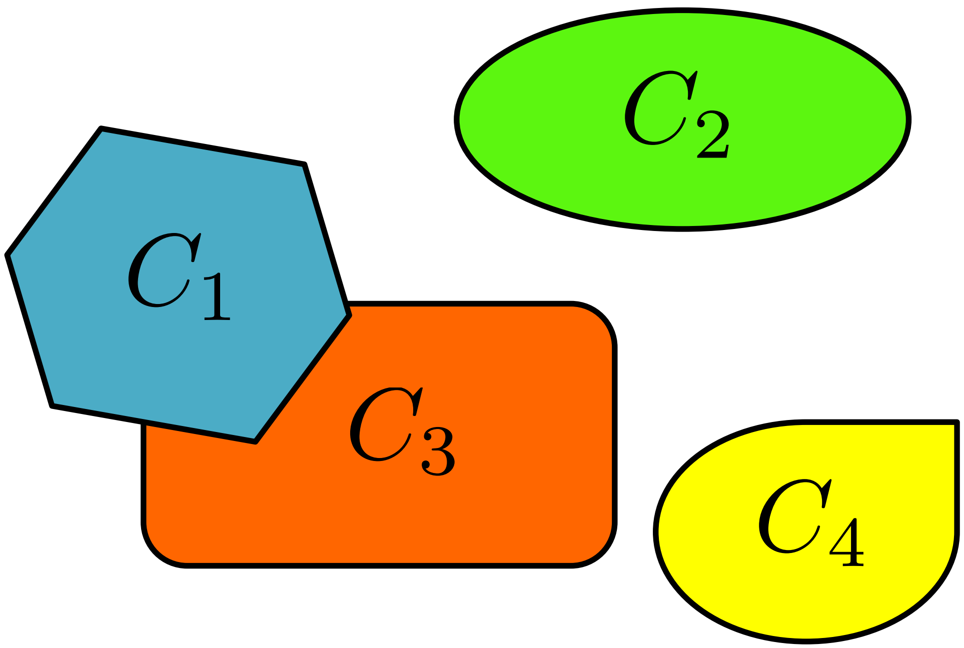



- Fastest Open Source MISOCP Solver!
- Pajarito can also solve MISDPs and MI-“EXP”

# Nonlinear Mixed 0-1 Integer Formulations

- Modeling Finite Alternatives = Unions of Convex Sets

$$x \in \bigcup_{i=1}^{n} C_i \subseteq \mathbb{R}^d$$



# Extended and Non-Extended Formulations for $\bigcup_{i=1}^{n} C_i$

$$C_i = \{x \in \mathbb{R}^d \ : \ f_i(x) \leq 0\}$$

| Extended | Non-Extended |
|---|---|
| $\tilde{f}_i\left(x^i, y_i\right) \leq 0 \quad \forall i \in [n]$ <br> $\sum_{i=1}^{n} x^i = x$ <br> $\sum_{i=1}^{n} y_i = 1$ <br> $y \in \{0,1\}^n$ <br> $x,\ x^i \in \mathbb{R}^d \quad \forall i \in [n]$ | $f_i\left(x\right) \leq M_i\left(1 - y_i\right) \quad \forall i \in [n]$ <br> $\sum_{i=1}^{n} y_i = 1$ <br> $y \in \{0,1\}^n$ <br> $x \in \mathbb{R}^d \quad \forall i \in [n]$ |
| Strong, but large | Small, but weak? |

# Extended Formulations: Perspective "v/s" Cones

- e.g. Ceria and Soares '99

$$C_i = \{x \in \mathbb{R}^d \ : \ f_i(x) \leq 0\}$$

$$\tilde{f}(x,y) = \begin{cases} y f(x/y) & \text{if } y > 0 \\ \lim_{\alpha \downarrow 0} \alpha f(x' - x + x/\alpha) & \text{if } y = 0 \\ +\infty & \text{if } y < 0 \end{cases}$$

$$\boxed{\begin{array}{lr} \tilde{f}_i\left(x^i, y_i\right) \leq 0 & \forall i \in [n] \\ \sum_{i=1}^n x^i = x & \\ \sum_{i=1}^n y_i = 1 & \\ y \in \{0,1\}^n & \\ x,\ x^i \in \mathbb{R}^d & \forall i \in [n] \end{array}}$$

- e.g. Ben-tal and Nemirovski '01, Helton and Nie '09

$$C_i = \left\{ x \in \mathbb{R}^d \ : \ \begin{array}{r} \exists u \in \mathbb{R}^{p_i} \ \text{s.t.} \\ A^i x + D^i u - b \in K^i \end{array} \right\}$$

$K^i$ closed convex cone

$$\boxed{\begin{array}{lr} A^i x^i + D^i u^i - b y_i \in K^i & \forall i \in [n] \\ \sum_{i=1}^n x^i = x & \\ \sum_{i=1}^n y_i = 1 & \\ y \in \{0,1\}^n & \\ x,\ x^i \in \mathbb{R}^d & \forall i \in [n] \\ u^i \in \mathbb{R}^{p_i} & \forall i \in [n] \end{array}}$$

- Both formulations are ideal (extreme points of continuous relaxation satisfy integrality constraints)

# Cones Can Mitigate Unintended Numerical Issues

- Let $C_i = \{x \in \mathbb{R}^2 \ : \ f_i(x) \leq 0\}$
  where $f_i(x) = x_1^2 - x_2 - 1$

$$\tilde{f}_i(x,y) = \begin{cases} y(x_1\boxed{/y)}^2 - x_2 - y & \text{if } y > 0 \\ -x_2 & \text{if } y = x_1 = 0 \\ +\infty & \text{if } o.w. \end{cases}$$

- Conic (SOCP) representation

$$C_i = \left\{x \in \mathbb{R}^2 \ : \ \sqrt{x_2^2 + 4x_1^2} \leq 2 + x_2\right\}$$

$$\sqrt{\left(x_2^i\right)^2 + 4\left(x_1^i\right)^2} \leq 2y_i + x_2$$

# Very Stable and Fast Conic Solvers

- Matrices grow quadratically so you can easily run out of memory.
- If it fits in memory you can probably solve relaxation

mosek

- And then the MIP

# Advanced Convex MINLP Formulations

# A Classical Strong Formulation for $\bigcup_{i=1}^{k} C_i$

$$C_i = \left\{x \in \mathbb{R}^n \,:\, A^i x \preceq_i b^i\right\}, \quad C_i^\infty = C_j^\infty$$

$$A^i x^i \preceq_i b^i z_i, \quad \forall i \in [k]$$

$$\sum_{i=1}^{k} x^i = x,$$

$$\sum_{i=1}^{k} z_i = 1, \quad z \in \{0,1\}^k$$

$$x,\ x^i \in \mathbb{R}^n, \quad \forall i \in [k]$$

- Auxiliary continuous variables are copies of original variables
  - $y = \left(x^i\right)_{i=1}^{k}$
- "Ideal" Formulation Strength:
  - Extreme points of continuous relaxation satisfy **integrality constraints** on $z$
  - **Variable copies** crucial here, but **slow down computations** (usually worse than Big-M)

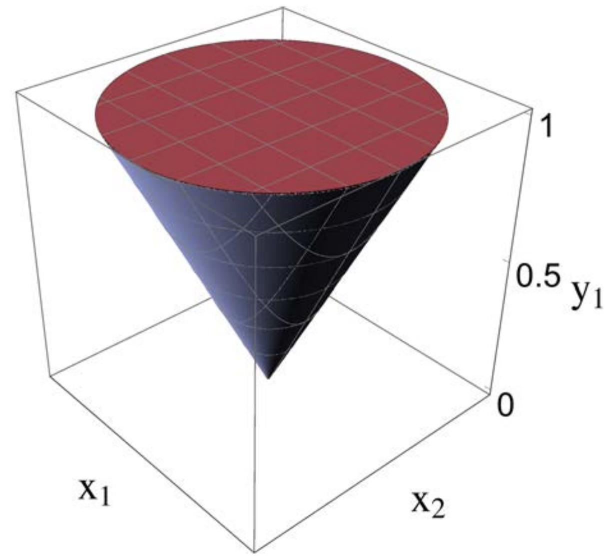
- Balas, Jeroslow and Lowe (Polyhedral), Ben-tal, Nemirovski, Helton, Nie (Conic)

# Generic Geometric Formulation = Gauge Functions

- For $C$ such that $\mathbf{0} \in \operatorname{int}(C)$ let:

  $$\gamma_C(x) := \inf\{\lambda > 0 \, : \, x \in \lambda C\}$$

  $$\operatorname{epi}(\gamma_C) =$$




- If $b^i \in \operatorname{int}(C_i)$ then ideal formulation:

$$\gamma_{C^i - \{b^i\}}\left(x^i - z_i b^i\right) \leq z_i \qquad \forall i \in [k]$$

$$\sum_{i=1}^{k} x^i = x$$

$$\sum_{i=1}^{k} z_i = 1$$

$$z \in \{0,1\}^k$$

$$x, \; x^i \in \mathbb{R}^n \qquad \forall i \in [k]$$

# Non-Extended and Big-M Formulations

- Gauge Big-M = no copies $x^i$ of original variables:

$$\gamma_{C^i - \{b^i\}}\left(x - \sum_{j=1}^n y_j \overline{M}^{i,j}\right) \leq \sum_{j=1}^n y_j \underline{M}^{i,j} \quad \forall i \in [n]$$

$$\sum_{i=1}^n y_i = 1$$

$$y \in \{0,1\}^n$$

- Can be stronger than standard big-M (e.g. $C_i = \{x \in \mathbb{R}^n \,:\, \|x\|_2 \leq r_i\}$ or $C_i = \left\{x \in \mathbb{R}^n \,:\, \|x\|_2^2 \leq r_i^2\right\}$ )
- But may not be ideal
- What about ideal non-extended formulations?

# Simple Non-Extended Ideal Formulation

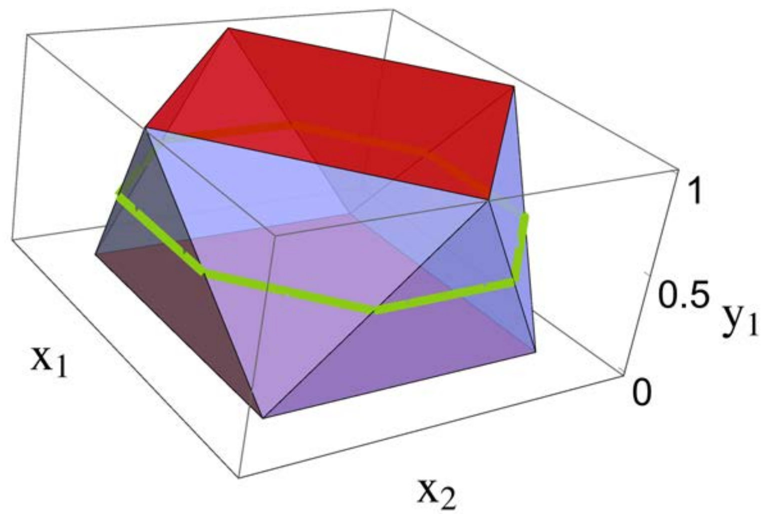
- Unions of (nearly) Homothetic Closed Convex Sets:

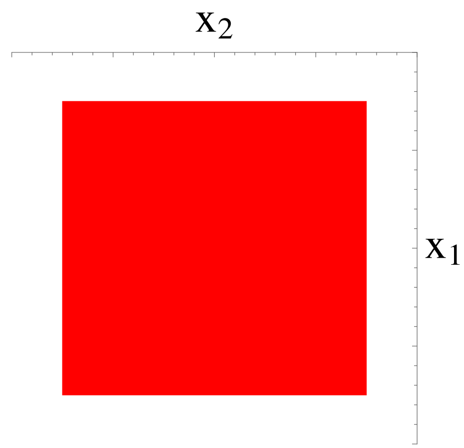
$$C_i = \lambda_i C + b^i + C_\infty$$


$$\gamma_C\left(x - \sum_{i=1}^n y_i b^i\right) \leq \sum_{i=1}^n \lambda_i y_i$$

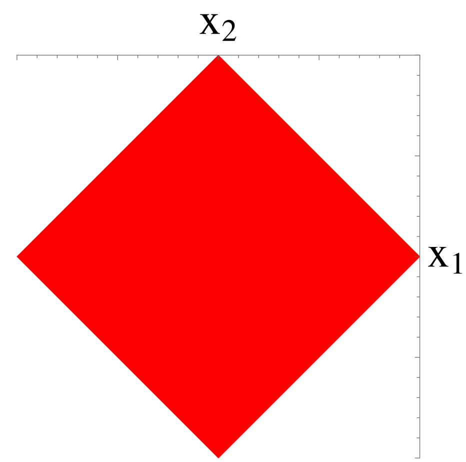
$$\sum_{i=1}^n y_i = 1,\ y \in \{0,1\}^n$$

# Unary Encoding, Minkowski Sum and Cayley Trick

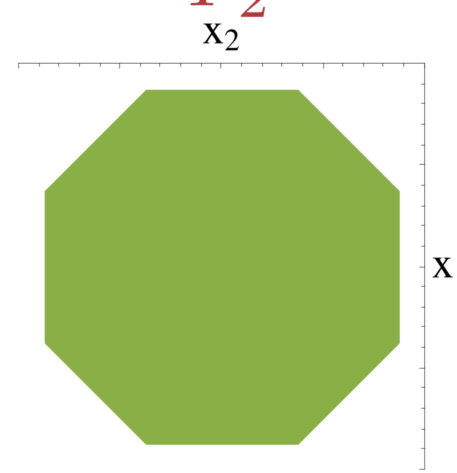



$P_1$




$P_2$




$$Q \cap \left(\mathbb{R}^2 \times \{0.5\}\right) \equiv P_1 + P_2 =$$

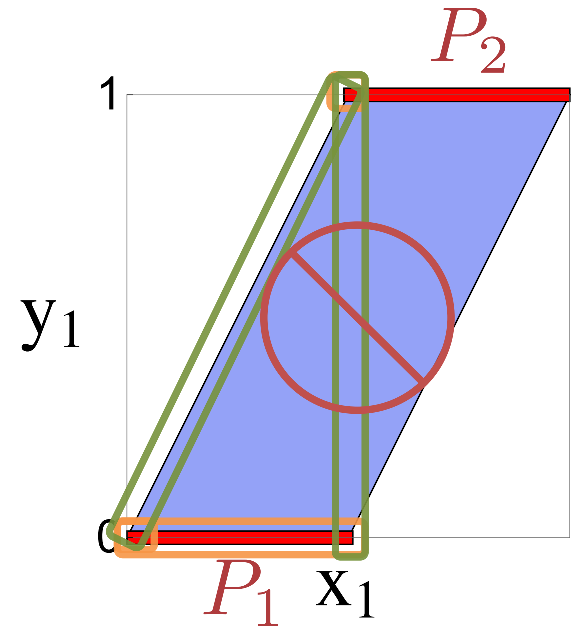
$$H = \{e^i\}_{i=1}^n$$

$$Q\left(H\right) \cap \left(\mathbb{R}^d \times \left\{\tfrac{1}{n}\sum_{i=1}^n \mathbf{e}^i\right\}\right) \equiv \sum_{i=1}^n P_i$$

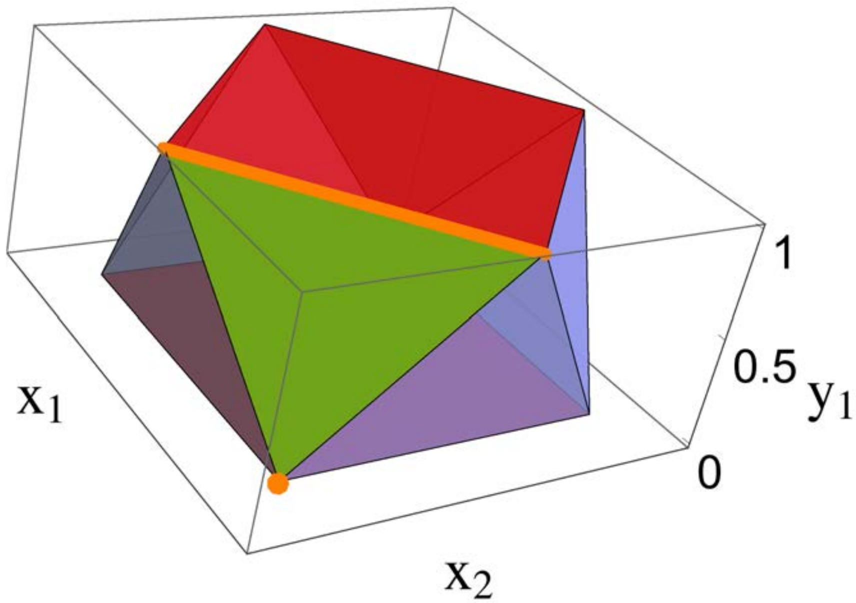
# Faces of Cayley Embedding

- Two types of facets (or faces):
  - $P_1 \times \{0\} \equiv y_i \geq 0$
  - $\mathrm{conv}\left(\left(F_1 \times 0\right) \cup \left(F_2 \times 1\right)\right)$

    $F_i$ proper face of $P_i$
  - Not all combinations of faces
  - Which ones are valid?



# Valid Combinations = Common Normals

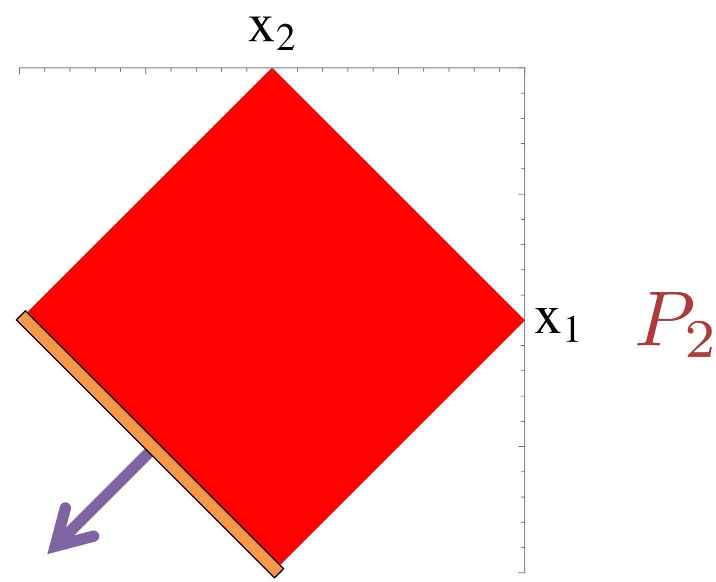


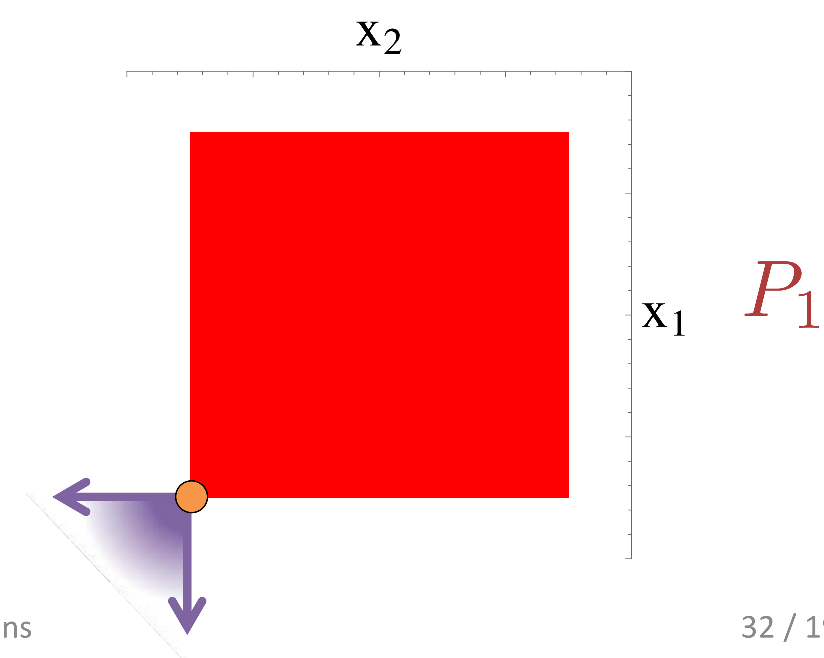

$$N\left(F_1\right) \cap N\left(F_2\right) \neq \emptyset$$

$$\Updownarrow$$

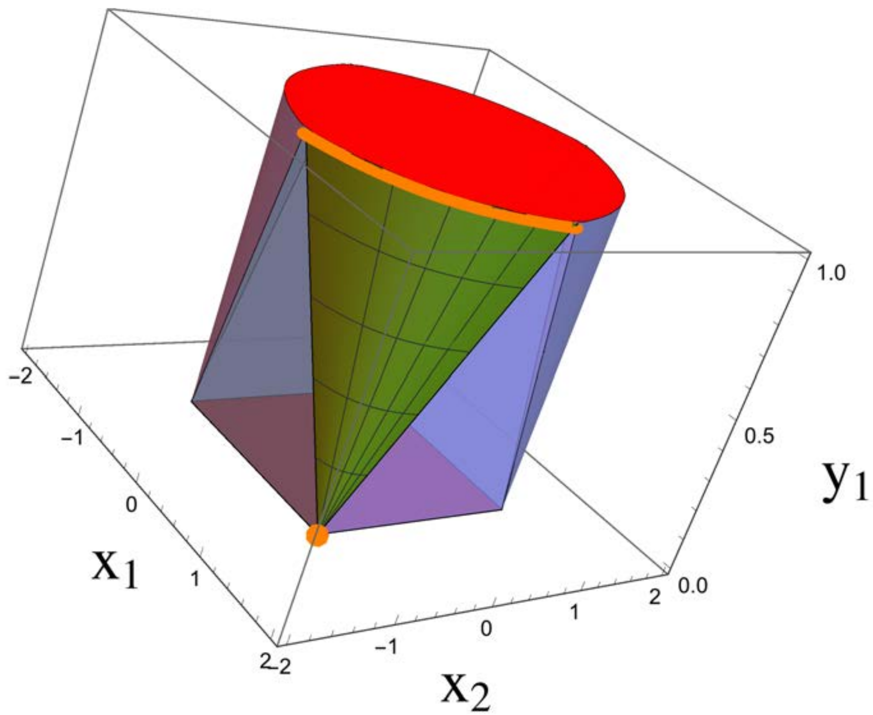
$$\operatorname{conv}\left(\left(F_1 \times 0\right) \cup \left(F_2 \times 1\right)\right)$$

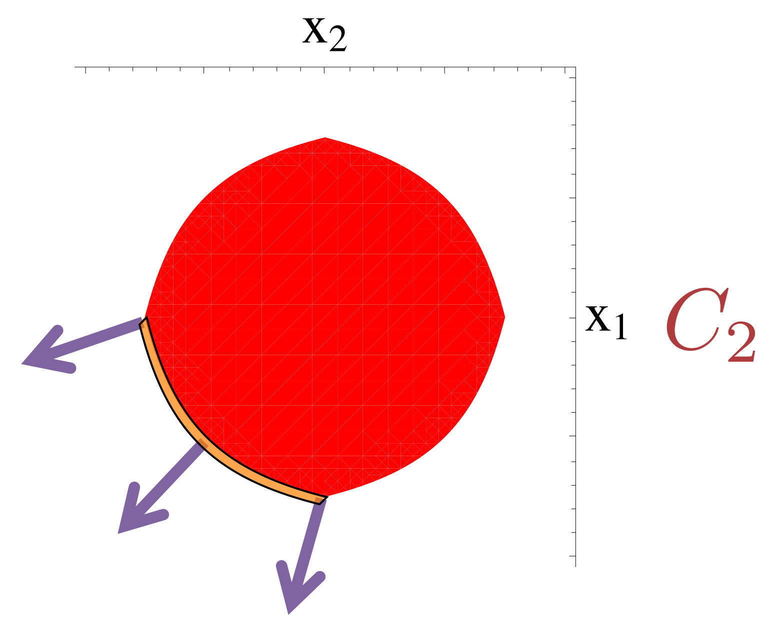
is face of $Q\left(H\right)$

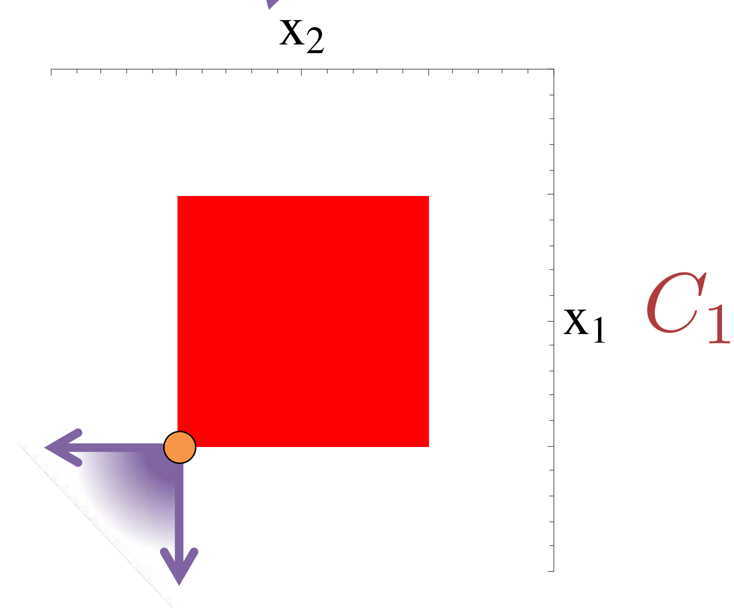


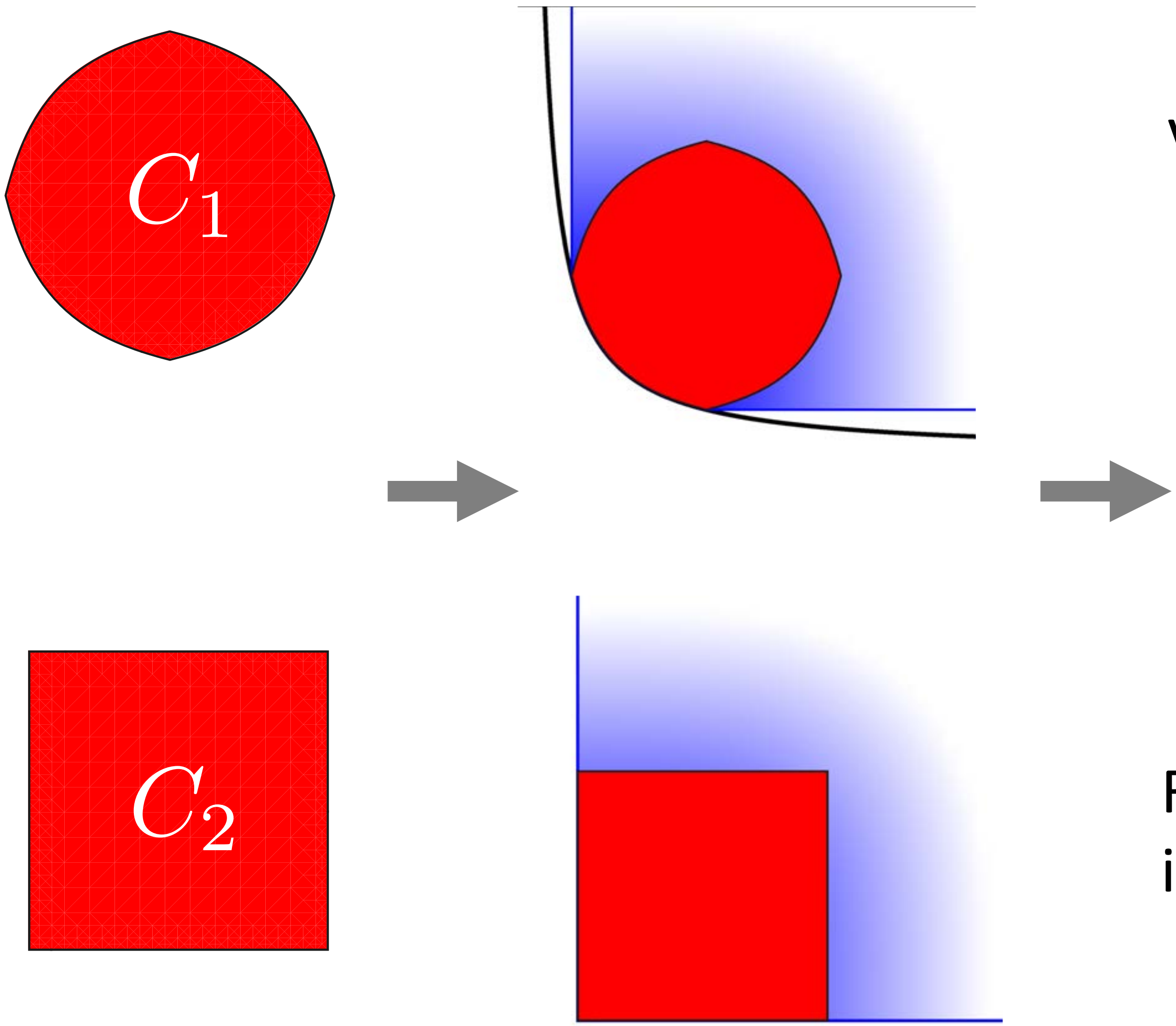
# Characterization Extends to Closed Convex Sets



# Sticking Homothetic Formulations Together



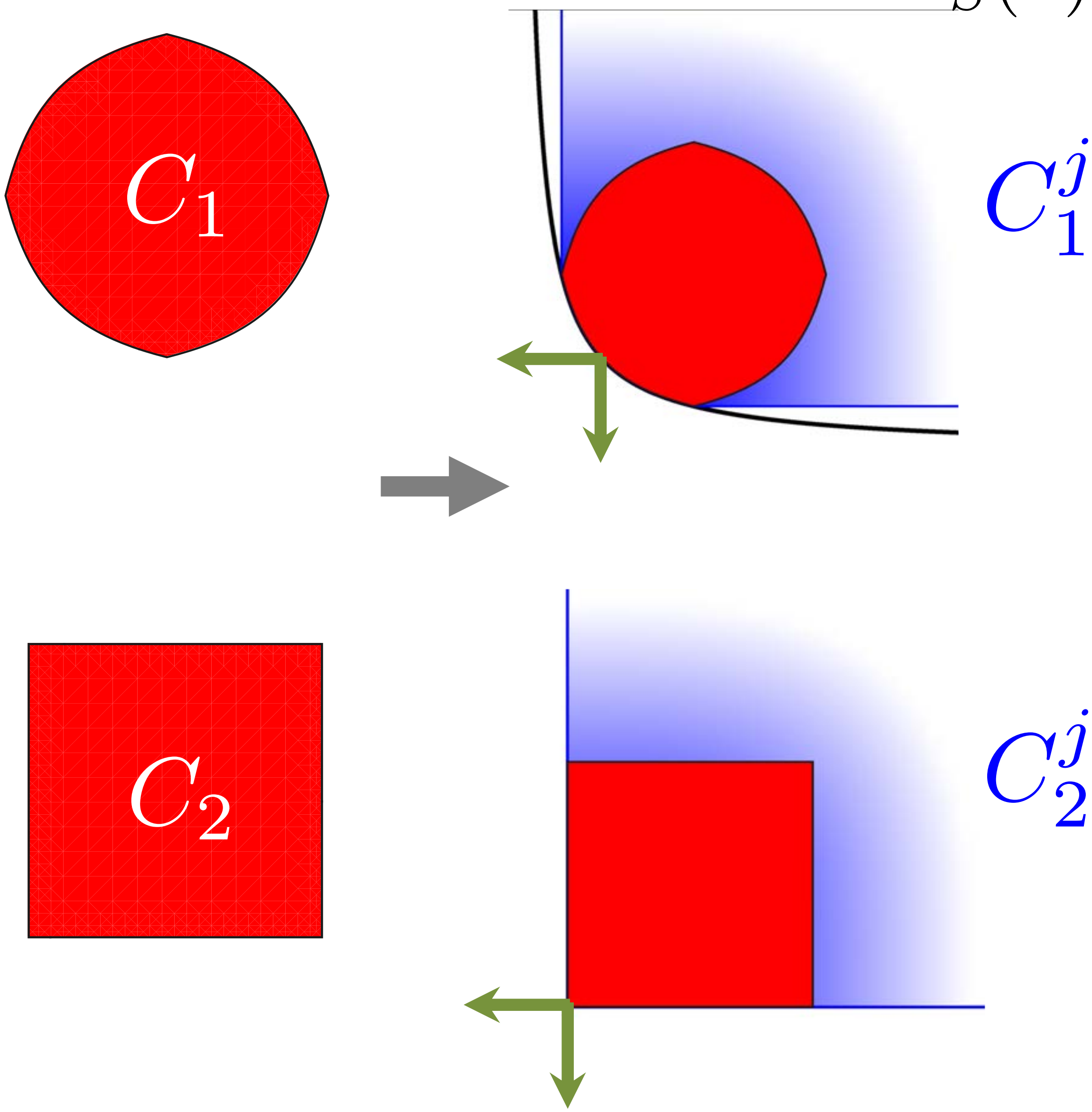

Valid, but not ideal!

Combine 4 homothetic formulations

Right relaxations yield ideal formulation

# Sufficient Conditions For Ideal Formulation

$$\sigma_S(u) := \sup\{u \cdot x \ : \ x \in S\}$$




$$\forall u \in \mathbb{R}^n \quad \exists j$$

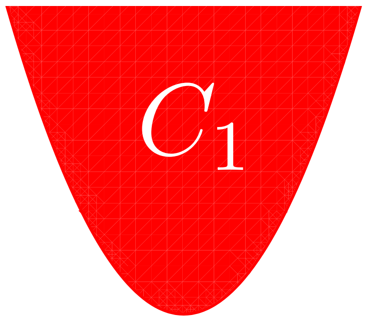
$$s.t.$$

$$\sigma_{C_i}(u) = \sigma_{C_i^j}(u)$$

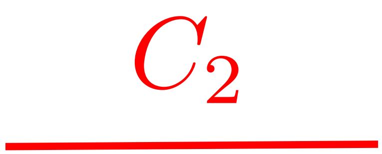
$$\forall i \in \{1, 2\}$$

*Similar* to “lifting” of e.g. Tawarmalani et al. ‘10

# May Need to “Find” Homothetic Constraints




$$x_1^2 \leq x_2 \leq 1$$

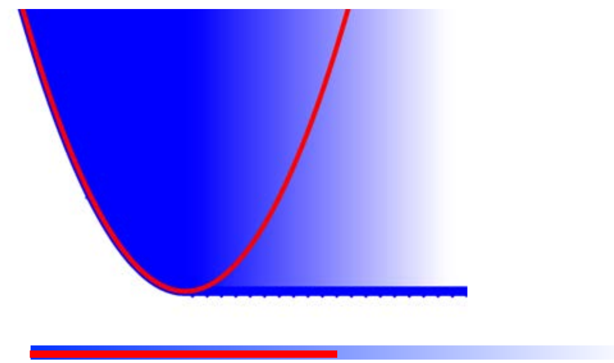



$$[-1, 1] \times 0$$

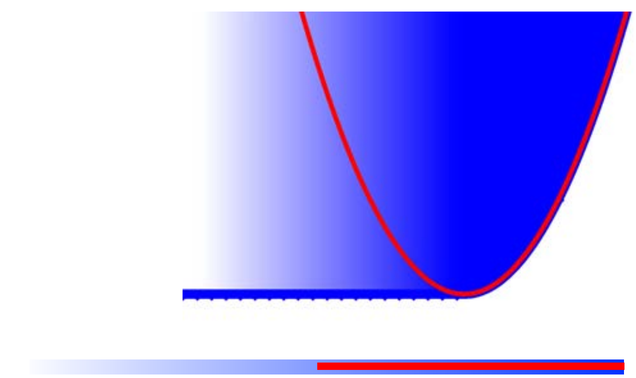
$$C_1 + (\mathbb{R}_+ \times \{0\}) :$$

$$(\max\{x_1, 0\})^2 \leq x_2 \leq 1$$

$$C_i + (\mathbb{R}_+ \times \{0\})$$

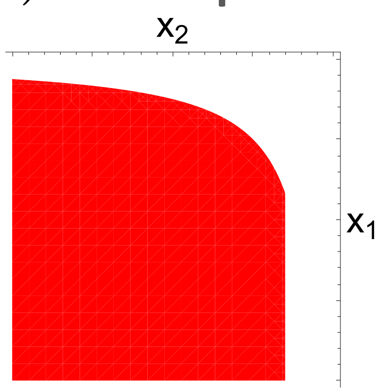


$$C_i + (\mathbb{R}_- \times \{0\})$$

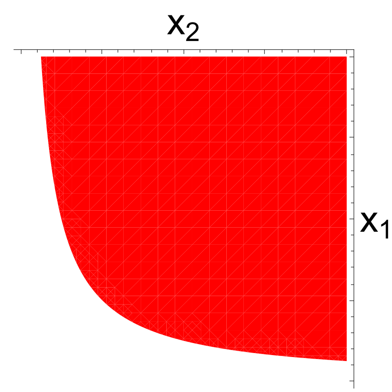


# Existing Small Ideal Formulations (Isotone Sets)

- Studied by Hijazi et al. ‘12 and Bonami et al. ’15 (n=1, 2):
  - $C_i = \{x \in \mathbb{R}^d \,:\, l^i \leq x \leq u^i, \quad f_i(x) \leq 0\}$
- $f_i(x)$ component-wise monotonous (i=1,2 opposite).




$C_1$




$C_2$

- Ideal Formulation

$$
\begin{aligned}
& y_1 l^1 + y_2 l^2 \leq x \leq y_1 u^1 + y_2 u^2 \\
& f_J^i(x, y) \leq 0 && \forall J \subseteq [d], i \in [2] \\
& y_1 + y_2 = 1 \\
& y_i \in \{0, 1\} && i \in [2]
\end{aligned}
$$

# Generalization and Simplification

- More than 2 sets (with general “opposite condition”).
- Generalization of the monotone/isotone condition (beyond affine transformation)
- Significantly smaller formulation: One non-linear constraint per set.

$$y_1 l^1 + y_2 l^2 \leq x \leq y_1 u^1 + y_2 u^2$$

$$\cancel{f_J^i(x, y) \leq 0 \qquad \forall J \subseteq [d], i \in [2]}$$

$$y_1 + y_2 = 1$$

$$y_i \in \{0, 1\} \qquad i \in [2]$$

$$\hat{f}^i(x, y) \leq 0 \quad \forall i \in [2]$$

# Details of Size Reduction

$$C_i = \{x \in \mathbb{R}^d \;:\; l^i \le x \le u^i, \quad f_i(x) \le 0\}$$

$$G_i = \{x \in \mathbb{R}^d \;:\; f_i(x) \le 0\}$$

- Original formulation:

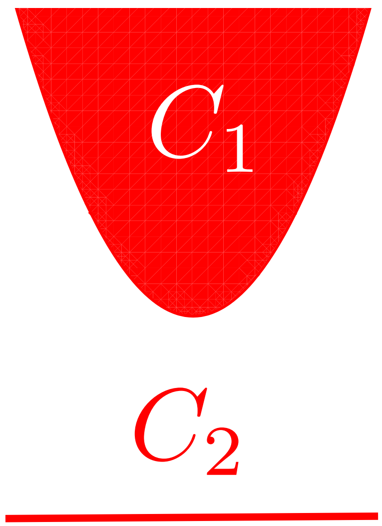
$$\gamma_{G_i}\left([x]_J\right) \le y_i, \boxed{\forall J \subseteq [d]} \qquad \left([x]_J\right)_j := \begin{cases} x_j & j \in J \\ 0 & o.w. \end{cases}$$

- Smaller formulation:

$$\gamma_{G_i}\left([x]^+\right) \le y_i \qquad \left([x]^+\right)_j := \max\{x_j, 0\}$$

  - max can cause representability issues.

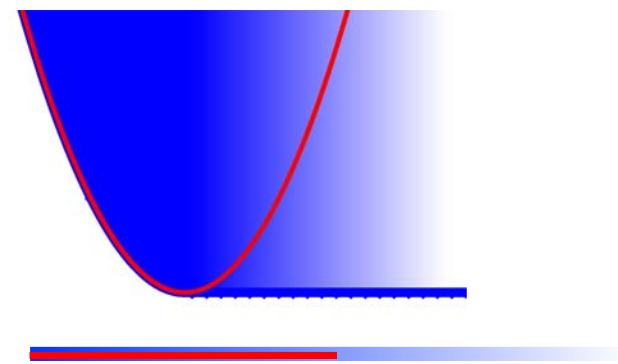
# Algebraic Representation Issues




$$x_1^2 \le x_2 \le 1$$

$$[-1, 1] \times 0$$

$$C_i + (\mathbb{R}_+ \times \{0\})$$



$$C_1 + (\mathbb{R}_+ \times \{0\}) : \ (\max\{x_1, 0\})^2 \le x_2 \le 1$$

- Non-basic semi-algebraic set contained in formulation.
- Finite polynomial inequalities requires max or auxiliary vars.