Reformulations for Mixed-Integer Nonlinear Programs: a surprisingly simple one with surprisingly good results in (quite) a few different applications

Antonio Frangioni

Dipartimento di Informatica, Università di Pisa

with J. Castro, C. D'Ambrosio, F. Furini, L. Galli, C. Gentile, E. Grande, A. Pacifici, M.G. Scutellà, G. Stea

X Brazilian Workshop on Continuous Optimization
March 21, 2014

Outline

Take I: NonLinear Convex-Cost Semi-Continuous Variable

- (Vector of) variable(s) p: either 0 or $\in$ some compact polyhedron $\mathcal{P}$
- Convex cost function $f(p)+$ possibly fixed cost for being $\neq 0$
- Obvious formulation: Mixed-Integer NonLinear Program fragment

$$\min \left\{ f(p) + cu \;:\; Ap \leq bu \;,\; u \in \{0,1\} \right\} \tag{1}$$

$p \in \mathcal{P} = \{ p \in \mathbb{R}^n \;:\; Ap \leq b \}$ compact $\equiv \{ p \;:\; Ap \leq 0 \} = \{0\}$

- Equivalently, minimize the nonconvex function

$$f(p,u) = \begin{cases} 0 & \text{if } u = 0 \text{ and } p = 0 \\ f(p) + c & \text{if } u = 1 \text{ and } Ap \leq b \\ +\infty & \text{otherwise} \end{cases} \tag{2}$$

- Issue: continuous relaxation can be very weak

Take II: Convex Indicator (On/Off) Constraints

- (Vector of) variable(s) p, convex function $f(p)$, one binary variable u
- Constraint $f(p) \leq 0$ "active" $\iff$ $u = 1$, or more in general, constraint $\ldots + s + \ldots \leq d$ with $s = f(p)$ if $u = 1$, $s = 0$ otherwise
- Special case: $p = 0$ if $u = 0$ (but can be generalized)[1]
- Obvious MINLP formulations (plus $Ap \leq bu$, $u \in \{0, 1\}$)

$$f(p) \leq M(1 - u) \tag{3}$$

$$\ldots + s + \ldots \leq d \; , \; s \geq 0 \; , \; s \geq f(p) - M(1 - u) \tag{4}$$

(yet again $\mathcal{P}$ compact for M to exist)

- Same issue: continuous relaxation can be very weak (M large)

[1] Hijazi, Bonami, Cornuéjols, Ouorou "Mixed Integer NonLinear Programs featuring "On/Off" constraints: convex analysis and applications" *Electronic Notes in Discrete Mathematics*, 2010

Outline

The Perspective Reformulation (Relaxation)

- What can we do to improve on this? If f is linear, nothing ...

[2] F., Gentile "Perspective Cuts for a Class of Convex 0-1 Mixed Integer Programs", *Math. Prog.*, 2006

[3] Günlük, Linderoth "Perspective reformulations of MINLPs with indicator variables", *Math. Prog.*, 2010

The Perspective Reformulation (Relaxation)

- What can we do to improve on this? If f is linear, nothing ...

  ... but if f is nonlinear, we can indeed do something

- Best possible convex relaxation of (1): the convex envelope[2]

$$h(p,u) = \begin{cases} 0 & \text{if } p = 0 \text{ and } u = 0 \\ uf(p/u) + cu & \text{if } Ap \leq bu \text{ , } u \in (0,1] \\ +\infty & \text{otherwise} \end{cases} \qquad (5)$$

  leads to the Perspective Reformulation (assuming $0f(0/0) = 0$)

$$\min \{ \, uf(p/u) + cu \; : \; Ap \leq bu \; , \; u \in \{0,1\} \, \} \qquad (6)$$

- Similar[3] for (3)/(4): given $\mathcal{P}_0 = \{ \, 0 \, \}$, $\mathcal{P}_1 = \mathcal{P} \cap \{ \, f(p) \leq 0 \, \}$

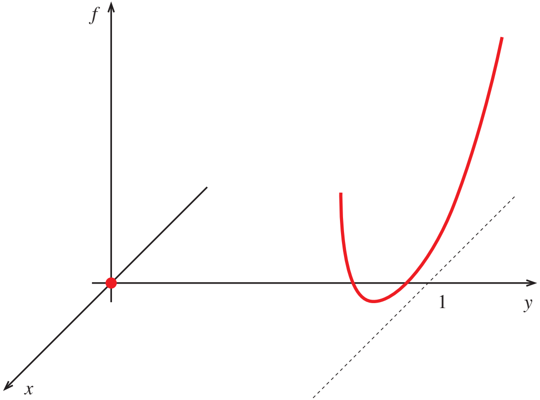
$$conv(\, \mathcal{P}_0 \cup \mathcal{P}_1 \,) = \left\{ \, p \; : \; Ap \leq bu \; , \; uf(p/u) \leq 0 \; , \; u \in \{0,1\} \, \right\} \qquad (7)$$

[2] F., Gentile "Perspective Cuts for a Class of Convex 0-1 Mixed Integer Programs", *Math. Prog.*, 2006

[3] Günlük, Linderoth "Perspective reformulations of MINLPs with indicator variables", *Math. Prog.*, 2010

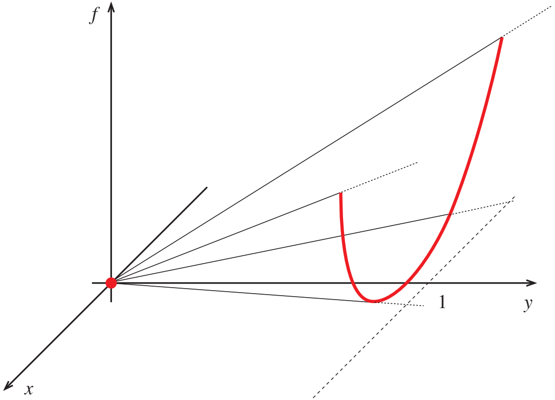
The Perspective what?

- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



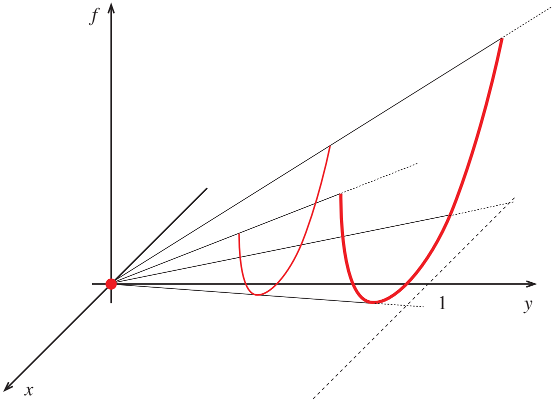
The Perspective what?

- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



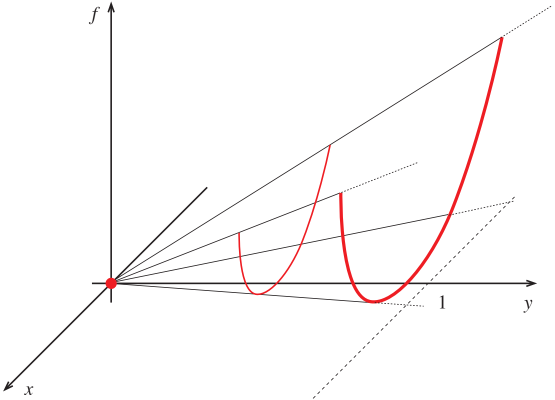
The Perspective what?

- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



The Perspective what?

- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$

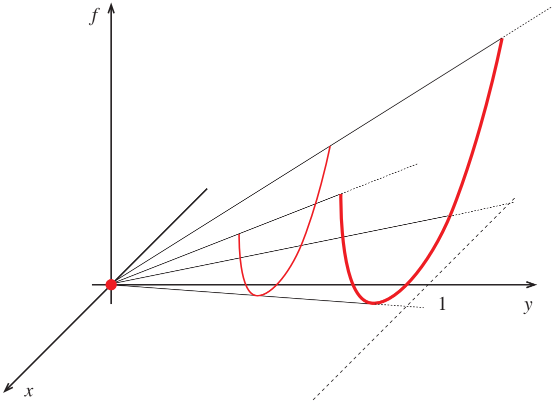


- $h(p, u)$ convex but "much more nonlinear" than $f(p) + cu$

example: $f(p) = ap^2 + bp \quad \Longrightarrow \quad h(p, u) = (a/u)p^2 + bp + cu$

The Perspective what?

- $h(p,u)$ is a section of the perspective function $f(x,\lambda) = \lambda f(x/\lambda)$



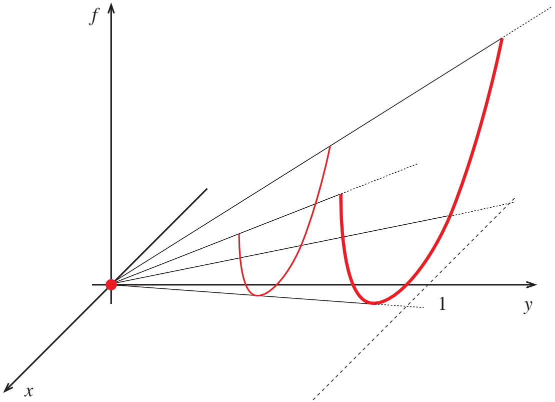
- $h(p,u)$ convex but "much more nonlinear" than $f(p) + cu$

example: $f(p) = ap^2 + bp \quad \Longrightarrow \quad h(p,u) = (a/u)p^2 + bp + cu$

notes: I) $a/u > a$ for $u < 1$;

The Perspective what?

- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



- $h(p, u)$ convex but "much more nonlinear" than $f(p) + cu$

example: $f(p) = ap^2 + bp \quad \Longrightarrow \quad h(p, u) = (a/u)p^2 + bp + cu$

notes: I) $a/u > a$ for $u < 1$; II) for $a = 0$ nothing happens

Solving the Perspective Relaxation I

- Continuous relaxation of (6): convex, more nonlinear

$$\min \{\ uf(p/u) + cu\ :\ Ap \leq bu\ ,\ u \in [0,1]\ \} \tag{8}$$

lower bound much better than ordinary continuous relaxation

$$\min \{\ f(p) + cu\ :\ Ap \leq bu\ ,\ u \in [0,1]\ \} \tag{9}$$

- But how to solve it efficiently?

[4] Ben Tal "Lecture Notes of the Course on Conic and Robust Optimization", 2002

Solving the Perspective Relaxation I

- Continuous relaxation of (6): convex, more nonlinear

$$\min \left\{ \, uf(p/u) + cu \;:\; Ap \leq bu \,,\; u \in [0,1] \, \right\} \qquad (8)$$

lower bound much better than ordinary continuous relaxation

$$\min \left\{ \, f(p) + cu \;:\; Ap \leq bu \,,\; u \in [0,1] \, \right\} \qquad (9)$$

- But how to solve it efficiently?
- Good news: $\lambda f(x/\lambda)$ is SOCP-representable if f is[4]
- Example: for $f(p) = ap^2 + bp$, (6) becomes

$$\min \left\{ \, t + bp + cu \;:\; ap^2 \leq tu \,,\; Ap \leq bu \,,\; u \in \{0,1\} \, \right\} \qquad (10)$$

a Mixed-Integer (rotated) Second-Order Cone Program: Cplex/Gurobi can solve it

[4] Ben Tal "Lecture Notes of the Course on Conic and Robust Optimization", 2002

Solving the Perspective Relaxation II

- Is it the only way? Of course not.
- Every convex function is the supremum of its affine minorants

$(v, p, u) \in epi\ h \iff Ap \leq bu,\ u \in [0, 1]$, and $\forall \bar{p} \in \mathcal{P}$

$$v \geq f(\bar{p}) + c + \left[\, s\ ,\ c + f(\bar{p}) - s\bar{p} \,\right] \left[\begin{array}{c} p - \bar{p} \\ u - 1 \end{array} \right] \qquad \forall s \in \partial f(\bar{p}) \quad (11)$$

- Infinitely many inequalities: looks hard, but actually OK for B&C
- The quadratic case: Perspective Cuts (P/C)

$$v \geq (2a\bar{p} + b)p + (c - a\bar{p}^2)u \qquad \forall \bar{p} \in \mathcal{P} \qquad (12)$$

- Basically the same thing as linearizing the cones in (10), which can be done automatically . . . but does not work nearly as well (don't know why)

Outline

Application 1: Quadratic Radar Sensors Location

- Sometimes it's basically the only structure you have
- Optimal placement of n circles to cover a line segment
- Activation constraints $p_i \leq u_i$ (p_i: coverage diameter)
- Install cost $c_i u_i + a_i p_i^2$ for covering diameter p_i (circle)

$$\begin{array}{ll} \min & \sum_{i=1}^{n} c_i u_i + a_i p_i^2 \\ & \sum_{i=1}^{n} p_i \geq 1 \\ & 0 \leq p_i \leq u_i \;\; , \;\; u_i \in \{0,1\} \quad i = 1, \ldots, n \end{array} \tag{13}$$

- $\mathcal{NP}$-hard (from Knapsack, so non that hard ...)
- Continuous relaxation: at optimum $u_i = p_i \Longrightarrow$ weak bounds
- PR: i-th term of objective function $a_i p_i^2 / u_i + c_i u_i$
- PR helps, but knapsack-like structure destroyed

Application 2: Quadratic (1-Commodity) Network Design

- More often you have (1) + other combinatorial structures
- Directed graph $G = (N, A)$, deficit b_i for $i \in N$, arc capacity $\bar{p}_{ij}$ with fixed-charge cost $c_{ij} > 0$, quadratic routing cost

$$\begin{array}{lll} \min & \sum_{(i,j)\in A} c_{ij} u_{ij} + b_{ij} p_{ij} + a_{ij} p_{ij}^2 & \\ & \sum_{(j,i)\in A} p_{ji} - \sum_{(i,j)\in A} p_{ij} = b_i \quad i \in N & \\ & 0 \le p_{ij} \le \bar{p}_{ij} u_{ij} \;\; , \;\; u_{ij} \in \{0,1\} \quad (i,j) \in A & \end{array} \tag{14}$$

- Again $\mathcal{NP}$-hard (from Network Design, so harder)
- In continuous relaxation, $u_{ij} = p_{ij} / \bar{p}_{ij} \Longrightarrow$ a quadratic flow problem
- Very weak bound, PR improves it but destroys flow structure
- Perhaps sounds too artificial to you?

Application 3: The Unit Commitment problem

- Can be a small part of a large and complicated problem
- Set P of units, iscretized time horizon $\mathcal{T}$, energy demand $\bar{d}_t$ for $t \in \mathcal{T}$

$$
\begin{array}{lr}
\min \sum_{i \in P} \sum_{t \in \mathcal{T}} a^i_t (p^i_t)^2 + b^i_t p^i_t + c^i_t u^i_t & \\
\quad \bar{p}^i_{min} u^i_t \leq p^i_t \leq \bar{p}^i_{max} u^i_t & t \in \mathcal{T} \\
\quad p^i_t \leq p^i_{t-1} + u^i_{t-1} \Delta^i_+ + (1 - u^i_{t-1}) \bar{l}^i & t \in \mathcal{T} \\
\quad p^i_{t-1} \leq p^i_t + u^i_t \Delta^i_- + (1 - u^i_t) \bar{u}^i & t \in \mathcal{T} \\
\quad u^i_t \leq 1 - u^i_{r-1} + u^i_r & t \in \mathcal{T},\ r \in [t - \tau^i_+, t-1] \\
\quad u^i_t \geq 1 - u^i_{r-1} - u^i_r & t \in \mathcal{T},\ r \in [t - \tau^i_-, t-1] \\
\quad \sum_{i \in P} p^i_t = \bar{d}_t & t \in \mathcal{T} \\
\quad u^i_t \in \{0, 1\} & i \in P \ , \quad t \in \mathcal{T}
\end{array}
$$

- Simplified version (hydro units, time-dependent start-up costs[5], . . .)

[5] Nowak, Römisch "Stochastic Lagrangian Relaxation Applied to Power Scheduling in a Hydro-thermal System Under Uncertainty", *Annals of O.R.*, 2000

A First Glimpse to the Computational Results

- P/C (8) (via (12)) vs. ordinary continuous relaxation (9)

P/C				CPLEX				
r.time	r.gap	time	nodes	r.time	r.gap	time	nodes	gap
4.17	0.28	15.61	3	1.41	2.36	10000	264179	1.27
4.29	0.13	4.53	1	1.80	0.49	62	1205	-
2.07	0.69	178.12	136	0.85	1.24	216	4083	-
8.64	0.28	37.14	4	1.61	2.40	10000	331732	1.43
8.42	0.20	23.75	2	1.71	1.63	10000	245582	0.87
6.71	0.24	12.59	2	1.58	1.37	10000	268516	0.73
4.83	0.28	12.71	3	0.87	2.23	10000	475400	1.45
5.97	0.18	19.35	3	1.74	1.06	6137	189898	-
6.73	0.23	44.35	44	1.55	2.60	10000	337915	1.69
7.96	0.26	141.69	73	1.64	2.28	10000	286651	1.02
5.98	0.28	48.98	57	1.48	1.77	7642	240516	0.85

- Root gap node greatly reduced at a small expense in running time
- Nice thing is: you only need a `cutcallback`

Computational Results: Conic Program Reformulation

P/C				CP			
nds	LPs	time	t/LP	nds	CPs	time	t/CP
612	936	18.36	0.0196	812	1323	24.62	0.0186
630	1106	18.42	0.0167	709	1212	24.38	0.0201
374	724	13.92	0.0192	943	1488	28.80	0.0194
280	585	11.50	0.0197	149	416	12.43	0.0299
277	527	6.62	0.0126	285	600	10.11	0.0169
40527	47005	6855.43	0.1458	58794	67516	14242.65	0.2110
5670	7460	581.84	0.0780	9966	12611	1765.41	0.1400
163414	183594	23066.99	0.1256	170972	193056	41283.69	0.2138
37004	43191	3179.55	0.0736	82600	95947	9121.58	0.0951
5850	7446	341.81	0.0459	5836	8215	545.88	0.0664

- LPs somewhat faster than CPs $\Longrightarrow$ P/C better than CP[6]
- Also somewhat less nodes despite the same bound (or worse)
  Mysteries of the B&C (heuristics?)

[6] F., Gentile "A Computational Comparison of [. . .]: SOCP vs. Cutting Planes" *Op. Res. Letters*, 2009

Application 4: Delay Constrained Routing Problem

- IP Network $\equiv$ directed graph $G = (N, A)$, MTU L, node delay n_i
- One "new" s–d flow arriving, burst σ and rate ρ, deadline δ
- link delay l_{ij}, link speed w_{ij}, reservable capacity c^k_{ij} ($\leq w_{ij}$)
- Worst-case delay of the flow depends on:
  1. the selected s-d path P in G;
  2. the reserved rate (capacity) $r_{ij} \in [0, c_{ij}]$ for each $(i, j) \in P$
- Necessary assumption for finite delay: $r_{ij} \geq \rho$ for each $(i, j) \in P$ ($\rho \equiv$ rate $\equiv$ "steady-state" flow demand in classical flow models)
- Delay constraint (Strictly Rate-Proportional protocol, may be $\neq$)

$$\frac{\sigma}{\min\{ r_{ij} \; : \; (i,j) \in P\}} + \sum_{(i,j)\in P} \left(\frac{L}{r_{ij}} + \frac{L}{w_{ij}} + l_{ij} + n_i \right) \leq \delta \qquad (15)$$

convex and SOCP-representable, but nonlinear

A (partial) SOCP Model of DCR

- Path binary variables x_{ij}, reserve continuous variables r_{ij}

$$\min \textstyle\sum_{(i,j)\in A} r_{ij} \tag{16}$$

$$\sum_{(j,i)\in BS(i)} x_{ji} - \sum_{(i,j)\in FS(i)} x_{ij} = \begin{cases} -1 & \text{if } i = s \\ 1 & \text{if } i = d \\ 0 & \text{otherwise} \end{cases} \qquad i \in N \tag{17}$$

$$0 \leq r_{ij} \leq c_{ij} x_{ij} \qquad (i,j) \in A \tag{18}$$

$$\rho \leq r_{min} \leq r_{ij} + c_{max}(1 - x_{ij}) \qquad (i,j) \in A \tag{19}$$

$$t + \sum_{(i,j)\in A} \left(\theta_{ij} + \left(\frac{L}{w_{ij}} + l_{ij} + n_i \right) x_{ij} \right) \leq \delta \tag{20}$$

$$t\, r_{min} \geq \sigma \quad , \quad t \geq 0 \tag{21}$$

$$x_{ij} \in \{0,1\} \quad , \quad r_{ij} \in \mathbb{R} \qquad (i,j) \in A$$

- (21) rotated SOCP constraint $\equiv t \geq \sigma / r_{min}$ (since $t \geq 0$)
- Issue: how to write "$x_{ij} = 1 \Longrightarrow \theta_{ij} \geq L / r_{ij}$, $x_{ij} = 1 \Longrightarrow \theta_{ij} = 0$"

Modeling the Delay Contribution (On/Off Constraint)

- Issue: can't use $r_{ij}\,\theta_{ij} \geq L$ for this $\Longrightarrow$ $\theta_{ij} > 0$ always
- Solution: two extra sets of variables s_{ij} and r'_{ij}

$$0 \leq \theta_{ij} \leq M x_{ij}$$

$$\theta_{ij} \geq s_{ij} - M(1 - x_{ij})$$

$$s_{ij}\, r'_{ij} \geq L \quad , \quad s_{ij} \geq 0$$

$$0 \leq r'_{ij} \leq r_{ij} + M(1 - x_{ij})$$

$M = \max(\sqrt{L}\,,\ L/\rho)$ suffices, still it's big-M

- The PR variant: apply (7), after a little tedious algebra

$$\rho x_{ij} \leq r_{ij} \leq c_{ij} x_{ij} \quad , \quad 0 \leq \theta_{ij} \leq (L/\rho) x_{ij} \quad , \quad \theta_{ij} r_{ij} \geq L x_{ij}^2$$

(now θ_{ij} can be 0 when $x_{ij} = 0$, x^2/r convex for $r > 0$)

- Original variables + a(nother rotated) SOCP constraint

Sample of the Results (Real or Realistic Networks, Cplex)

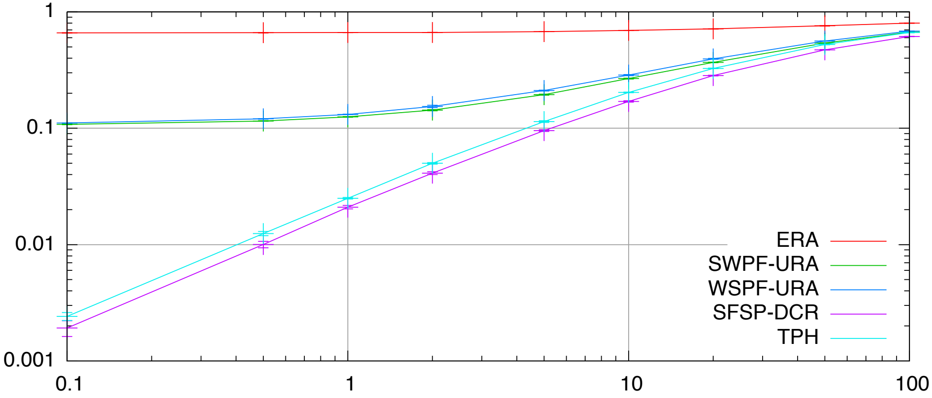
	PR				bM			
	avg		max		avg		max	
	t	n	t	n	t	n	t	n
abilene	0.011	0.000	0.03	0	0.02	0.03	0.09	1
atlanta	0.015	0.044	0.18	1	0.03	0.07	0.17	1
di-yuan	0.051	1.190	0.34	18	0.11	1.36	0.62	31
giul39	0.245	0.547	0.99	13	1.27	15.33	6.68	610
india35	0.021	0.036	0.27	1	0.08	0.07	0.58	4
janos-us	0.093	0.108	0.63	7	0.43	2.65	1.55	30
janos-us-ca	0.141	0.138	0.83	8	0.80	5.76	2.76	243
pdh	0.042	0.444	0.38	8	0.11	0.74	0.38	13
pioro40	0.019	0.039	0.27	1	0.10	0.14	0.57	6
sun	0.165	0.587	0.89	13	0.65	7.68	2.36	257
ta2	0.020	0.015	0.13	1	0.12	0.08	0.89	4
w1-100-04	1.854	3.176	43.14	85	8.88	164.49	43.87	2585
w1-200-04	24.231	25.366	413.95	4075	231.09	2714.68	9088.54	127429

Data @ http://www.di.unipi.it/optimize/Data/MMCF.html#UMMCF[7]

[7] F., Galli, Scutellà "Delay-Constrained Shortest Paths: Approximation Algorithms and Second-Order Cone Models", submitted

Does it really matter in practice?

- Large-scale simulation of the network behavior[8]
- SOCP model (+ Quick&Dirty modification) vs. the literature
- Blocking probability as a function of network load

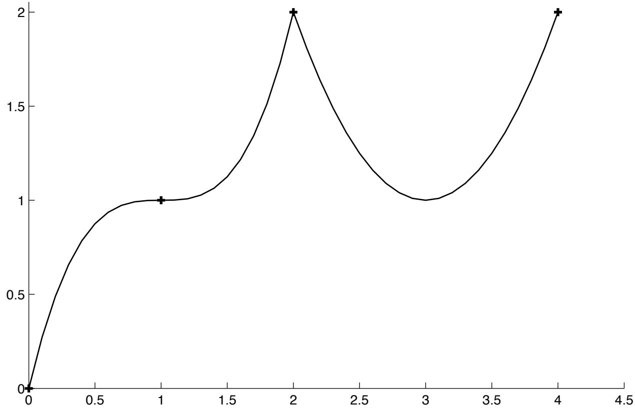


- Similar on all topologies, and network conditions

[8] F., Galli, Stea "Optimal Joint Path Computation and Rate Allocation for Real-time Traffic", submitted

Application 5: Univariate NonConvex MINLPs

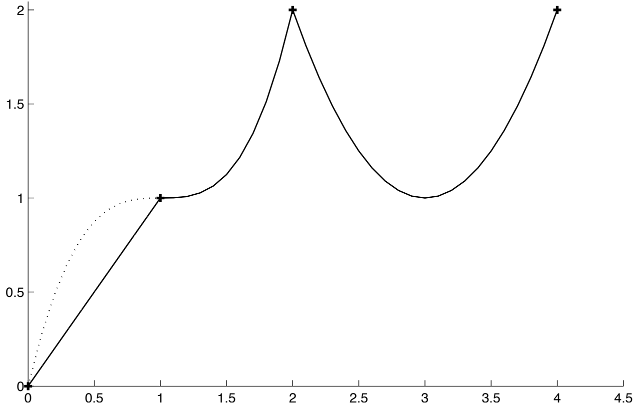
- Nonconvex (MI)NLP where the nonconvexity is only univariate
- Idea: relax into a Convex MINLP, refine as needed[9]



[9] D'Ambrosio, Lee, Wächter "An Algorithmic Framework for MINLP with Separable Non-Convexity" *IMA Volumes*, 2012

Application 5: Univariate NonConvex MINLPs

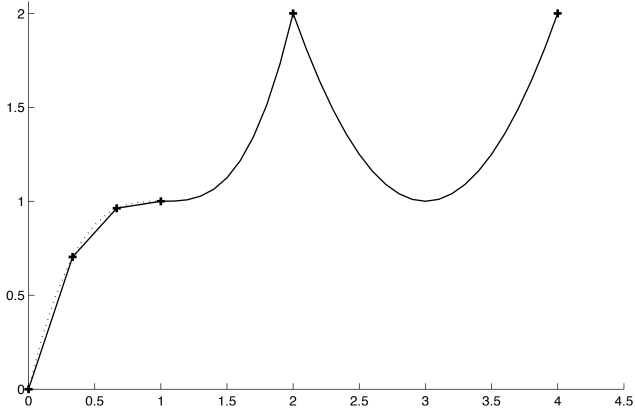
- Nonconvex (MI)NLP where the nonconvexity is only univariate
- Idea: relax into a Convex MINLP, refine as needed[9]



[9] D'Ambrosio, Lee, Wächter "An Algorithmic Framework for MINLP with Separable Non-Convexity" *IMA Volumes*, 2012

Application 5: Univariate NonConvex MINLPs

- Nonconvex (MI)NLP where the nonconvexity is only univariate
- Idea: relax into a Convex MINLP, refine as needed[9]



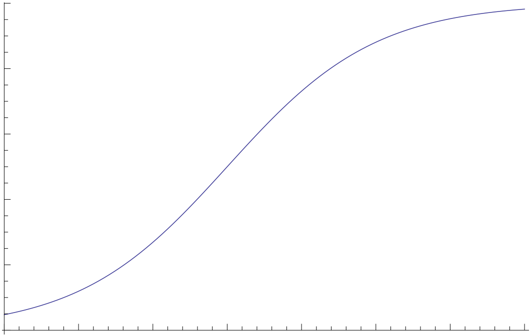
- One new nonlinear semicontinuous variable for each convex(ified) piece

[9] D'Ambrosio, Lee, Wächter "An Algorithmic Framework for MINLP with Separable Non-Convexity" *IMA Volumes*, 2012

Example 5.1: Nonlinear Continuous Knapsack

$$\max \left\{ \sum_{j \in N} g_j(x_j) \;:\; \sum_{j \in N} x_j \leq C \,,\; 0 \leq x_j \leq U \quad j \in N \right\}$$

$$\text{where } g_j(x_j) = \frac{c_j}{1 + b_j e^{-a_j(x_j + d_j)}}$$

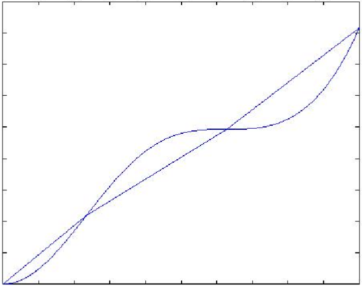


- Similar to Application 1, continuous but nonconvex

Example 5.2: Nonlinear Uncapacitated Facility Location

- K facilities, T customers, nonlinear transportation costs $g_{kt}(w_{kt})$

$$
\begin{aligned}
\min \sum_{k \in K} C_k y_k \; + \; & \sum_{t \in T} \sum_{k \in K} g_{kt}(w_{kt}) \\
0 \leq w_{kt} \leq y_k \qquad & t \in T \, , \, k \in K \\
\sum_{k \in K} w_{kt} = 1 \qquad & t \in T \\
y_k \in \{0, 1\} \qquad & k \in K
\end{aligned}
$$



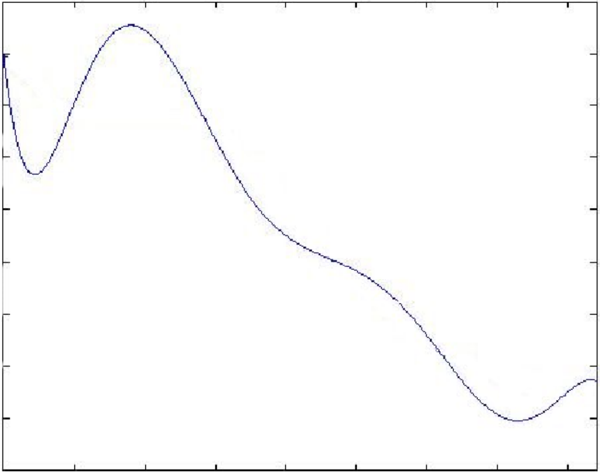
- Similar to Application 2, continuous but nonconvex

Example 5.3: Hydro Unit Commitment

- J units, $\mathcal{T}$ instances, nonlinear efficiency $\varphi(q_{jt})$

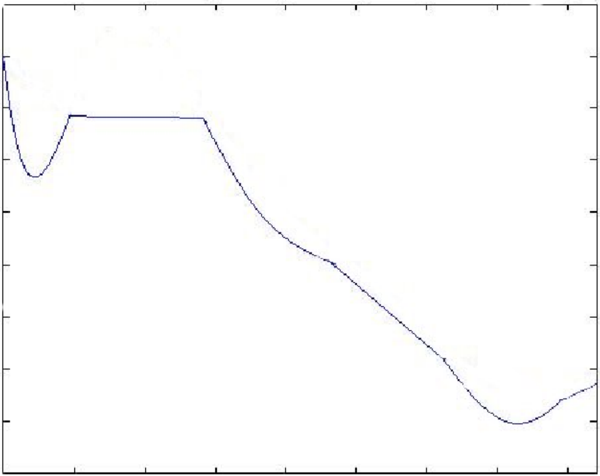
$$
\begin{array}{lr}
\min \sum_{j\in J}\sum_{t\in\mathcal{T}} \Delta t \Pi_t \varphi(q_{jt}) + C_j\, \widetilde{w}_{jt} + (D_j + \Pi_t E_j)\widetilde{y}_{jt} & \\
\sum_{j\in J}\sum_{t\in\mathcal{T}} v_t - v_{t-1} = 3600\,\Delta t\,(I_t - \sum_{j\in J} q_{jt} - s_t) & t\in\mathcal{T} \\
(Q_j^-\, u_{jt} + \underline{Q}_j\, g_{jt}) \le q_{jt} \le (Q_j^-\, u_{jt} + \overline{Q}_j\, g_{jt}) & j\in J, t\in\mathcal{T} \\
\sum_{j\in J}(q_{jt} - q_{j(t-1)}) + \Delta q^- \ge 0 & t\in\mathcal{T} \\
\sum_{j\in J}(q_{jt} - q_{j(t-1)}) - \Delta q^+ \le 0 & t\in\mathcal{T} \\
s_t - \sum_{j\in J}(W_j\, \widetilde{w}_{jt} + Y_j\, \widetilde{y}_{jt}) \ge 0 & t\in\mathcal{T} \\
\sum_{j\in J} q_{jt} + s_t - \underline{\Theta} \ge 0 & t\in\mathcal{T} \\
g_{jt} - g_{j(t-1)} = (\widetilde{w}_{jt} - w_{jt}) & j\in J\,,\ t\in\mathcal{T} \\
\widetilde{w}_{jt} + w_{jt} \le 1 & j\in J\,,\ t\in\mathcal{T} \\
u_{jt} - u_{j(t-1)} = (\widetilde{y}_{jt} - y_{jt}) & j\in J\,,\ t\in\mathcal{T} \\
\widetilde{y}_{jt} + y_{jt} \le 1 & j\in J\,,\ t\in\mathcal{T} \\
g_{jt} + u_{kt} \le 1 & j,k\in J\,,\ t\in\mathcal{T} \\
\sum_{j\in J} u_{jt} \le \overline{n} - 1 & t\in\mathcal{T}
\end{array}
$$

- A complicated problem with a complicated function



Hydro Unit Commitment and Scheduling (cont.d)

- A complicated problem with a complicated function



- Similar to Application 3, nonconvex and mixed-integer

Preliminary Results for Example 5.1 (very fresh!)

- Solving a single step (no refinement phase as yet)
- Solving the nonlinear problem by gradient linearization in Cplex (for fairness, but it turns out to be quite effective)

	PR				Cplex			
N	nds	CPU	LPs	cuts	nds	CPU	LPs	cuts
100	0	0.070	49	928	1.3	0.204	105	1206
200	0	0.116	44	1780	1.5	0.610	124	2391
500	0	0.394	52	4585	6.3	3.599	183	7334

- Once again, PR seems to uncontestedly dominate

Application 6: Mean-Variance Portfolio problem

- Mean-Variance problem with min and max buy-in thresholds

$$\min \left\{ x^T Q x \;\middle|\; \begin{array}{l} ex = 1 \;\;,\;\; \mu x \geq \rho \;, \\ l_i y_i \leq x_i \leq u_i y_i \;,\; y_i \in \{0,1\} \;\; i = 1, \ldots, n \end{array} \right\}$$

μ = expected return, Q = covariance matrix, ρ = desired return

- But f is nonseparable, so Perspective Reformulation not applicable

Application 6: Mean-Variance Portfolio problem

- Mean-Variance problem with min and max buy-in thresholds

$$\min \left\{ x^T Q x \;\middle|\; \begin{array}{l} ex = 1 \;\;,\;\; \mu x \geq \rho \;, \\ l_i y_i \leq x_i \leq u_i y_i \;,\; y_i \in \{0, 1\} \;\; i = 1, \ldots, n \end{array} \right\}$$

$\mu =$ expected return, $Q =$ covariance matrix, $\rho =$ desired return

- But f is nonseparable, so Perspective Reformulation not applicable

- Dirty trick: choose $D \succeq 0$ diagonal s.t. $R = Q - D \succeq 0$

$$\min \left\{ x^T D x + z^T R z \;\middle|\; \begin{array}{l} ex = 1 \;\;,\;\; \mu x \geq \rho \;\;,\;\; z = x \\ l_i y_i \leq x_i \leq u_i y_i \;,\; y_i \in \{0, 1\} \;\; i = 1, \ldots, n \end{array} \right\}$$

move nonseparability to new variables p, let D "as large as possible"

- Simple choice $D = \lambda_{min}(Q) I$, much better the solution of SDP

$$\max \left\{ \; \sum_{i=1}^{n} d_i \;:\; Q - \sum_{i=1}^{n} d_i (e_i e_i^T) \succeq 0 \; [\; , \; d \geq 0 \;] \; \right\}$$

$$\min \left\{ \; tr(QX) \;:\; diag(X) \geq (=) \, e \;,\; X \succeq 0 \; \right\}$$

Results of P/C

	P/C				Cplex			
	time	nodes	d.gap	r.gap	nodes	p.gap	d.gap	r.gap
200^+	164	1.2e+4		1.14	1.9e+7	0.14	45.33	85.63
200^0	161	1.1e+4		2.14	8.5e+6	0.38	51.27	84.47
200^-	1902	1.3e+5		3.65	8.9e+6	0.24	42.09	78.88
300^+	818	2.9e+4		4.54	4.0e+6	0.41	64.68	92.01
300^0	856	2.7e+4		1.97	3.6e+6	0.43	59.91	87.87
300^-	1709	5.2e+4		2.68	3.6e+6	0.43	59.91	87.87
400^+	2264	7.0e+4		4.79	3.0e+6	0.53	45.11	78.77
400^0	4378	7.2e+4	0.10	2.29	1.5e+6	1.18	68.68	90.03
400^-	6311	1.0e+5	0.23	3.06	1.5e+6	1.60	65.88	88.47

- Root node gap divided by 20: $\leq 4\%$ w.r.t. $> 80\%$
- Effectiveness worsens as Q less dominant, could not solve all the 400 instances, but much better than Cplex: none solved in 10000s
- Choosing the right D via SDP important: results much worse otherwise, although still much better that non-PR

Comparing CP with P/C on MV

	P/C				CP				
	nodes	QPs	time	t/QP	nodes	CPs	time	t/CP	gap
200^+	1.9e4	1.9e4	194	0.0008	9.2e3	1.1e3	17961	1.578	0.15(1)
200^0	1.7e4	1.8e4	90	0.0007	2.7e4	3.2e4	30785	1.648	0.32(2)
200^-	1.2e5	1.3e5	835	0.0006	1.6e4	1.9e5	55144	1.719	1.02(5)
300^+	3.4e4	3.5e4	433	0.0014	1.1e4	1.4e4	72075	8.334	0.58(7)
300^0	3.1e5	3.3e4	378	0.0019	1.0e4	1.3e4	59591	4.464	0.53(6)
300^-	5.5e5	5.8e4	654	0.0014	1.1e4	1.3e4	66863	5.272	0.81(7)
400^+	7.9e4	8.2e4	2066	0.0032	4.7e3	5.9e3	61810	10.397	1.01(6)
400^0	2.3e5	2.4e5	3974	0.0020	6.1e3	7.6e3	83782	10.588	1.79(9)
400^-	3.3e5	3.4e5	8092	0.0026	6.3e3	7.9e3	80382	10.764	2.71(8)

- Same SDP solved for CP and P/C
- QPs awesomely faster than CPs to solve ("quadratic simplex") even invoking the built-in SOCP linearization (???)
- Faster bound $\Longrightarrow$ more nodes $\Longrightarrow$ faster convergence
- Best case (very unstructured problem)

Outline

Projected Perspective Reformulation

- We tried quite a few ideas (Newton, ...) before "discovering" CP

Projected Perspective Reformulation

- We tried quite a few ideas (Newton, . . .) before “discovering” CP
- The (failed) Newton idea eventually led us to projecting

Projected Perspective Reformulation

- We tried quite a few ideas (Newton, ...) before "discovering" CP
- The (failed) Newton idea eventually led us to projecting
- Only works under strong assumptions (often true):

  A1 p is a single variable, $\bar{p}_{min} \geq 0$

  A2 f is quadratic

  A3 there are no constraints linking different u

- Basic idea: recast the PR as

$$\min \{ \ z(p) \ : \ p \in [0, \bar{p}_{max}] \ \}$$

where $z(p)$ (partial minimization of a convex function $\Rightarrow$ convex) is

$$z(p) = bp + \min_u \{ \ ap^2/u + cu \ : \ u\bar{p}_{min} \leq p \leq u\bar{p}_{max} \ , \ p \in [0, 1] \ \}$$

(intuition: the projection will be "less nonlinear")

Derivation

- Algebraic characterization of $z(p)$ out of optimal solution $u^*(p)$[10]
- In turn, $u^*(p)$ out of unconstrained minimizer $\tilde{u}(p)$, i.e., solution to

$$\frac{\partial h(p,u)}{\partial u} = c - ap^2/p^2 = 0$$

(if any)

  - if $\tilde{u}(p)$ is feasible, then $u^*(p) = \tilde{u}(p)$
  - otherwise, $u^*(p)$ is the projection of $\tilde{u}(p)$ over the feasible region (easy because the feasible region is very simple)
  - note: $ap^2/u^2 \geq 0$

- Basically, a few algebraic computations depending on a, c, ...

[10] F., Gentile, Grande, Pacifici "Projected Perspective Reformulations with Applications in Design Problems" *Operations Research* 2010

The form of $z(p)$

1) $c \leq 0 \Longrightarrow \tilde{u}(p)$ undefined $\Longrightarrow u^*(p) = 1 \Longrightarrow$

$$\boxed{z(p) = ap^2 + bp + c}$$

2) $c > 0 \Longrightarrow \tilde{u}(p) = p\sqrt{a/c}$

2.1) $\tilde{u} \leq p/\bar{p}_{max} \Longleftrightarrow \bar{p}_{max} \leq \sqrt{c/a} \Longleftrightarrow u^*(p) = p/\bar{p}_{max} \Longrightarrow$

$$\boxed{z(p) = (\ b + a\bar{p}_{max} + c/\bar{p}_{max}\ )\, p}$$

2.2) $0 \geq \tilde{u}(p) \geq p/\bar{p}_{max} \Longleftrightarrow \bar{p}_{max} \geq \sqrt{c/a}\ (\geq \bar{p}_{min})$.

- $(\bar{p}_{max} \geq)\ p \geq \sqrt{c/a}\ (\geq 0) \Longrightarrow \tilde{u}(p) \geq 1 \Longrightarrow u^*(p) = 1$;
- $0 \leq p \leq \sqrt{c/a}\ (\leq \bar{p}_{max}) \Longrightarrow \tilde{u}(p) \leq 1 \Longrightarrow u^*(p) = \tilde{u}(p)$.

$$\Longrightarrow \boxed{z(p) = \begin{cases} (\ b + 2\sqrt{ac}\ )\, p & \text{if } 0 \leq p \leq \sqrt{c/a} \\ ap^2 + bp + c & \text{if } \sqrt{c/a} \leq p \leq \bar{p}_{max} \end{cases}}$$

The form of $z(p)$

1) $c \leq 0 \Longrightarrow \tilde{u}(p)$ undefined $\Longrightarrow u^*(p) = 1 \Longrightarrow$

$$\boxed{z(p) = ap^2 + bp + c}$$

2) $c > 0 \Longrightarrow \tilde{u}(p) = p\sqrt{a/c}$

2.1) $\tilde{u} \leq p/\bar{p}_{max} \Longleftrightarrow \bar{p}_{max} \leq \sqrt{c/a} \Longleftrightarrow u^*(p) = p/\bar{p}_{max} \Longrightarrow$

$$\boxed{z(p) = (\ b + a\bar{p}_{max} + c/\bar{p}_{max}\ )p}$$

2.2) $0 \geq \tilde{u}(p) \geq p/\bar{p}_{max} \Longleftrightarrow \bar{p}_{max} \geq \sqrt{c/a}\ (\geq \bar{p}_{min})$.

- $(\bar{p}_{max} \geq)\ p \geq \sqrt{c/a}\ (\geq 0) \Longrightarrow \tilde{u}(p) \geq 1 \Longrightarrow u^*(p) = 1$;
- $0 \leq p \leq \sqrt{c/a}\ (\leq \bar{p}_{max}) \Longrightarrow \tilde{u}(p) \leq 1 \Longrightarrow u^*(p) = \tilde{u}(p)$.

$$\Longrightarrow \boxed{z(p) = \begin{cases} (\ b + 2\sqrt{ac}\ )p & \text{if } 0 \leq p \leq \sqrt{c/a} \\ ap^2 + bp + c & \text{if } \sqrt{c/a} \leq p \leq \bar{p}_{max} \end{cases}}$$

- $z(p)$ convex differentiable piecewise-quadratic with ≤ 2 pieces

Solving the PR: specialized algorithm

- In continuous relaxation, $u_{ij} = p_{ij} / \bar{p}_{ij} \Longrightarrow$ Very weak bound
- PR improves it but destroys structure (knapsack, flow)
- P^2R for Sensor Placement: Separable Convex Quadratic Knapsack with at most twice the variables $\equiv$ easy to solve[11]

[11] F., Gorgone "A Library for Continuous Convex Separable Quadratic Knapsack Problems" EJOR, 2013

Solving the PR: specialized algorithm

- In continuous relaxation, $u_{ij} = p_{ij} / \bar{p}_{ij} \Longrightarrow$ Very weak bound
- PR improves it but destroys structure (knapsack, flow)
- P^2R for Sensor Placement: Separable Convex Quadratic Knapsack with at most twice the variables $\equiv$ easy to solve[11]
- P^2R for Network Design: Separable Convex Quadratic MCF on $G' = (N, A')$ with $|A'| \leq 2|A|$ (same nodes, duplicated arcs)

$$
\begin{array}{lll}
\min & \sum_{(i,j) \in A'} b'_{ij} r_{ij} + a'_{ij} r_{ij}^2 & \\
& \sum_{(j,i) \in A'} r_{ji} - \sum_{(i,j) \in A'} r_{ij} = b_i & i \in N \\
& 0 \leq r_{ij} \leq \bar{r}_{ij} & (i,j) \in A'
\end{array}
$$

[11] F., Gorgone "A Library for Continuous Convex Separable Quadratic Knapsack Problems" EJOR, 2013

Computational results: Sensor Placement

name	time	nodes	t/n	time	nodes	t/n	gap
		P^2/R			CPLEX		
2000-h	0.39	1	0.39	1020.51	223293	0.01	4.03
2000-l	0.09	1	0.09	101.58	3713	0.03	0.00
3000-h	0.92	1	0.92	1057.09	144406	0.01	7.18
3000-l	0.21	1	0.21	270.49	5724	0.05	0.00
PTN-2000	0.43	1	0.43	1018.13	4149	0.25	2.98
PTN-3000	1.02	1	1.02	1008.42	568	1.79	3.14
		P/C			CP		
2000-h	47.74	924	30.43	1066.02	507	2.11	207.04
2000-l	17.02	1	17.02	49.32	38	7.60	0.00
3000-h	91.24	88	74.09	1069.73	332	3.24	412.54
3000-l	40.27	1	40.27	135.95	72	12.08	0.00
PTN-2000	94.30	6	56.93	23.79	1	23.80	0.00
PTN-3000	202.63	6	114.72	53.74	1	53.74	0.00

- Randomly generated instances + “hard” Partition instances
- P^2/R a lot faster despite rudimentary, hand-made B&B (max. 1000s)

Computational results: Network Design

- Randomly generated Network Design instances:
  - standard DIMACS random MCF problem generator (`netgen`)
  - linear costs randomly generated in a given interval
  - quadratic/fixed costs randomly generated with respect to linear costs (two different ways, "h" – high and "l" – low)
- Not as difficult as expected (many solved at root node)
- 2-pieces linear/quadratic P^2/R solved with `CPLEX-qpopt`
  $\Longrightarrow$ not really a specialized algorithm
- Using "true" MCF solver possible with further piecewise-linearization but static version not competitive (dynamic may be)
- P^2/R very efficient (but helped by few branching/cuts)
- Non P/R B&C and Conic Program formulation of P/R much slower

Network Design — the table

name	time	nodes	t/n	time	nodes	t/n	gap
	P^2/R			CPLEX			
2000-h-h	0.10	1	0.10	690.09	101868	0.11	0.00
2000-h-l	45.42	278	1.10	1031.75	141485	0.01	0.06
2000-l-h	0.09	1	0.09	858.22	131954	0.03	0.00
2000-l-l	8.78	63	0.10	1036.79	140877	0.01	0.04
3000-h-h	0.15	1	0.15	1041.96	88541	0.01	0.00
3000-h-l	71.02	269	0.17	1051.93	73591	0.01	0.12
3000-l-h	0.15	1	0.15	988.74	89209	0.12	0.00
3000-l-l	19.05	79	0.16	1062.45	85878	0.01	0.04
	P/C			CP			
2000-h-h	57.09	7	13.84	895.70	8	207.60	0.01
2000-h-l	51.60	348	0.72	252.98	36	27.65	0.00
2000-l-h	42.30	6	16.57	525.35	9	63.35	0.00
2000-l-l	20.60	131	0.51	252.82	193	40.02	0.00
3000-h-h	117.30	11	18.90	564.41	2	407.97	0.01
3000-h-l	140.47	584	1.39	366.95	27	36.76	0.00
3000-l-h	101.18	12	12.01	372.16	4	89.53	0.01
3000-l-l	45.43	153	0.89	292.41	83	62.39	0.00

Application 7: Controlled Tabular Adjustment

- Statistical table, $\mathcal{N}$ cells, data d satisfies system $Ad = b$
- Find some perturbation z such that:
  - $A(d + z) = b \iff Az = 0$
  - for sensitive cells $i \in \mathcal{S} \subseteq \mathcal{N}$, either $z_i \geq u_i$ or $z_i \leq -l_i$
  - minimize some (weighted) $||z||$
- Positive and negative perturbation variables: $z_i = z_i^+ - z_i^-$ for $i \in \mathcal{N}$
- Obvious MIQP formulation (given weights w)

$$\begin{array}{lll}
\min & \sum_{i \in \mathcal{N}} w_i (z_i^+ - z_i^-)^2 & \\
& A(z^+ - z^-) = 0 & \\
& 0 \leq z_i^+ \leq \bar{u}_i \;,\; 0 \leq z_i^- \leq \bar{l}_i & i \in \mathcal{U} = \mathcal{N} \setminus \mathcal{S} \\
& \left.\begin{array}{l} u_i y_i^+ \leq z_i^+ \leq \bar{u}_i y_i^+ \;,\; l_i y_i^- \leq z_i^- \leq \bar{l}_i y_i^- \\ y_i^+ + y_i^- = 1 \;,\; y_i^\pm \in \{0, 1\} \end{array}\right\} & i \in \mathcal{S}
\end{array}$$

- Disjunctive constraints, mutually exclusive semicontinuous variables

Perspective Reformulation

- PR not directly applicable: $w_i(z_i^+ - z_i^-)^2 \Longrightarrow w_i\big((z_i^+)^2 + (z_i^-)^2\big)$

$$\min \textstyle\sum_{i \in \mathcal{U}} \; w_i\big(\, (z_i^+)^2 + (z_i^-)^2 \,\big) + \textstyle\sum_{i \in \mathcal{S}} \; w_i \left(\frac{(z_i^+)^2}{y_i^+} + \frac{(z_i^-)^2}{y_i^-} \right)$$

Perspective Reformulation

- PR not directly applicable: $w_i(z_i^+ - z_i^-)^2 \Longrightarrow w_i\big((z_i^+)^2 + (z_i^-)^2\big)$

$$\min \textstyle\sum_{i \in \mathcal{U}} w_i\big(\, (z_i^+)^2 + (z_i^-)^2 \,\big) + \sum_{i \in \mathcal{S}} w_i \left(\frac{(z_i^+)^2}{y_i^+} + \frac{(z_i^-)^2}{y_i^-} \right)$$

- (Trivial) theorem: substitute $y_i = y_i^+$ and $1 - y_i = y_i^-$

$$f(\, z^+ \,,\, z^- \,,\, y \,) = \begin{cases} w(z^+ - z^-)^2 & \text{if } u \le z^+ \le \bar{u},\ z^- = 0,\ y = 1 \\ w(z^+ - z^-)^2 & \text{if } l \le z^- \le \bar{l},\ \ z^+ = 0,\ y = 0 \\ +\infty & \text{otherwise} \end{cases}$$

$$h(\, z^+ \,,\, z^- \,,\, y \,) = w \left(\frac{(z^+)^2}{y} + \frac{(z^-)^2}{1-y} \right)$$

for $uy \le z^+ \le \bar{u}y$, $l(1-y) \le z^- \le \bar{l}(1-y)$, $y \in [0,1]$

Perspective Reformulation

- PR not directly applicable: $w_i(z_i^+ - z_i^-)^2 \Longrightarrow w_i\big((z_i^+)^2 + (z_i^-)^2\big)$

$$\min \textstyle\sum_{i\in\mathcal{U}} \; w_i\big(\,(z_i^+)^2 + (z_i^-)^2\,\big) + \textstyle\sum_{i\in\mathcal{S}} \; w_i\left(\frac{(z_i^+)^2}{y_i^+} + \frac{(z_i^-)^2}{y_i^-}\right)$$

- (Trivial) theorem: substitute $y_i = y_i^+$ and $1 - y_i = y_i^-$

$$f(\,z^+\,,\,z^-\,,\,y\,) = \begin{cases} w(z^+ - z^-)^2 & \text{if } u \leq z^+ \leq \bar{u},\, z^- = 0,\, y = 1 \\ w(z^+ - z^-)^2 & \text{if } l \leq z^- \leq \bar{l},\;\; z^+ = 0,\, y = 0 \\ +\infty & \text{otherwise} \end{cases}$$

$$h(\,z^+\,,\,z^-\,,\,y\,) = w\left(\frac{(z^+)^2}{y} + \frac{(z^-)^2}{1-y}\right)$$

for $uy \leq z^+ \leq \bar{u}y$, $l(1-y) \leq z^- \leq \bar{l}(1-y)$, $y \in [0,1]$

- Can perform the projection analysis

$$g(z^+, z^-) = \min_{y\in[0,1]} \quad w\left(\frac{(z^+)^2}{y} + \frac{(z^-)^2}{1-y}\right)$$

$$uy \leq z^+ \leq \bar{u}y \;\;,\;\; l(1-y) \leq z^- \leq \bar{l}(1-y)$$

The projected function

cond.	$g(z^+, z^-)$
$\bar{u} \leq l$	$\begin{cases} \bar{u}((z^-)^2/(\bar{u}-z^+)+z^+) & \text{if } lz^+ + \bar{u}z^- \geq \bar{u}l \\ l((z^+)^2/(l-z^-)+z^-) & \text{if } lz^+ + \bar{u}z^- \leq \bar{u}l \end{cases}$
$\bar{l} \leq u$	$\begin{cases} u((z^-)^2/(u-z^+)+z^+) & \text{if } \bar{l}z^+ + uz^- \leq \bar{l}u \\ \bar{l}((z^+)^2/(\bar{l}-z^-)+z^-) & \text{if } \bar{l}z^+ + uz^- \geq \bar{l}u \end{cases}$
$l \leq u \leq \bar{u}$	$\begin{cases} u((z^-)^2/(u-z^+)+z^+) & \text{if } l \leq z^+ + z^- \leq u \\ (z^+ + z^-)^2 & \text{if } u \leq z^+ + z^- \leq \bar{u} \\ \bar{u}((z^-)^2/(\bar{u}-z^+)+z^+) & \text{if } \bar{u} \leq z^+ + z^- \leq \bar{l} \end{cases}$
$l \leq u \leq \bar{l} \leq \bar{u}$	$\begin{cases} u((z^-)^2/(u-z^+)+z^+) & \text{if } l \leq z^+ + z^- \leq u \\ (z^+ + z^-)^2 & \text{if } u \leq z^+ + z^- \leq \bar{l} \\ \bar{l}((z^+)^2/(\bar{l}-z^-)+z^-) & \text{if } \bar{l} \leq z^+ + z^- \leq \bar{u} \end{cases}$
$u \leq l \leq \bar{l} \leq \bar{u}$	$\begin{cases} l((z^+)^2/(l-z^-)+z^-) & \text{if } u \leq z^+ + z^- \leq l \\ (z^+ + z^-)^2 & \text{if } l \leq z^+ + z^- \leq \bar{l} \\ \bar{l}((z^+)^2/(\bar{l}-z^-)+z^-) & \text{if } \bar{l} \leq z^+ + z^- \leq \bar{u} \end{cases}$
$u \leq l \leq \bar{u} \leq \bar{l}$	$\begin{cases} l((z^+)^2/(l-z^-)+z^-) & \text{if } u \leq z^+ + z^- \leq l \\ (z^+ + z^-)^2 & \text{if } l \leq z^+ + z^- \leq \bar{u} \\ \bar{u}((z^-)^2/(\bar{u}-z^+)+z^+) & \text{if } \bar{u} \leq z^+ + z^- \leq \bar{l} \end{cases}$

The projected function: wrap-up

- $g(z^+, z^-)$ is piecewise-SOCP-quadratic with at most three pieces
- Under reasonable hypotheses, the "central" one is $(z^+ + z^-)^2$
- Symmetric bounds ($\bar{u}_i = \bar{l}_i$, $l_i = u_i$) $\Longrightarrow g(z^+, z^-) = (z^+ + z^-)^2$
- $w_i(z_i^+ - z_i^-)^2 \Longrightarrow w_i\big(z_i^+ + z_i^-\big)^2$ applicable also to nonsensitive cells
- Four possible "basic" formulations:

MIQP	MIQP+
$\min \sum_{i \in \mathcal{N}} \; w_i\big(\, (z_i^+)^2 + (z_i^-)^2 \,\big)$	$\min \sum_{i \in \mathcal{N}} \; w_i\big(\, z_i^+ + z_i^- \,\big)^2$
PR	PR+
$\min \sum_{i \in \mathcal{U}} \; w_i\big((z_i^+)^2 + (z_i^-)^2\big) + \sum_{i \in \mathcal{S}} w_i\big(\frac{(z_i^+)^2}{y_i} + \frac{(z_i^-)^2}{1-y_i}\big)$	$\min \sum_{i \in \mathcal{U}} \; w_i(z_i^+ + z_i^-)^2 + \sum_{i \in \mathcal{S}} w_i\big(\frac{(z_i^+)^2}{y_i} + \frac{(z_i^-)^2}{1-y_i}\big)$

- PR and PR+ solved with P/C or CP: P/C, CP, P/C+, CP+

Computational experiments: symmetric instances

- Random one-hierarchical-two-dimensional data (1H2D), realistic size
- r-c-s(-a): rows-columns-%sensitive(-asymmetry $u_i = a \cdot l_i$), 10 each
- CP and CP+ always not competitive

Computational experiments: symmetric instances

- Random one-hierarchical-two-dimensional data (1H2D), realistic size
- r-c-s(-a): rows-columns-%sensitive(-asymmetry $u_i = a \cdot l_i$), 10 each
- CP and CP+ always not competitive

	MIQP+			P/C+			MIQP			P/C		
	gap	time	nodes	gap	time	nodes	gap	time	nodes	gap	time	nodes
10-20-3	0.01	442	474	0.00	486	357	6.49	9686	10365	0.00	1331	1973
10-20-5	0.01	765	690	0.01	1016	611	67.62	10000	2649	0.16	6695	8675
10-20-10	0.01	3852	10507	2.21	7660	2676	72.75	10000	5536	12.39	10000	3230
10-30-3	0.01	1470	760	0.01	1749	457	127.03	10000	778	0.98	9070	3022
10-30-5	0.01	4850	4003	0.07	7102	4769	118.53	10000	1422	15.80	10000	1853
10-30-10	2.44	10000	3512	8.26	10000	889	128.67	10000	1619	35.30	10000	643
20-20-3	0.00	1710	260	0.00	1874	291	158.64	10000	636	17.84	8559	596
20-20-5	0.01	3543	1507	1.27	7237	1185	138.59	10000	625	12.33	8808	481
20-20-10	7.10	10000	1968	24.51	10000	504	142.82	10000	777	38.22	10000	262
20-30-3	0.40	6113	738	3.60	6800	458	138.85	10000	726	27.17	10000	379
20-30-5	7.39	8791	751	15.19	8885	379	156.73	10000	801	32.83	10000	406
20-30-10	19.92	10000	674	32.04	10000	102	153.79	10000	496	44.06	10000	56

Computational experiments: asymmetric instances 1

	MIQP+			P/C+			MIQP			P/C		
	gap	time	nodes	gap	time	nodes	gap	time	nodes	gap	time	nodes
10-20-3-2	0.00	23	9	0.00	58	1	0.01	1218	7823	0.00	106	17
10-20-3-5	0.01	19	1	0.00	82	1	0.01	322	197	0.00	111	1
10-20-3-10	0.00	15	7	0.00	55	1	0.01	270	124	0.00	78	1
10-20-5-2	0.01	58	30	0.00	119	9	0.04	10000	113601	0.00	152	32
10-20-5-5	0.01	21	15	0.00	79	1	0.01	1293	2332	0.00	81	1
10-20-5-10	0.00	20	2	0.00	106	1	0.01	1483	660	0.00	111	1
10-20-10-2	0.01	438	556	0.00	637	181	0.04	10000	67541	1.49	2904	370
10-20-10-5	0.01	4315	31344	0.00	142	1	0.08	10000	102641	0.00	142	1
10-20-10-10	0.01	416	2135	0.00	120	1	0.04	5044	26508	0.00	109	1
10-30-3-2	0.00	115	28	0.00	271	5	0.02	10000	55266	0.00	391	35
10-30-3-5	0.00	40	4	0.00	220	1	0.01	2447	1333	0.00	237	1
10-30-3-10	0.00	31	1	0.00	232	1	0.01	1468	565	0.00	258	1
10-30-5-2	0.00	193	103	0.00	377	19	0.05	10000	28721	0.00	455	72
10-30-5-5	0.01	119	39	0.00	333	1	0.01	4055	24181	0.00	258	1
10-30-5-10	0.01	63	46	0.00	207	1	0.01	1855	1104	0.00	216	1
10-30-10-2	0.01	1158	1035	0.00	1905	230	7.03	10000	27461	0.82	3066	986
10-30-10-5	0.01	6489	38818	0.00	401	1	8.53	10000	60347	0.00	311	1
10-30-10-10	0.01	4806	22519	0.00	522	1	0.09	10000	52141	0.00	372	1

- Easier than (not really) symmetric ones (under practitioners' control)

Computational experiments: asymmetric instances 2

	MIQP+			P/C+			MIQP			P/C		
	gap	time	nodes	gap	time	nodes	gap	time	nodes	gap	time	nodes
20-20-3-2	0.00	136	25	0.00	393	1	0.03	10000	13721	0.00	502	9
20-20-3-5	0.01	72	1	0.00	625	1	0.01	4074	1207	0.00	691	1
20-20-3-10	0.00	76	1	0.00	574	1	2.18	5356	465	0.00	644	1
20-20-5-2	0.00	257	47	0.00	601	4	1.40	10000	14362	0.00	598	24
20-20-5-5	0.01	117	10	0.00	690	1	1.19	10000	15635	0.00	638	1
20-20-5-10	0.01	128	54	0.00	736	1	0.52	6434	2076	0.00	623	1
20-20-10-2	0.01	1448	212	0.00	2802	138	63.41	10000	1006	0.00	2525	228
20-20-10-5	0.02	9203	22462	0.00	943	1	3.40	10000	9950	0.00	634	1
20-20-10-10	0.03	7910	19421	0.00	1327	1	7.33	10000	9801	0.00	801	1
20-30-3-2	0.01	439	28	0.00	1477	1	13.94	10000	1203	0.00	1649	16
20-30-3-5	0.01	140	1	0.00	1597	1	5.39	8400	1767	0.00	1510	1
20-30-3-10	0.00	157	8	0.00	1601	1	8.34	9321	691	0.00	1547	1
20-30-5-2	0.00	777	65	0.00	2160	17	48.34	10000	612	0.00	2111	34
20-30-5-5	0.01	618	462	0.00	1800	1	19.74	10000	1692	0.00	1622	1
20-30-5-10	0.01	622	243	0.00	1988	1	2.14	9815	2623	0.00	1625	1
20-30-10-2	1.23	7575	1454	3.67	8407	297	79.80	10000	422	4.16	7705	262
20-30-10-5	0.52	10000	12890	0.00	2784	1	36.91	10000	718	0.00	1915	1
20-30-10-10	0.04	10000	17526	0.00	2619	1	27.08	10000	1441	0.00	1817	1

- MIQP+ better except for many sensitive cells & very asymmetric

Approximated P^2R: Project&Lift

- Approximating the PR may work
- Biggest roadblocks in P^2R:
  - separability of u (UC does not have it)
  - need for a special structure to be exploited (MV does not have it)
  - no $u \Longrightarrow$ home-made branch, cut, fixing, ...
- But for non-separable u, computing the projection is too complicated

Approximated P^2R: Project&Lift

- Approximating the PR may work
- Biggest roadblocks in P^2R:
  - separability of u (UC does not have it)
  - need for a special structure to be exploited (MV does not have it)
  - no $u \Longrightarrow$ home-made branch, cut, fixing, ...
- But for non-separable u, computing the projection is too complicated
- Possible solution: project as if u were separable, re-introduce them

Approximated P^2R: Project&Lift

- Approximating the PR may work
- Biggest roadblocks in P^2R:
  - separability of u (UC does not have it)
  - need for a special structure to be exploited (MV does not have it)
  - no $u \Longrightarrow$ home-made branch, cut, fixing, ...
- But for non-separable u, computing the projection is too complicated
- Possible solution: project as if u were separable, re-introduce them
- Meanwhile, a few useful generalizations:
  - Extend to nonquadratic f provided that for fixed $g^{\pm}$

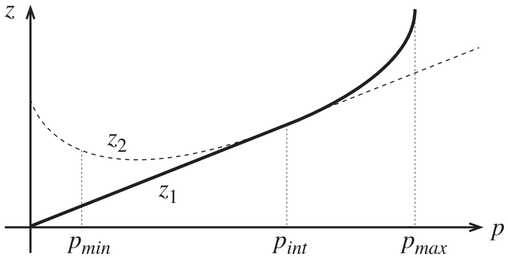
$$\tilde{u}(p) = \begin{cases} pg^{+} & \text{if } p \geq 0 \\ -pg^{-} & \text{if } p \leq 0 \end{cases} \tag{22}$$

    is the unique stationary point of $h(p, u)$ with respect to u
    (many cases: p^k, e^p, Kleinrock delay function, ...)
  - Extend to $\bar{p}_{min} < 0$ (4-pieces z instead of 2-pieces one)

Project&Lift

- The general form ($\bar{p}_{min} \geq 0$): for $p_{int} \in \{ \ \bar{p}_{min} \ , \ 1/g^+ \ , \ \bar{p}_{max} \ \}$

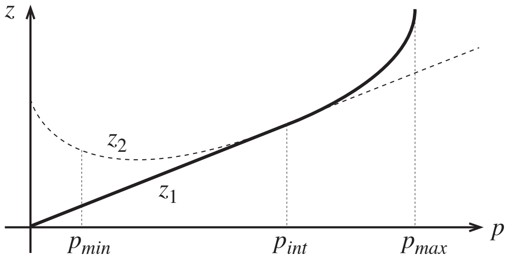
$$z(p) = \begin{cases} z_1(p) = \big(\ b + f(p_{int})/p_{int} + c/p_{int} \ \big) p & 0 \leq p \leq p_{int} \\ z_2(p) = f(p) + bp + c & p_{int} \leq p \leq \bar{p}_{max} \end{cases}$$



Project&Lift

- The general form ($\bar{p}_{min} \geq 0$): for $p_{int} \in \{ \ \bar{p}_{min} \ , \ 1/g^+ \ , \ \bar{p}_{max} \ \}$

$$z(p) = \begin{cases} z_1(p) = \left(\ b + f(p_{int})/p_{int} + c/p_{int} \ \right) p & 0 \leq p \leq p_{int} \\ z_2(p) = f(p) + bp + c & p_{int} \leq p \leq \bar{p}_{max} \end{cases}$$



- Theorem: can be written as a NLP with the u "making" z_1:

$$z(p) = \begin{cases} \min \ h(u,q) = & uz(p_{int}) + z_2(q + p_{int}) - z(p_{int}) \\ & (\bar{p}_{min} - p_{int})u \leq q \leq (\bar{p}_{max} - p_{int})u \\ & p = p_{int}u + q \quad , \quad u \in [0,1] \end{cases}$$

Approximated Projected Perspective Reformulation

- With the integrality constraints, a reformulation of the block

$$\begin{aligned} \min \quad & u z(p_{int}) + z_2(q + p_{int}) - z(p_{int}) \\ & (\bar{p}_{min} - p_{int}) u \leq q \leq (\bar{p}_{max} - p_{int}) u \\ & p = p_{int} u + q \qquad , \qquad u \in \{0, 1\} \end{aligned}$$

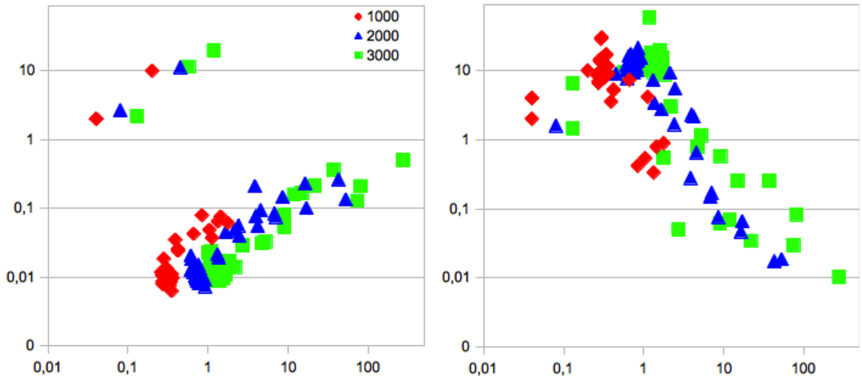
just use this instead of the original constraints

- Not as strong as PR, but same number of variables and easy to do
- 4-pieces version for $\bar{p}_{min} < 0$ (duplicate u)
- Just properly translating one variable improves the LB: blatant violation of the no-free-lunch principle!
- Funny observation: $f(p) = ap^2$, just redefine $p = p_{int} u + q$ and use

$$u^2 = u \qquad\qquad qu = q$$

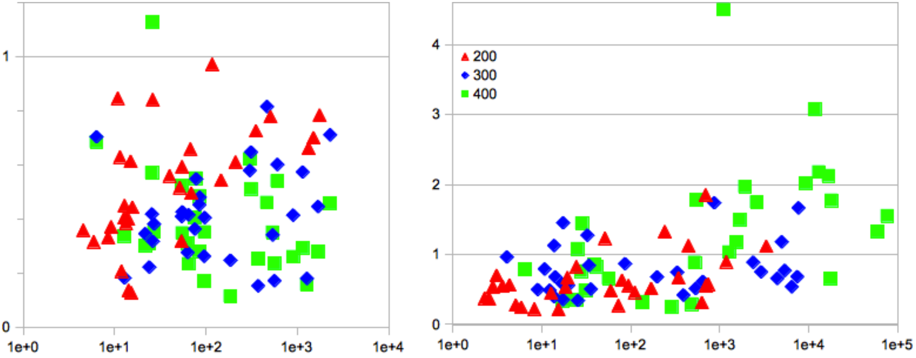
valid because $u \in \{0, 1\}$ and strengthening (RLT?)

Computational results: ND



- left: (AP^2R time) / (P/C time) plotted against AP^2R time
- right = (AP^2R time) / (P^2R time) plotted against AP^2R time

Computational results: MV



- (AP^2R time) / (P/C time) plotted against AP^2R time
- right = (tight) constraint on max assets (10), not separable

Computational results: UC

		AP^{2}R									
		NCNH		CNH		CH		B&C			
t	h	time	dgap	time	dgap	time	pgap	nodes	time	pgap	gap
10	0	0.31	1.49	1.50	0.29	2.59	0.04	1229	284	0.00	0.01
20	0	0.75	1.25	6.97	0.29	13.99	0.15	4635	9999	0.01	0.17
50	0	3.19	1.19	50.53	0.22	65.12	0.23	1078	9999	0.02	0.23
20	10	0.88	0.58	3.04	0.15	7.91	0.16	16477	1078	0.00	0.01
50	20	2.91	0.58	18.97	0.09	28.33		3780	9999	0.00	0.07
75	35	5.46	0.49	39.18	0.06	45.73		1727	9999	0.03	0.08
		PC									
10	0	0.17	1.48	0.99	0.23	1.25	0.40	365	17	0.00	0.01
20	0	0.49	1.24	3.93	0.25	5.38		15607	4851	0.00	0.02
50	0	2.85	1.16	16.59	0.19	20.63		14286	9986	0.00	0.13
20	10	0.52	0.56	1.92	0.13	3.14	0.51	8107	240	0.00	0.01
50	20	2.05	0.57	6.17	0.07	13.11		66945	6649	0.00	0.02
75	35	4.19	0.48	11.23	0.05	20.22	0.08	57456	9999	0.00	0.02

- No Cuts No Heuristic, Cuts No Heuristic, Cuts & Heuristic, B&C
- Sometimes a better root node solution is found, all the rest is bad

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999

[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013

[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective
- The extra nonlinearity can often be "projected away"

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999
[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013
[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective
- The extra nonlinearity can often be "projected away"
- Basically, no reason not to use it (can even be simpler): should be the formulation, rather than "a reformulation"

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999
[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013
[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective
- The extra nonlinearity can often be "projected away"
- Basically, no reason not to use it (can even be simpler): should be the formulation, rather than "a reformulation"
- Still lots of work to do in PR (nonseparable, . . .), several clever people did[12]/does it[13,14]

---

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999

[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013

[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective
- The extra nonlinearity can often be "projected away"
- Basically, no reason not to use it (can even be simpler):
  should be the formulation, rather than "a reformulation"
- Still lots of work to do in PR (nonseparable, ...),
  several clever people did[12]/does it[13,14]
- MINLP a nice place to be in:
  nonlinear + combinatorial = many useful structures

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999

[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013

[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010

Conclusions

- Perspective Reformulation is very simple, yet surprisingly effective
- The extra nonlinearity can often be "projected away"
- Basically, no reason not to use it (can even be simpler): should be the formulation, rather than "a reformulation"
- Still lots of work to do in PR (nonseparable, ...), several clever people did[12]/does it[13,14]
- MINLP a nice place to be in: nonlinear + combinatorial = many useful structures
- Enough work to be back for ClovisFest 80++ :-)

---

[12] Stubbs, Mehrotra "A branch-and-cut method for 0-1 mixed convex programming" *Math. Prog.*, 1999

[13] Khajavirad, Sahinidis "Convex envelopes generated from finitely many compact convex sets" *Math. Prog.* 2013

[14] Luedtke, Namazifar, Linderoth "Some results on the strength of relaxations of multilinear functions" *TR UW-Madison*, 2010