Disjunctive Conic Cuts

Ted Ralphs

Joint work with:
Pietro Belotti, Julio C. Góez, Imre Pólik, Tamás Terlaky

INFORMS Computing Society Conference
January 7, 2013

Agenda

Agenda

1. Introduction
2. DCCs for MISOCO
3. Computational Experience

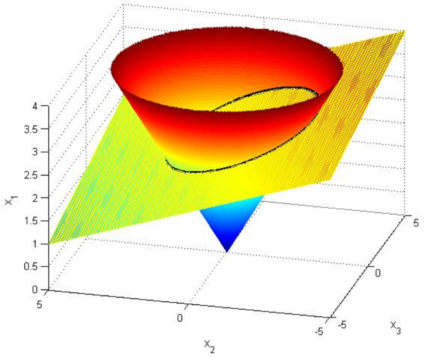
$$\begin{aligned} \text{min: } & c^T x \\ \text{s.t.: } & Ax = b \qquad\qquad \text{(SOCO)} \\ & x \in \mathcal{K} \end{aligned}$$

where

- $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$
- $x = (x^i, \dots, x^n)$
- $\mathcal{K} = \{\mathbb{L}^{n_i} \times \cdots \times \mathbb{L}^{n_k}\}$
- $\mathbb{L}^{n_i} = \{x^i | x_1^i \geq \|x_{2:n_i}^i\|\}$
- Rows of A are linearly independent

Example

$$
\begin{array}{lllll}
\text{min:} & x_1 & -2x_2 & +x_3 & \\
\text{s.t.:} & x_1 & -0.1x_2 & +0.2x_3 & = 2.5 \\
 & x_1 \geq \|(x_2, x_3)\| & & &
\end{array}
$$



Feasible set

Intersection of an affine space and a second order cone

- All points satisfying $Ax = b$ are in the set

$$\mathcal{H} := \{x \in \mathbb{R}^n \mid x = x_0 + Hz,\ \forall z \in \mathbb{R}^{n-m}\},$$

  where $Ax_0 = b$ and $H \in \mathbb{R}^{n \times n-m}$ is a basis for $\mathrm{Null}(A)$.
- There exist a matrix $P \in \mathbb{R}^{n-m \times n-m}$, $p \in \mathbb{R}^{n-m}$, $\rho \in \mathbb{R}$, s.t.

$$\mathcal{H} \cap \mathbb{L}^n \subset \{z \in \mathbb{R}^{n-m} \mid z^\top Pz + 2p^\top z + \rho \leq 0\},$$

  and P has at most one negative eigenvalue.
- The set $\mathcal{Q} = \{z \in \mathbb{R}^{n-m} \mid z^\top Pz + 2p^\top z + \rho \leq 0\}$ is a quadric and we denote it as (P, p, ρ).

Mixed Integer Second Order Cone Optimization

$$\begin{aligned} \text{min: } & c^T x \\ \text{s.t.: } & Ax = b \qquad\qquad \text{(MISOCO)} \\ & x \in \mathcal{K} \\ & x \in \mathbb{Z}^d \times \mathbb{R}^{n-d}, \end{aligned}$$

where

- $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$
- $x = (x^i, \dots, x^n)$
- $\mathcal{K} = \{\mathbb{L}^{n_i} \times \cdots \times \mathbb{L}^{n_k}\}$
- $\mathbb{L}^{n_i} = \{x^i | x^i_1 \geq \|x^i_{2:n_i}\|\}$
- Rows of A are linearly independent

Agenda

1 Introduction

2 DCCs for MISOCO

3 Computational Experience

Algorithmic Framework

- We propose an algorithm similar to a standard branch-and-cut algorithm.
  - Solve the continuous relaxation (a SOCO problem).
  - Identify a violated disjunction (fractional variable).
  - Either branch or generate a disjunctive constraint.
- Procedure for cut generation is similar to lift and project for mixed integer linear optimization (MILO) problems.
- The convex hull of the disjunctive set associated with a variable disjunction can be obtained by the addition of a single conic constraint.
- This constraint is easy to obtain.

Step 1: Solve the relaxed problem

Find the optimal solution x^*_{soco} for the continuous relaxation

$$\begin{array}{llllll} \text{min:} & 3x_1 & +2x_2 & +2x_3 & +x_4 & \\ \text{s.t.:} & 9x_1 & +x_2 & +x_3 & +x_4 & = 10 \\ & & (x_1, x_2, x_3, x_4) & & \in \mathbb{L}^4 & \\ & & & x_4 & \in \mathbb{Z}. & \end{array}$$

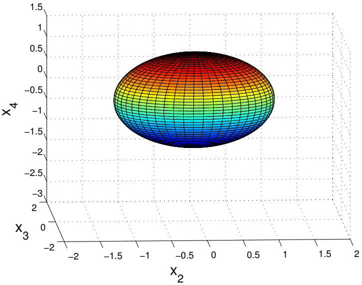
Relaxing the integrality constraint we get the optimal solution:

$$x^*_{soco} = (1.36, -0.91, -0.91, -0.45),$$

with and optimal objective value: $z^* = 0.00$.

Reformulation of the relaxed problem

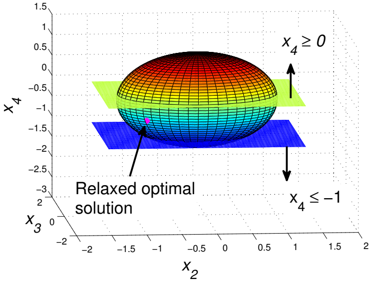
$$
\begin{aligned}
\text{min:} \quad & \frac{1}{3}\left(10 + 5x_2 + 5x_3 + 2x_4\right) \\
\text{s.t.:} \quad & \begin{bmatrix} x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} 8 & -\frac{1}{10} & -\frac{1}{10} \\ -\frac{1}{10} & 8 & -\frac{1}{10} \\ -\frac{1}{10} & -\frac{1}{10} & 8 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix} + 2\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix} - 10 \quad \leq 0 \\
& x_4 \in \mathbb{Z} \qquad .
\end{aligned}
$$



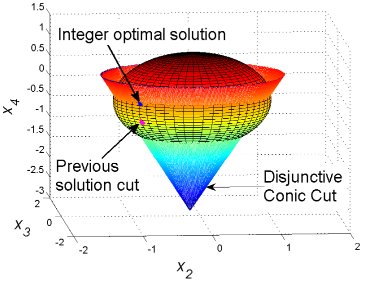
Feasible set of the reformulated problem

Step 2: Find a violated disjunction

The disjunction $x_4 \leq -1 \bigvee x_4 \geq 0$ is violated by x^*_{soco}



(A) Disjunction



(B) Disjunctive conic cut

Step 3: Apply the disjunction and Convexify

The constraints in red represent the disjunctive conic cut.

$$
\begin{array}{rccccccccl}
\text{min:} & 3x_1 & +2x_2 & +2x_3 & +x_4 & & & & & \\
\text{s.t:} & 9x_1 & +x_2 & +x_3 & +x_4 & & & & & = 10 \\
 & & -0.04x_2 & -0.04x_3 & -3.56x_4 & +x_5 & & & & = 10.14 \\
 & & -6.28x_2 & -6.28x_3 & +0.14x_4 & & +x_6 & & & = 1.65 \\
 & & 6.36x_2 & -6.36x_3 & & & & +x_7 & & = 0 \\
\multicolumn{10}{c}{(x_1, x_2, x_3, x_4) \in \mathbb{L}^4} \\
\multicolumn{10}{c}{(x_5, x_6, x_7) \in \mathbb{L}^3} \\
\multicolumn{10}{c}{x_4 \in \mathbb{Z}.}
\end{array}
$$

An integer optimal solution is obtained after adding one cut:

$$x^*_{misoco} = x^*_{soco} = (1.32,\ -0.93,\ -0.93,\ 0.00,\ 10.06,\ -10.06,\ 0.00),$$

with an optimal objective value: $z^*_{misoco} = x^*_{soco} = 0.24.$

Uni-parametric family of quadrics

Theorem

Let (P, p, ρ) *be a quadric and consider two hyperplanes*

$$\mathcal{A}^{=} = \{z \mid a^\top z = \alpha\} \text{ and } \mathcal{B}^{=} = \{z \mid d^\top z = \beta\}.$$

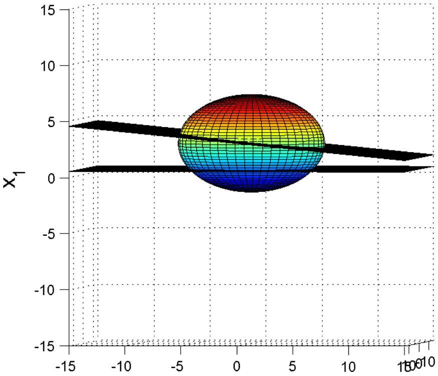
The family of quadrics $(P(\tau), p(\tau), \rho(\tau))$ *parametrized by* $\tau \in \mathbb{R}$ *having the same intersection with* $\mathcal{A}^{=}$ *and* $\mathcal{B}^{=}$ *as the quadric* (P, p, ρ) *is given by*

$$P(\tau) = P + \tau \frac{ad^T + da^T}{2}$$

$$p(\tau) = p - \tau \frac{\beta a + \alpha d}{2}$$

$$\rho(\tau) = \rho + \tau\alpha\beta.$$

Uni-parametric family of quadrics



Sequence of quadrics $z^\top P(\tau)z + 2p(\tau)^\top z + \rho(\tau) \leq 0$,
for $-106.863 \leq \tau \leq 1617$

Classification of Shapes

Range	$(P(\tau), p(\tau), \rho(\tau))$
$\tau > 1617$	*Two sheets hyperboloids*
$\tau = 1617$	*Paraboloid*
$-8.9946 < \tau < 1617$	*Ellipsoids*
$\tau = -8.9946$	*Paraboloid*
$-9.581 < \tau < -8.9946$	*Two sheet hyperboloids*
$\tau = -9.581$	*Cone*
$-106.863 < \tau < -9.581$	*One sheet hyperboloids*
$\tau = -106.863$	*Cone*
$\tau < -106.863$	*Two sheets hyperboloids*

Theorem

Let $\mathcal{A}^= = \{z|a^\top z = \alpha\}$ *and* $\mathcal{B}^= = \{z|a^\top z = \beta\}$ *be two parallel hyperplanes. The quadric*

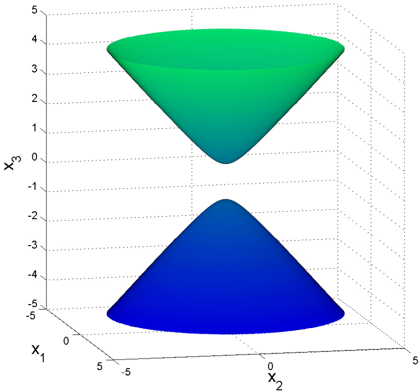
$$(Q(\hat{\tau}), q(\hat{\tau}), \rho(\hat{\tau}))$$

defines a disjunctive conic cut for MISOCO, where $\hat{\tau}$ *is the larger root of the equation*

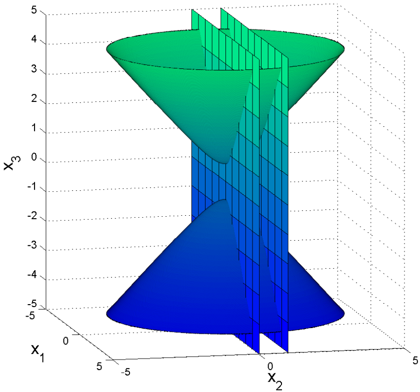
$$q(\tau)^\top Q(\tau) q(\tau) - \rho(\tau) = 0,$$

which is a second degree polynomial in τ*.*

Hyperboloid with Unbounded Intersection

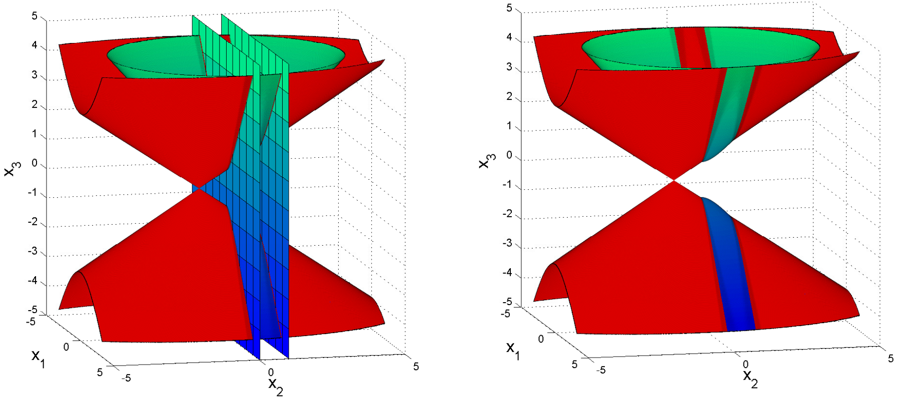


Feasible set



Disjunction

Hyperboloid with Unbounded Intersection



Disjunctive Conic Cut

Agenda

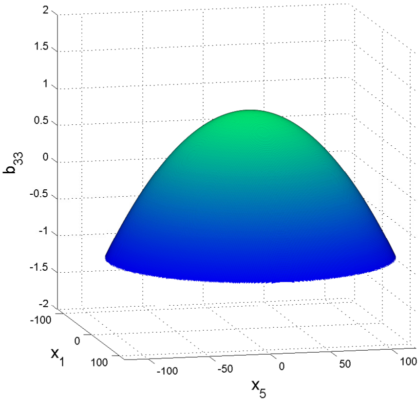
CLay Problems (Bonami et al. 2008)

- Constrained layout problems
- Quadratic constraints corresponding to Euclidean-distance

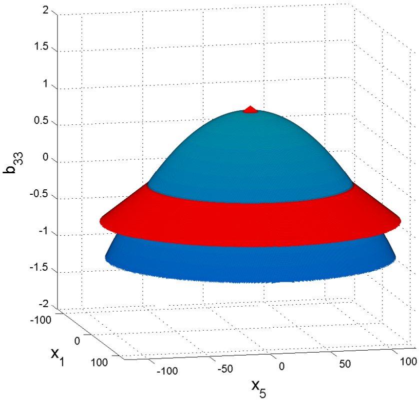
$$(x1 - 17.5)^2 + (x5 - 7)^2 + 6814 * b33 \leq 6850$$

	0203M	0204M	0205M	0303M	0304M	0305M
Var	31	52	81	34	57	86
Binary	18	21	50	21	36	55
Constraints	55	91	136	67	107	156
Quad	24	32	40	36	48	60

CLay Problems (Bonami et al. 2008)

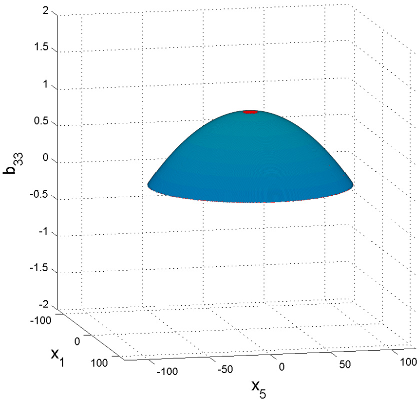


CLay Quadratic Constraints

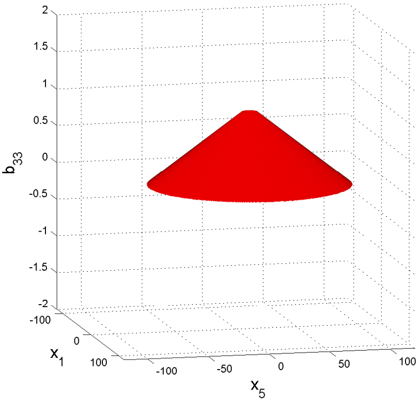


DCC cut

CLay Problems (Bonami et al. 2008)



Original Formulation



DCC Formulation

Clay Problems Solved With CPLEX 12.4

Original Formulation

	0203M	0204M	0205M	0303M	0304M	0305M
Time	0.84	1.12	2.03	1.001	2.02	4.04
Nodes	167	738	8212	453	2549	11188
Iter	1677	3601	47125	6483	23560	65174
Obj	41572.98	6545.00	8092.5	26668.75	40261.08	8029.5

DCC Formulation

	0203M	0204M	0205M	0303M	0304M	0305M
Time	0.44	0.41	1.56	0.467	1.19	1.80
Nodes	165	656	6244	481	1336	8957
Iter	1285	3302	37118	3190	11336	62290
Obj	41565.61	6545.00	8092.5	26662.49	40241.57	8092.5

Difference

	0203M	0204M	0205M	0303M	0304M	0305M
Time	48%	63%	23%	53%	41%	55%
Nodes	1%	11%	24%	-6%	47%	20%
Iter	23%	8%	21%	51%	52%	4%

CLay Problems Solved With MOSEK 6.0

Original Formulation

	0203M	0204M	0205M	0303M	0304M	0305M
Time	3.06	16.91	339.40	7.15	101.98	621.41
Nodes	484	1974	25400	868	8467	38184
Iter	6981	28450	377914	12674	130714	570935
Obj	41573.26	6545.00	8092.5	26669.10	40262.38	8092.50

DCC Formulation

	0203M	0204M	0205M	0303M	0304M	0305M
Time	2.29	15.10	207.90	5.84	76.74	487.46
Nodes	400	2194	20528	838	7013	32875
Iter	5272	27714	271433	10944	104978	455239
Obj	41565.75	6545.00	8092.50	26652.50	40241.57	8092.50

Difference

	0203M	0204M	0205M	0303M	0304M	0305M
Time	25%	11%	39%	18%	25%	22%
Nodes	17%	-11%	19%	3%	17%	14%
Iter	24%	3%	28%	14%	20%	20%

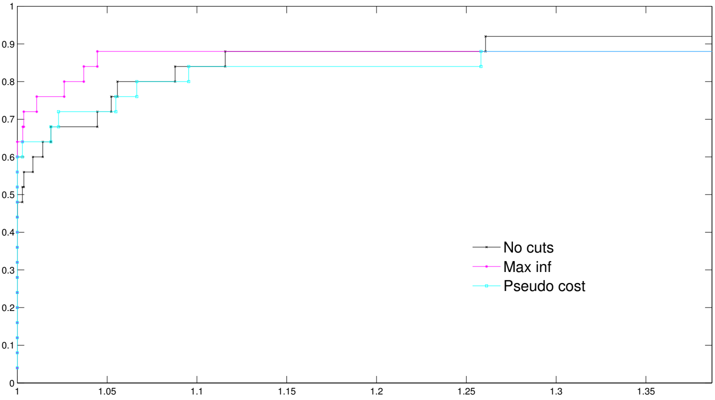
Branch and Cut Solver

- Solver built using the COIN High Performance Parallel Search (CHiPPs) framework on top of the BiCePS layer.
- MOSEK 6 is used to solve the relaxations.
- Experimental Setup
    - Randomly generated problems
    - Naming convention is:
      R(num Rows).C(num Cols).Con(num Cones).Int(num IntVar)
    - One conic cut is added every 10 nodes with a limit of 10 conic cuts in total.

Branching rule: Strong Branching

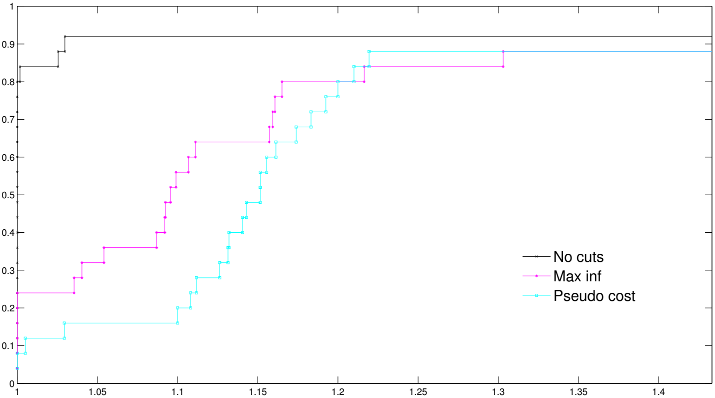
		Selection of Disjunctive Conic Cut		
Rows.Cols.Cones.IntV		No cuts added	Max Inf	Pseudo Cost
R14C18Cone3Int15	Number of Nodes	377	319	375
	CPU time (s)	0.42	0.44	0.48
R17C30Cone3Int15	Number of Nodes	845	845	1035
	CPU time (s)	1.21	1.42	1.69
R17C30Cone3Int21	Number of Nodes	540405	540039	393405
	CPU time (s)	4736.22	5282.25	2110.55
R23C45Cone3Int24	Number of Nodes	1121	1113	1115
	CPU time (s)	1.99	2.44	2.42
R27C50Cone5Int35	Number of Nodes	2226749	2227761	2186683
	CPU time (s)	67741.79	85598.07	84121.83
R27C50Cone5Int50	Number of Nodes	2795427	NaN	3021913
	CPU time (s)	135873.60	NaN	145516.09
R32C60Cone15Int45	Number of Nodes	217115	216787	214887
	CPU time (s)	893.78	936.39	1012.89
R52C75Cone5Int60	Number of Nodes	359195	418927	418865
	CPU time (s)	2140.95	3179.29	3253.12

Branching rule: Strong Branching



Performance profile using the size of the tree

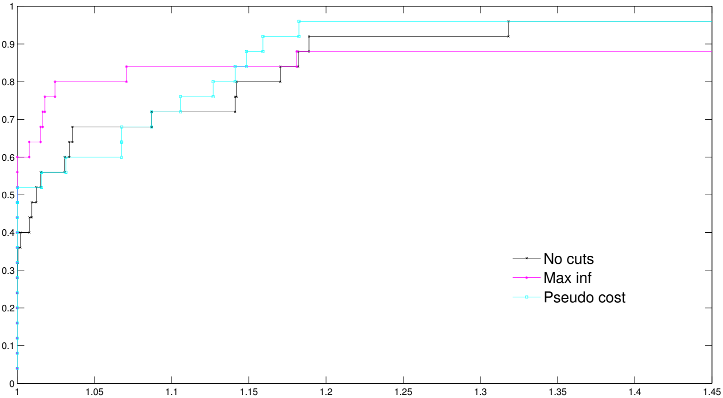
Branching rule: Strong Branching



Performance profile using the solution time

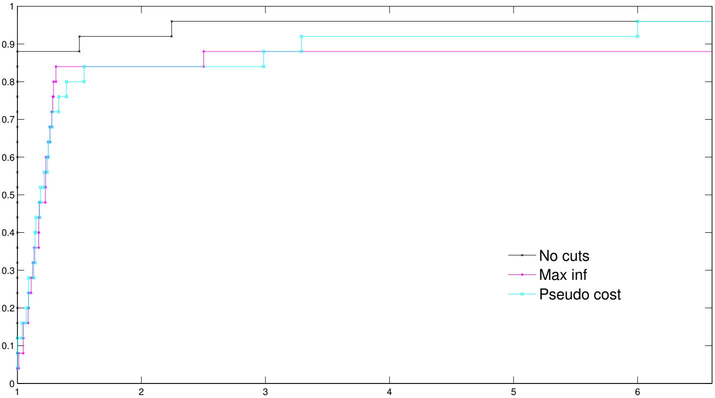
Branching rule: Pseudo Cost

Rows.Cols.Cones.IntV		Selection of Disjunctive Conic Cut: No cuts added	Max Inf	Pseudo Cost
R14C18Cone3Int15	Num Nodes	109	107	109
	CPU time (s)	0.70	0.81	0.77
R17C30Cone3Int15	Num Nodes	536	504	516
	CPU time (s)	6.97	7.66	7.85
R17C30Cone3Int21	Num Nodes	314773	314763	313707
	CPU time (s)	4707.02	5463.58	5439.03
R23C45Cone3Int24	Num Nodes	6154	5530	6042
	CPU time (s)	209.59	217.01	247.99
R27C50Cone5Int35	Num Nodes	NaN	NaN	NaN
	CPU time (s)	NaN	NaN	NaN
R27C50Cone5Int50	Num Nodes	3796593	3492829	NaN
	CPU time (s)	211121.87	205883.77	NaN
R32C60Cone15Int45	Num Nodes	89307	NaN	94583
	CPU time (s)	1445.38	NaN	1648.51
R52C75Cone5Int60	Num Nodes	NaN	NaN	NaN
	CPU time (s)	NaN	NaN	NaN



Performance profile using the size of the tree

Branching rule: Pseudo Cost



Performance profile using the solution time

Conclusions

- The computational experiments show that conic cuts can help to decrease the size of the tree.
- The criteria for selecting the disjunction is important to the effectiveness of the cuts.
- In this case, using the most fractional variables seems to be the best option.
- The addition of conic cuts can significantly increase the solution time of the relaxations, negating the decrease in the size of the tree.
- Numerical issues also arise when adding too many cuts.
- These issues are similar to the ones seen when adding disjunctive cuts in MILO.
- Controlling them will require active management of the relaxation.

Future work

- Investigate more criteria for the construction of the conic cut.
- Investigate methods for actively managing the relaxation to maintain efficiency and numerical stability.
- Investigate the potential to use disjunctive conic cuts in the reformulation of special quadratic constraints like the ones in the CLay problems
- The availability of a good test set for MISOCO problems is needed for a better evaluation and comparison of the cutting techniques available