Hybrid Lagrangian-MILP Approaches to Unit Commitment: Challenges and Opportunities

Antonio Frangioni

Dipartimento di Informatica, Università di Pisa

with C. Gentile and M. De Santis

PGMO's Days 2013

October 4, 2013

Outline

The (basic) Hydro-thermal Unit Commitment problem

- Operate a set of generating units (P thermal, H hydro) over a discretized time horizon $\mathcal{T}$ so as to satisfy (forecasted) demand
- Simple transmission constraints (bus), spinning reserve constraints

[1] Nowak, Römisch "Stochastic Lagrangian Relaxation Applied to Power Scheduling in a Hydro-thermal System Under Uncertainty", *Annals of Operations Research*, 2000

The (basic) Hydro-thermal Unit Commitment problem

- Operate a set of generating units (P thermal, H hydro) over a discretized time horizon $\mathcal{T}$ so as to satisfy (forecasted) demand
- Simple transmission constraints (bus), spinning reserve constraints
- A MIQP formulation: main variables
  - $u^i_t \in \{0, 1\}$: ON/OFF state of thermal unit $i \in P$
  - $p^i_t \in \mathbb{R}_+$: power level of thermal unit $i \in P$
  - $q^j_t \in \mathbb{R}_+$: water discharge for hydro unit $j \in H(h)$ for cascade $h \in H$
- Objective function:

$$f(p, u) = \sum_{i \in P} c^i(p^i, u^i) = \sum_{i \in P} \left(s^i(u^i) + \sum_{t \in \mathcal{T}} \left(a^i_t (p^i_t)^2 + b^i_t p^i_t + c^i_t u^i_t \right) \right) \quad (1)$$

  - nonlinear convex energy cost ($a^i_t > 0$), fixed costs
  - time-dependent start-up costs $s^i(u^i)$ (only a few extra constraints and continuous variables with nifty formulation[1])

[1] Nowak, Römisch "Stochastic Lagrangian Relaxation Applied to Power Scheduling in a Hydro-thermal System Under Uncertainty", *Annals of Operations Research*, 2000

A MIQP formulation of UC (2)

Thermal units:

- Maximum and minimum power output:

$$\bar{p}^i_{min} u^i_t \leq p^i_t \leq \bar{p}^i_{max} u^i_t \qquad t \in \mathcal{T} \qquad (2)$$

- Ramp-up constraints (Δ^i_+ = ramp-up threshold):

$$p^i_t \leq p^i_{t-1} + u^i_{t-1}\Delta^i_+ + (1 - u^i_{t-1})\bar{l}^i \qquad t \in \mathcal{T} \qquad (3)$$

- Ramp-down constraints (Δ^i_- = ramp-down threshold):

$$p^i_{t-1} \leq p^i_t + u^i_t\Delta^i_- + (1 - u^i_t)\bar{u}^i \qquad t \in \mathcal{T} \qquad (4)$$

- Min up-time constraints (τ^i_+ = min up-time):

$$u^i_t \leq 1 - u^i_{r-1} + u^i_r \qquad t \in \mathcal{T},\ r \in [t - \tau^i_+, t - 1] \qquad (5)$$

- Min down-time constraints (τ^i_- = min down-time):

$$u^i_t \geq 1 - u^i_{r-1} - u^i_r \qquad t \in \mathcal{T},\ r \in [t - \tau^i_-, t - 1] \qquad (6)$$

A MIQP formulation of UC (3)

Hydro units:

- Maximum discharge:

$$0 \leq q_t^j \leq \bar{q}_{max}^j \qquad t \in \mathcal{T} \qquad (7)$$

- Maximum and minimum reservoir volume:

$$\bar{v}_{min}^j \leq v_t^j \leq \bar{v}_{max}^j \qquad t \in \mathcal{T} \qquad (8)$$

- Water conservation ($\bar{w}_t^j$ = inflow, w_t^j = spillage, t_{kj} = time delay):

$$v_t^j - v_{t-1}^j = \bar{w}_t^j - w_t^j - q_t^j + \sum_{k \in \mathcal{S}(j)} \left(q_{t-t_{kj}}^k + w_{t-t_{kj}}^k \right) \qquad t \in \mathcal{T} \quad (9)$$

A MIQP formulation of UC (3)

Hydro units:

- Maximum discharge:

$$0 \leq q^j_t \leq \bar{q}^j_{max} \qquad t \in \mathcal{T} \qquad (7)$$

- Maximum and minimum reservoir volume:

$$\bar{v}^j_{min} \leq v^j_t \leq \bar{v}^j_{max} \qquad t \in \mathcal{T} \qquad (8)$$

- Water conservation ($\bar{w}^j_t$ = inflow, w^j_t = spillage, t_{kj} = time delay):

$$v^j_t - v^j_{t-1} = \bar{w}^j_t - w^j_t - q^j_t + \sum_{k \in \mathcal{S}(j)} \left(q^k_{t-t_{kj}} + w^k_{t-t_{kj}} \right) \qquad t \in \mathcal{T} \quad (9)$$

System-wide constraints:

- Demand satisfaction (α^j = constant power-to-discharged water):

$$\sum_{i \in P} p^i_t + \sum_{h \in H} \sum_{j \in H(h)} \alpha^j q^j_t = \bar{d}_t \qquad t \in \mathcal{T} \qquad (10)$$

Lagrangian Relaxation

- Large-scale, difficult problem, but amenable to decomposition

[2] F. "Generalized Bundle Methods" *SIAM Journal on Optimization*, 2002

- Large-scale, difficult problem, but amenable to decomposition
- Denote (2)—(6) as $\mathcal{U}^i$, (7)—(9) as $\mathcal{H}^h$
- Lagrangian Relaxation of demand constraints (10), multipliers λ
- The problem decomposes by unit:

$$\phi(\lambda) = \sum_{i \in P} \phi^1_i(\lambda) + \sum_{h \in H} \phi^2_h(\lambda) + \sum_{t \in \mathcal{T}} \lambda_t \bar{d}_t$$

$$\phi^1_i(\lambda) = \min \left\{ c^i(p^i, u^i) - \lambda p^i \; : \; (p^i, u^i) \in \mathcal{U}^i \right\} \quad (11)$$

$$\phi^2_h(\lambda) = \min \left\{ -\lambda \sum_{j \in H(h)} \alpha^j q^j \; : \; [q^j]_{j \in H(h)} \in \mathcal{H}^h \right\} \quad (12)$$

[2] F. "Generalized Bundle Methods" *SIAM Journal on Optimization*, 2002

Lagrangian Relaxation

- Large-scale, difficult problem, but amenable to decomposition
- Denote (2)—(6) as $\mathcal{U}^i$, (7)—(9) as $\mathcal{H}^h$
- Lagrangian Relaxation of demand constraints (10), multipliers λ
- The problem decomposes by unit:

$$\phi(\lambda) = \sum_{i \in P} \phi^1_i(\lambda) + \sum_{h \in H} \phi^2_h(\lambda) + \sum_{t \in \mathcal{T}} \lambda_t \bar{d}_t$$

$$\phi^1_i(\lambda) = \min \left\{ c^i(p^i, u^i) - \lambda p^i \ : \ (p^i, u^i) \in \mathcal{U}^i \right\} \quad (11)$$

$$\phi^2_h(\lambda) = \min \left\{ -\lambda \sum_{j \in H(h)} \alpha^j q^j \ : \ [q^j]_{j \in H(h)} \in \mathcal{H}^h \right\} \quad (12)$$

- Lagrangian Dual:

$$\max \{ \phi(\lambda) \ : \ \lambda \in \mathbb{R}^n \} \quad (13)$$

($n = O(|\mathcal{T}|)$) efficiently solvable e.g. by Bundle methods[2], provided that $\phi(\lambda)$ (and a subgradient) is efficiently computable

[2] F. "Generalized Bundle Methods" *SIAM Journal on Optimization*, 2002

Solving the Lagrangian Problem

- Hydro Single-Unit Subproblems (12): Network Flow algorithms
- Thermal Single-Unit Subproblems (11): Dynamic Programming
  - easy if there are no ramping constraints (3)–(4) (trick: optimal power level if unit on computable a-priori, known for 15+ years)
  - more complicated but doable with (3)–(4): $O(n^2)$ nodes, $O(n^4)$ arcs but structured into levels, $O(n^3)$ overall with specialized DP for ED[3] quite a lot faster than a general-purpose solver

	DP		CPLEX			
h	time	st.dev.	time	st.dev.	gap%	fail
24	.002	4e-3	0.08	0.05		0
96	0.04	3e-3	32.21	76.12	0.03	6
168	0.20	6e-3	117.46	141.87	1.23	35

- + fast convergence of Bundle methods & heuristics or Augmented-Lagrangian-type stuff $\Longrightarrow$ good solutions quickly enough[4]

[3] F., Gentile "Solving Nonlinear Single-Unit Commitment Problems with Ramping Constraints" *Op. Res.*, 2006

[4] F., Gentile, Lacalandra "Solving Unit Commitment Problems with General Ramp Contraints", *IJEPES*, 2008

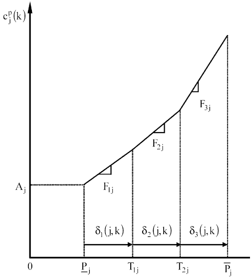
A traditional MILP formulation

- Traditionally MIQP has been deemed much harder than MILP, which suggest the standard piecewise-linearization[5]

[5] Carrión, Arroyo "A Computationally Efficient Mixed-integer Linear Formulation for the Thermal Unit Commitment Problem" *IEEE Transactions on Power Systems*, 2006

A traditional MILP formulation

- Traditionally MIQP has been deemed much harder than MILP, which suggest the standard piecewise-linearization[5]



- k new variables (i, t fixed), redefine p

$$\begin{aligned} & p = \textstyle\sum_{l=1}^{k} \delta_l + \bar{p}_{min} u \\ & 0 \leq \delta_l \leq \bar{p}^l - \bar{p}^{l-1} \qquad l = 1, \ldots, k \end{aligned} \tag{14}$$

- cost coefficient of u set to $f(\bar{p}_{min})$

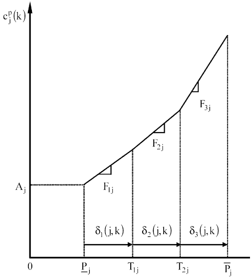
- cost coefficient of each δ_l set to

$$F_l = \frac{f(\bar{p}^l) - f(\bar{p}^{l-1})}{\bar{p}^l - \bar{p}^{l-1}} = a(\bar{p}^l + \bar{p}^{l-1}) + b$$

[5] Carrión, Arroyo "A Computationally Efficient Mixed-integer Linear Formulation for the Thermal Unit Commitment Problem" *IEEE Transactions on Power Systems*, 2006

A traditional MILP formulation

- Traditionally MIQP has been deemed much harder than MILP, which suggest the standard piecewise-linearization[5]



- k new variables (i, t fixed), redefine p

$$p = \sum_{l=1}^{k} \delta_l + \bar{p}_{min} u$$
$$0 \leq \delta_l \leq \bar{p}^l - \bar{p}^{l-1} \qquad l = 1, \dots, k \tag{14}$$

- cost coefficient of u set to $f(\bar{p}_{min})$
- cost coefficient of each δ_l set to

$$F_l = \frac{f(\bar{p}^l) - f(\bar{p}^{l-1})}{\bar{p}^l - \bar{p}^{l-1}} = a(\bar{p}^l + \bar{p}^{l-1}) + b$$

- How does it compare with LR at low accuracy (0.5%)?

[5] Carrión, Arroyo "A Computationally Efficient Mixed-integer Linear Formulation for the Thermal Unit Commitment Problem" *IEEE Transactions on Power Systems*, 2006

Comparing MILP and LR

		LR			Cplex MILP					
p	h	time	gap	iter	time	gap	ftime	fgap	nodes	LPs
10	0	0.75	0.99	187	0.95	0.33		1.18	0	23
20	0	1.83	0.46	189	3.72	0.36		1.00	0	23
50	0	4.84	0.28	195	21.93	0.21	15.98	0.36	0	25
75	0	9.41	0.34	206	56.31	0.20	47.08	1.62	10	59
100	0	14.74	0.33	213	94.09	0.17	69.75	2.18	16	76
150	0	21.20	0.17	277	218.69	0.12	177.35	6.58	16	115
200	0	34.80	0.09	317	267.78	0.09	247.12	1.85	6	87
20	10	1.76	0.39	170	93.53	0.21		0.59	140	258
50	20	6.36	0.06	160	17.98	0.06	17.98	0.06	0	60
75	35	15.01	0.04	198	96.86	0.11	96.86	0.11	170	300
100	50	24.74	0.04	209	130.86	0.06	130.86	0.06	180	266
150	75	37.41	0.02	189	467.62	0.06	467.62	0.06	300	554
200	100	50.91	0.01	175	427.71	0.05	427.71	0.05	205	321

- Gaps computed agains much stronger Lagrangian bound
- LP gap vastly worse but Cplex primal heuristic impressively effective
- Can we do better?

Perspective Reformulation for General MINLPs

- Convex function f, Mixed-Integer NonLinear Program fragment

$$\min \left\{ \; f(p) + cu \; : \; Ap \leq bu \; , \; u \in \{0,1\} \; \right\} \tag{15}$$

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

[6] Tawarmalani, Sahinidis "Convex extensions and envelopes of l.s.c. functions" *Math. Prog.*, 2002

Perspective Reformulation for General MINLPs

- Convex function f, Mixed-Integer NonLinear Program fragment

$$\min \{ \, f(p) + cu \; : \; Ap \leq bu \, , \; u \in \{0,1\} \, \} \tag{15}$$

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

- Equivalently, minimize the nonconvex function

$$f(p,u) = \begin{cases} 0 & \text{if } u = 0 \text{ and } p = 0 \\ f(p) + c & \text{if } u = 1 \text{ and } Ap \leq b \\ +\infty & \text{otherwise} \end{cases} \tag{16}$$

[6] Tawarmalani, Sahinidis "Convex extensions and envelopes of l.s.c. functions" *Math. Prog.*, 2002

Perspective Reformulation for General MINLPs

- Convex function f, Mixed-Integer NonLinear Program fragment

$$\min \{ \, f(p) + cu \ : \ Ap \leq bu \, , \ u \in \{0,1\} \, \} \tag{15}$$

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

- Equivalently, minimize the nonconvex function

$$f(p,u) = \begin{cases} 0 & \text{if } u = 0 \text{ and } p = 0 \\ f(p) + c & \text{if } u = 1 \text{ and } Ap \leq b \\ +\infty & \text{otherwise} \end{cases} \tag{16}$$

- Best possible convex relaxation of (15): use the convex envelope[6]

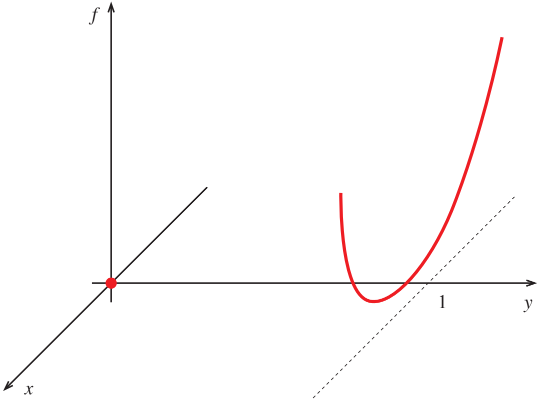
$$h(p,u) = \begin{cases} 0 & \text{if } p = 0 \text{ and } u = 0, \\ uf(p/u) + cu & \text{if } Ap \leq bu, \ \ u \in (0,1], \\ +\infty & \text{otherwise.} \end{cases} \tag{17}$$

(convex function minorizing $f(p,u)$ with smallest possible epigraph)

[6] Tawarmalani, Sahinidis "Convex extensions and envelopes of l.s.c. functions" *Math. Prog.*, 2002

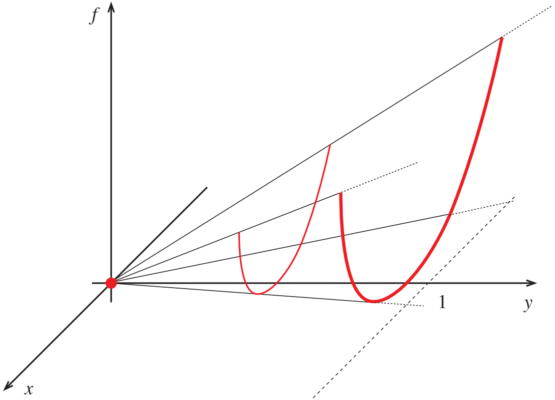
The Perspective what?

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



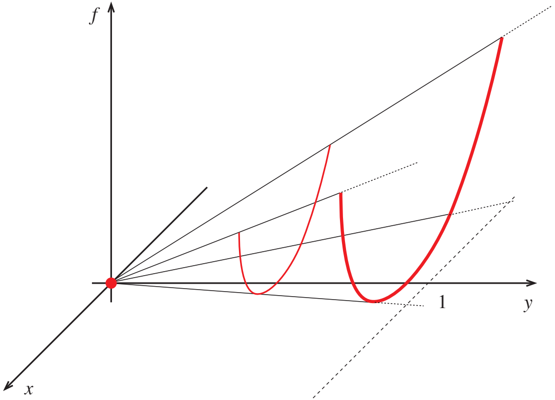
The Perspective what?

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



The Perspective what?

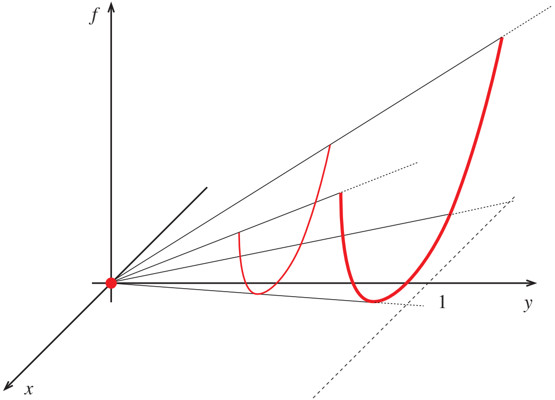
- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



- $h(p, u)$ convex but "more nonlinear" than $f(p) + cu$
  example: $f(p) = ap^2 + bp \quad \Rightarrow \quad h(p, u) = (a/u)p^2 + bp + cu$

The Perspective what?

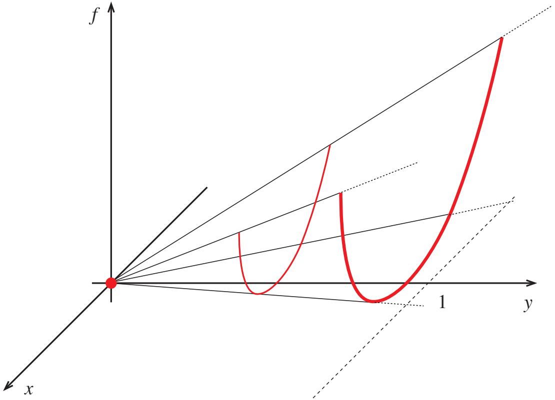
- $h(p, u)$ is a section of the perspective function $f(x, \lambda) = \lambda f(x/\lambda)$



- $h(p, u)$ convex but "more nonlinear" than $f(p) + cu$
  example: $f(p) = ap^2 + bp \quad \Rightarrow \quad h(p, u) = (a/u)p^2 + bp + cu$
- 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 "more nonlinear" than $f(p) + cu$
  example: $f(p) = ap^2 + bp \quad \Rightarrow \quad h(p, u) = (a/u)p^2 + bp + cu$
- Notes: I) $a/u > a$ for $u < 1$; II) for $a = 0$ nothing happens

The Perspective Relaxation (Reformulation)

- A convex, continuous, more nonlinear program

$$\min \left\{ \; uf(p/u) + cu \;\; : \;\; Ap \leq bu \; , \;\; u \in \{0,1\} \; \right\} \qquad (18)$$

$u \in \{0,1\}^n \Rightarrow$ a reformulation! (if $0f(0/0) = 0$)

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

The Perspective Relaxation (Reformulation)

- A convex, continuous, more nonlinear program

$$\min \{ \ uf(p/u) + cu \ : \ Ap \leq bu \ , \ u \in \{0,1\} \ \} \tag{18}$$

$u \in \{0,1\}^n \Rightarrow$ a reformulation! (if $0f(0/0) = 0$)

- Lower bound much better, but how to solve it efficiently?

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

The Perspective Relaxation (Reformulation)

- A convex, continuous, more nonlinear program

$$\min \{ \; uf(p/u) + cu \;:\; Ap \leq bu \;,\; u \in \{0,1\} \; \} \qquad (18)$$

$u \in \{0,1\}^n \Rightarrow$ a reformulation! (if $0f(0/0) = 0$)

- Lower bound much better, but how to solve it efficiently?

- 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

$$v \geq f(\bar{p}) + c + [s \;,\; c + f(\bar{p}) - s\bar{p}] \begin{bmatrix} p - \bar{p} \\ u - 1 \end{bmatrix} \qquad \begin{array}{c} \forall \bar{p} \in \mathcal{P} \\ \forall s \in \partial f(\bar{p}) \end{array} \qquad (19)$$

- The quadratic case: Perspective Cuts[7]

$$v \geq (2a\bar{p} + b)p + (c - a\bar{p}^2)u \qquad \forall \bar{p} \in \mathcal{P} \qquad (20)$$

infinitely many, but add them dynamically with a `cutcallback`

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

Impact of the Perspective Relaxation (P/C)

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
- Small instances ($p = 20$), no ramp constraints, gap $= 0.1\%$
- Time limit 10000s (unrealistic)

Comparing static formulations at lower accuracy (0.5%)

		$SPWF_4$				SP/C_4			
p	h	gap	nd	time	rgap	gap	nd	time	rgap
10	0	0.31	0	0.95	1.61	0.28	0	0.76	1.50
20	0	0.34	0	3.72	1.34	0.36	8	3.56	1.25
50	0	0.21	0	21.93	1.38	0.21	0	12.09	1.26
75	0	0.20	10	56.31	1.43	0.18	14	45.88	1.30
100	0	0.17	16	94.09	1.39	0.15	0	43.55	1.27
150	0	0.12	16	218.69	1.32	0.11	2	146.80	1.20
200	0	0.09	6	267.78	1.37	0.08	0	234.97	1.25
20	10	0.21	140	93.53	0.82	0.20	0	3.71	0.69
50	20	0.06	0	17.98	0.70	0.10	0	18.93	0.63
75	35	0.11	170	96.86	0.57	0.07	70	64.52	0.52
100	50	0.06	180	130.86	0.58	0.07	35	81.41	0.53
150	75	0.06	300	467.62	0.58	0.05	90	293.50	0.52
200	100	0.05	205	427.71	0.56	0.03	35	314.00	0.51

- P/C clearly better, both can be done at algebraic modeling level

Static vs. dynamic formulations at lower accuracy (0.5%)

		SP/C_4			P/C_4			P/C_∞		
p	h	gap	nd	time	gap	nd	time	gap	nd	time
10	0	0.28	0	0.76	0.30	0	0.86	0.28	0	0.80
20	0	0.36	8	3.56	0.36	0	2.51	0.33	0	3.00
50	0	0.21	0	12.09	0.19	0	14.17	0.18	0	13.08
75	0	0.18	14	45.88	0.19	2	36.62	0.22	0	22.58
100	0	0.15	0	43.55	0.17	0	34.31	0.20	0	36.51
150	0	0.11	2	146.80	0.11	4	104.68	0.12	10	169.68
200	0	0.08	0	234.97	0.10	0	183.01	0.14	12	235.60
20	10	0.20	0	3.71	0.30	5	4.18	0.15	0	2.51
50	20	0.10	0	18.93	0.10	10	19.06	0.13	0	10.93
75	35	0.07	70	64.52	0.05	115	70.55	0.03	95	64.80
100	50	0.07	35	81.41	0.05	15	47.62	0.04	40	60.78
150	75	0.05	90	293.50	0.05	115	194.10	0.05	115	216.33
200	100	0.03	35	314.00	0.02	0	155.36	0.03	135	342.69

- dynamic typically better (although requires `cutcallback`)

Results with higher accuracy (0.01%)

		$SPWF_4$		SP/C_4		P/C_4		P/C_∞	
p	h	gap	time	gap	time	gap	time	gap	time
10	0	0.01	22	0.01	15	0.01	12	0.01	16
20	0	0.01	3480	0.02	2969	0.02	3614	0.01	3481
50	0	0.09	10000	0.09	10000	0.08	10000	0.09	10000
75	0	0.09	10000	0.09	10000	0.08	10000	0.08	10000
100	0	0.07	10000	0.06	10000	0.06	10000	0.06	10000
150	0	0.07	10000	0.05	10000	0.05	10000	0.05	10000
200	0	0.07	10000	0.06	10000	0.05	10000	0.05	10000
20	10	0.01	288	0.01	383	0.01	238	0.01	317
50	20	0.01	9613	0.00	6855	0.00	7772	0.01	8326
75	35	0.01	10000	0.01	10000	0.01	10000	0.01	8326
100	50	0.01	10000	0.01	10000	0.01	10000	0.01	10000
150	75	0.01	10000	0.01	10000	0.01	10000	0.01	10000
200	100	0.01	10000	0.01	10000	0.01	10000	0.01	10000

- UC not well-solved with high accuracy

Alternative I: The Conic Program Reformulation

- $\lambda f(x/\lambda)$ is SOCP-representable if f is[8]

[8] Ben-Tal, Nemirovski "Lectures on Modern Convex Optimization" SIAM, 2001

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

Alternative I: The Conic Program Reformulation

- $\lambda f(x/\lambda)$ is SOCP-representable if f is[8]
- Formulation using one (rotated) Second-Order Cone constraint:

$$\min \left\{ \, t + bp + cu \, : \, \sqrt{ap^2 + (t-u)^2/4} \leq (t+u)/2 \ldots u \in \{0,1\} \, \right\}$$

$\Longrightarrow$ a Mixed-Integer SOCP: Cplex can solve it . . .

8 Ben-Tal, Nemirovski "Lectures on Modern Convex Optimization" SIAM, 2001

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

Alternative I: The Conic Program Reformulation

- $\lambda f(x/\lambda)$ is SOCP-representable if f is[8]
- Formulation using one (rotated) Second-Order Cone constraint:

$$\min \left\{ t + bp + cu : \sqrt{ap^2 + (t-u)^2/4} \leq (t+u)/2 \ldots u \in \{0,1\} \right\}$$

$\Longrightarrow$ a Mixed-Integer SOCP: Cplex can solve it ...

		P/C					CP				
p	h	gap	nodes	LPs	time	t/LP	gap	nodes	CPs	time	t/CP
10	0		4.3e2	7.8e2	14	0.018		5.8e2	1.0e3	20	0.021
20	0		5.0e4	5.8e4	6805	0.094		6.6e4	7.5e4	13392	0.145
50	0	0.08	1.7e5	2.1e5	86400	0.421	0.08	9.1e4	1.1e5	86400	0.781
20	10		1.1e4	1.3e4	161	0.014		1.4e4	1.8e4	626	1.937
50	20		5.5e5	6.6e5	29874	0.037	0.00	5.0e5	6.1e5	86400	0.460
75	35	0.01	8.5e5	1.0e6	73076	0.073	0.01	1.8e5	2.2e5	86400	0.314

... but it's not a great idea (24h = very unrealistic, gap = 0.01%)[9]

[8] Ben-Tal, Nemirovski "Lectures on Modern Convex Optimization" SIAM, 2001

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

Alternative II: Projected Perspective Reformulation

- Only works under strong assumptions: p is a single variable, $\bar{p}_{min} \geq 0$, f is quadratic, there are no explicit constraints linking different u

[10] F., Gentile, Grande, Pacifici "Projected Perspective Reformulations with Applications in Design Problems" *Op. Res.*, 2011

Alternative II: Projected Perspective Reformulation

- Only works under strong assumptions: p is a single variable, $\bar{p}_{min} \geq 0$, f is quadratic, there are no explicit constraints linking different u
- Basic idea[10]: recast PR as $\min \{ z(p) \ : \ p \in [0, \bar{p}_{max}] \}$ where
  $z(p) = bp + \min_u \{ ap^2/u + cu \ : \ u\bar{p}_{min} \leq p \leq u\bar{p}_{max} \ , \ p \in [0,1] \}$
  (partial minimization of a convex function $\Longrightarrow$ convex)

[10] F., Gentile, Grande, Pacifici "Projected Perspective Reformulations with Applications in Design Problems" *Op. Res.*, 2011

Alternative II: Projected Perspective Reformulation

- Only works under strong assumptions: p is a single variable, $\bar{p}_{min} \geq 0$, f is quadratic, there are no explicit constraints linking different u
- Basic idea[10]: recast PR as $\min \{ \, z(p) \; : \; p \in [0, \bar{p}_{max}] \, \}$ where
  $z(p) = bp + \min_u \{ \, ap^2/u + cu \; : \; u\bar{p}_{min} \leq p \leq u\bar{p}_{max} \; , \; p \in [0, 1] \, \}$
  (partial minimization of a convex function $\Longrightarrow$ convex)
- Algebraic characterization of $z(p)$ out of optimal solution $u^*(p)$
- In turn, $u^*(p)$ out of unconstrained minimizer $\tilde{u}(p)$, i.e., solution to

$$\frac{\partial h(p,u)}{\partial u} = c - a\frac{p^2}{u^2} = 0 \qquad \text{(if any)}$$

  - $\tilde{u}(p)$ feasible $\Longrightarrow u^*(p) = \tilde{u}(p)$
  - otherwise, $u^*(p) \equiv$ projection of $\tilde{u}(p)$ over $[0, 1]$
    (easy because the feasible region is very simple)
- Basically, a few algebraic computations depending on a, c, ...

[10] F., Gentile, Grande, Pacifici "Projected Perspective Reformulations with Applications in Design Problems" *Op. Res.*, 2011

The form of $z(p)$

1) $c \leq 0 \Rightarrow \tilde{u}(p)$ undefined $\Rightarrow u^*(p) = 1 \Rightarrow$ $\boxed{z(p) = ap^2 + bp + c}$

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

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

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

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

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

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

- $z(p)$ convex differentiable piecewise-quadratic with ≤ 2 pieces
- Separable Convex (piecewise) Quadratic problems with the same structure ($\Rightarrow$ specialized algorithms) and at most 2x the variables

Alternative III: Approximated P^2R ("Project&Lift")

- ... except UC fails the test: *u* are non-separable

Alternative III: Approximated P^{2}R ("Project&Lift")

- . . . except UC fails the test: u are non-separable
- Possible solution: project as if u were separable, re-introduce them

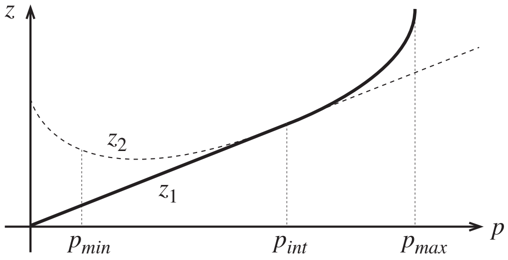
Alternative III: Approximated P^2R ("Project&Lift")

- ... except UC fails the test: u are non-separable
- Possible solution: project as if u were separable, re-introduce them
- Meanwhile, a few useful generalizations:
  - Extend to nonquadratic f with conditions on roots (p^k, e^p, Kleinrock)
  - Extend to $\bar{p}_{min} < 0$ (4-pieces z instead of 2-pieces one)

Alternative III: Approximated P^2R ("Project&Lift")

- ... except UC fails the test: u are non-separable
- Possible solution: project as if u were separable, re-introduce them
- Meanwhile, a few useful generalizations:
  - Extend to nonquadratic f with conditions on roots (p^k, e^p, Kleinrock)
  - Extend to $\bar{p}_{min} < 0$ (4-pieces z instead of 2-pieces one)
- The general form ($\bar{p}_{min} \geq 0$): for $p_{int} \in \{ \bar{p}_{min} , 1/g^+ , \bar{p}_{max} \}$

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



Approximated Projected Perspective Reformulation

- Theorem[11]: can be written as a NLP with the u "making" z_1:

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

[11] F., Furini, Gentile "Approximated Perspective Reformulations: a Project&Lift Approach", submitted, 2012

Approximated Projected Perspective Reformulation

- Theorem[11]: can be written as a NLP with the u "making" z_1:

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

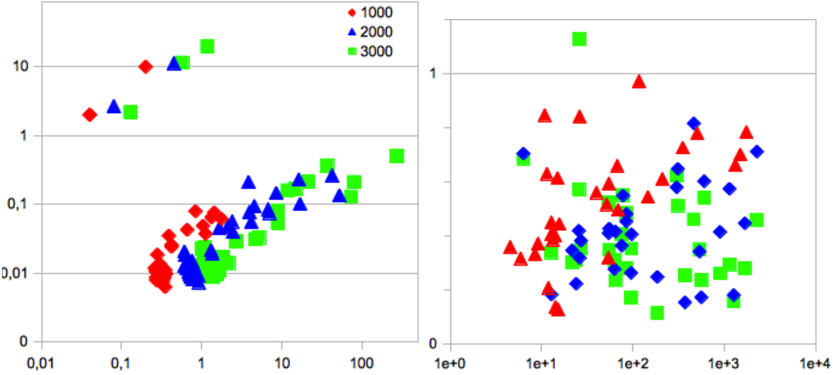
- With the integrality constraints, a reformulation of the block

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

  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 (!?!)

[11] F., Furini, Gentile "Approximated Perspective Reformulations: a Project&Lift Approach", submitted, 2012

Computational Results for Other Problems



Nonlinear Network Design — Mean-Variance Portfolio Problems

- (AP^2R time) / (P/C time) plotted against AP^2R time
- Works best if there are actually no linking constraints (not approximated)

Computational Results for UC

...which unfortunately is far from being the case for UC

		AP^2R				P/C			
p	h	nodes	time	pgap	gap	nodes	time	pgap	gap
10	0	1229	284	0.00	0.01	365	17	0.00	0.01
20	0	4635	9999	0.01	0.17	15607	4851	0.00	0.02
50	0	1078	9999	0.02	0.23	14286	9986	0.00	0.13
20	10	16477	1078	0.00	0.01	8107	240	0.00	0.01
50	20	3780	9999	0.00	0.07	66945	6649	0.00	0.02
75	35	1727	9999	0.03	0.08	57456	9999	0.00	0.02

- default 0.01% accuracy, time limit 10000 seconds
- Approximated PR bound too much weaker of true PR bound
- Silver lining: fewer relaxations solved (no Perspective Cuts)

Hybrid Approach I: Lagrangian, then MILP

- Lagrangian approach provides very good lower bounds, quickly
  but heuristic not so effective and efficient (many (ED), costly)

[12] F., Gentile, Lacalandra "Sequential Lagrangian-MILP Approaches for Unit Commitment Problems", *International Journal of Electrical Power and Energy Systems*, 2011

Hybrid Approach I: Lagrangian, then MILP

- Lagrangian approach provides very good lower bounds, quickly
  but heuristic not so effective and efficient (many (ED), costly)
- MILP approach provides very good feasible solutions, quickly
  but lower bounds loose, no termination for tight tolerance

[12] F., Gentile, Lacalandra "Sequential Lagrangian-MILP Approaches for Unit Commitment Problems", *International Journal of Electrical Power and Energy Systems*, 2011

Hybrid Approach I: Lagrangian, then MILP

- Lagrangian approach provides very good lower bounds, quickly
  but heuristic not so effective and efficient (many (ED), costly)
- MILP approach provides very good feasible solutions, quickly
  but lower bounds loose, no termination for tight tolerance

$$\Downarrow$$

Why not using both?

[12] F., Gentile, Lacalandra "Sequential Lagrangian-MILP Approaches for Unit Commitment Problems", *International Journal of Electrical Power and Energy Systems*, 2011

Hybrid Approach I: Lagrangian, then MILP

- Lagrangian approach provides very good lower bounds, quickly but heuristic not so effective and efficient (many (ED), costly)
- MILP approach provides very good feasible solutions, quickly but lower bounds loose, no termination for tight tolerance

$$\Downarrow$$

Why not using both?

- Not obvious how, but has to be a way (Lagrangian = large scale reformulation $\Rightarrow$ column/row generation ...)
- Dumbest possible way: compute LR bound (no heuristic = quick), use it to stop search as soon as good enough feasible solution found (admittedly[12] very coarse)

[12] F., Gentile, Lacalandra "Sequential Lagrangian-MILP Approaches for Unit Commitment Problems", *International Journal of Electrical Power and Energy Systems*, 2011

Results – 0.5%

		P/CD_4				P/CD_∞			
		NoLB		LB		NoLB		LB	
p	h	gap	time	gap	time	gap	time	gap	time
10	0	0.28	0.80	0.34	0.97	0.30	0.86	0.37	1.06
20	0	0.33	3.00	0.32	3.60	0.36	2.51	0.36	3.16
50	0	0.18	13.08	0.19	27.46	0.19	14.17	0.20	16.39
75	0	0.22	22.58	0.25	28.82	0.19	36.62	0.22	28.05
100	0	0.20	36.51	0.15	41.44	0.17	34.31	0.16	60.16
150	0	0.12	169.68	0.10	148.88	0.11	104.68	0.11	136.18
200	0	0.14	235.60	0.08	323.36	0.10	183.01	0.08	258.57
20	10	0.15	2.51	0.17	4.21	0.30	4.18	0.24	6.34
50	20	0.13	10.93	0.10	26.96	0.10	19.06	0.10	12.51
75	35	0.03	64.80	0.06	59.47	0.05	70.55	0.10	75.23
100	50	0.04	60.78	0.04	44.95	0.05	47.62	0.05	66.61
150	75	0.05	216.33	0.02	244.05	0.05	194.10	0.04	228.32
200	100	0.03	342.69	0.03	253.59	0.02	155.36	0.02	217.56

- Sizable relative (although small absolute) increase for small instances
- No clear positive effect

Results – 0.1%

		P/CD_4				P/CD_∞			
		NoLB		LB		NoLB		LB	
p	h	gap	time	gap	time	gap	time	gap	time
10	0	0.10	12.45	0.10	12.70	0.10	9.77	0.10	9.97
20	0	0.10	1295.28	0.10	2201.87	0.10	1169.94	0.10	1157.22
50	0	0.09	8279.78	0.11	4084.79	0.10	10000.00	0.11	4014.01
75	0	0.07	10000.00	0.09	3974.94	0.07	10000.00	0.09	2286.03
100	0	0.07	10000.00	0.09	289.01	0.06	10000.00	0.09	94.56
150	0	0.05	10000.00	0.06	193.38	0.05	10000.00	0.08	207.86
200	0	0.05	10000.00	0.07	337.33	0.06	10000.00	0.07	315.88
20	10	0.07	31.38	0.09	14.31	0.07	41.08	0.08	30.01
50	20	0.02	41.86	0.05	27.22	0.02	47.62	0.04	12.92
75	35	0.03	64.45	0.06	57.95	0.04	81.77	0.06	71.03
100	50	0.03	40.61	0.04	41.42	0.04	60.20	0.05	62.85
150	75	0.02	232.99	0.02	235.04	0.04	191.52	0.04	203.18
200	100	0.03	240.38	0.03	231.35	0.02	198.25	0.02	206.66

- Huge positive impact on large thermals, some effect on small hydro
- Gap worsens somewhat (which is expected)

Results – 0.05%

		P/CD_4				P/CD_∞			
		NoLB		LB		NoLB		LB	
p	h	gap	time	gap	time	gap	time	gap	time
10	0	0.06	15.42	0.06	15.72	0.06	11.63	0.06	11.85
20	0	0.06	2473.11	0.06	2440.86	0.06	2470.49	0.06	2499.97
50	0	0.09	10000.00	0.09	8113.35	0.09	10000.00	0.10	8489.08
75	0	0.09	10000.00	0.09	10002.22	0.08	8256.79	0.08	8259.00
100	0	0.07	10000.00	0.07	8018.89	0.06	10000.00	0.06	6538.84
150	0	0.05	10000.00	0.06	5151.71	0.05	10000.00	0.06	6151.20
200	0	0.05	10000.00	0.05	6255.99	0.06	10000.00	0.06	6271.77
20	10	0.06	73.26	0.06	73.00	0.06	71.19	0.06	68.40
50	20	0.01	623.95	0.02	34.44	0.01	269.34	0.03	44.53
75	35	0.02	177.50	0.03	59.37	0.02	124.85	0.03	100.47
100	50	0.02	438.39	0.04	39.45	0.02	665.37	0.05	60.00
150	75	0.02	1669.30	0.02	224.67	0.01	1144.10	0.04	201.31
200	100	0.02	1082.41	0.03	238.81	0.01	451.98	0.02	202.94

- Diminishing but still positive on large thermals, especially P/CD_∞
- Huge positive impact on all but the smallest hydro

Results – 0.01%

		P/CD_4				P/CD_∞			
		NoLB		LB		NoLB		LB	
p	h	gap	time	gap	time	gap	time	gap	time
10	0	0.02	16.66	0.02	16.84	0.02	12.49	0.02	12.80
20	0	0.02	3547.24	0.02	3699.26	0.02	3914.50	0.02	3946.51
50	0	0.09	10000.00	0.09	10001.25	0.09	10000.00	0.09	10001.25
75	0	0.09	10000.00	0.09	10002.22	0.08	10000.00	0.08	10002.22
100	0	0.07	10000.00	0.07	10003.68	0.06	10000.00	0.06	10003.68
150	0	0.05	10000.00	0.05	10006.14	0.05	10000.00	0.05	10006.14
200	0	0.05	10000.00	0.05	8248.37	0.06	10000.00	0.06	10008.52
20	10	0.02	268.49	0.02	263.40	0.02	248.75	0.02	255.95
50	20	0.00	7285.00	0.01	841.26	0.01	6495.96	0.01	121.86
75	35	0.01	10000.00	0.01	5033.34	0.01	10000.00	0.01	5045.42
100	50	0.01	10000.00	0.01	1198.73	0.01	10000.00	0.01	5789.69
150	75	0.01	10000.00	0.01	3376.87	0.01	10000.00	0.01	1145.61
200	100	0.01	10000.00	0.01	1182.27	0.01	10000.00	0.01	463.46

- No longer any impact (thus, slightly negative) on thermals
- Still huge positive impact on all but the smallest hydro

Hybrid Approach II: Lagrangian at Each B&B Node

- Logical next step: Lagrangian at each B&B node

[13] F. "About Lagrangian Methods in Integer Optimization" *Annals of Operations Research*, 2005

Hybrid Approach II: Lagrangian at Each B&B Node

- Logical next step: Lagrangian at each B&B node
- Entirely LR-based B&B (B&P) clearly possible, primals available[13]
- Yet, one would lose the effective primal heuristics (and flexibility)

[13] F. "About Lagrangian Methods in Integer Optimization" *Annals of Operations Research*, 2005

Hybrid Approach II: Lagrangian at Each B&B Node

- Logical next step: Lagrangian at each B&B node
- Entirely LR-based B&B (B&P) clearly possible, primals available[13]
- Yet, one would lose the effective primal heuristics (and flexibility)
- Simple idea: use the LP/QP duals of (10) as λ
- Obvious possibility: improve it somewhat (bundle/subgradient . . .)
- Nontrivial implementation issues:
  - Ram a better bound down the throat of `Cplex` (`branchcallback`)
  - Insert `Cplex` fixings into the LR
  - Exploit `Cplex cuts`: retrieve/recognize them, preprocessing, . . .
- Easier said than done, maybe alternative tools (SCIP)?

[13] F. "About Lagrangian Methods in Integer Optimization" *Annals of Operations Research*, 2005

Preliminary Results (on a much faster machine)

		P/C		P/C + LR		AP^2R		AP^2R + LR	
gap	n	nodes	time	nodes	time	nodes	time	nodes	time
5e-3	10	0	0.37	0	0.38	0	1.36	0	1.31
	20	0	1.47	0	1.48	0	5.21	0	5.00
1e-3	10	321	12.43	266	9.57	534	130.37	415	92.02
	20	3183	308.33	2895	253.00	6544	5012.21	3789	3191.12
5e-4	10	543	17.67	621	16.39	674	144.19	445	81.50
	20	6348	603.51	6237	504.71	11188	6406.15	6829	5276.84
1e-4	10	664	18.96	530	16.05	723	145.41	580	101.05
	20	10381	859.99	10706	966.32	15046	6755.73	8403	6173.50

- Have results up to $n = 200$ ... but there is a bug for $n > 20$
- Compute LR bound at every node, never try to improve it
- +LR always helps for AP^2R (pity it is not competitive to start with)
- +LR almost always helps for P/C (but more on "middle" gaps)
- All in all promising, but still a lot of work to do

Conclusions

- Energy problems motivate good methodological research, good methodologies help solving real-world problems
- Lagrangian techniques powerful, still useful, somewhat cumbersome
- MI(N)LP-based approaches "surf the wave" of general-purpose solvers development, powerful, more flexible, have their own quirks
- Combining the two makes sense, nontrivial but possible
- The applications require it
- Many other possible ways, interesting venue of research; one possible approach: (Stabilized) Structured Dantzig-Wolfe[14]

[14] F., Gendron "A Stabilized Structured Dantzig-Wolfe Decomposition Method" *Math. Prog.*, to appear