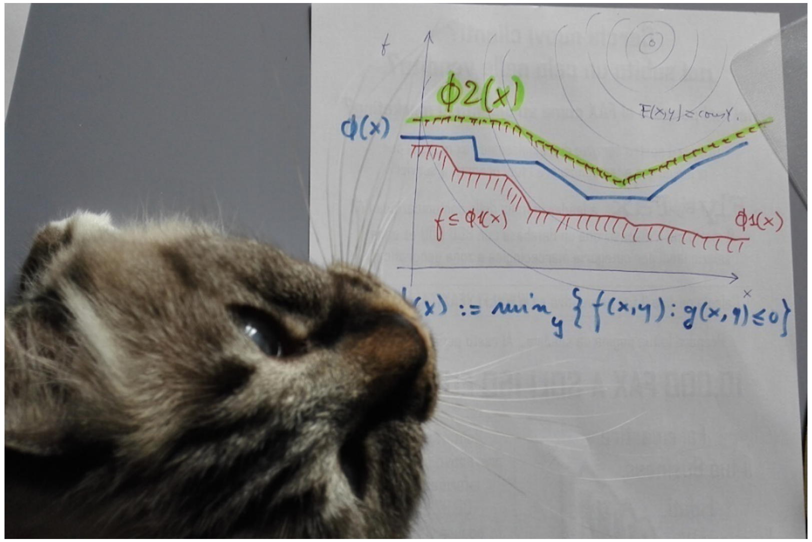
Intersection Cuts for Bilevel Optimization

Matteo Fischetti, University of Padova
Ivana Ljubic, ESSEC Paris

Michele Monaci, University of Padova
Markus Sinnl, University of Vienna



Bilevel Optimization

- The general **Bilevel Optimization Problem** (optimistic version) reads:

$$\min_{x \in \mathbb{R}^{n_1}, y \in \mathbb{R}^{n_2}} F(x,y)$$
$$G(x,y) \leq 0$$
$$y \in \arg \min_{y' \in \mathbb{R}^{n_2}} \{ f(x,y') : g(x,y') \leq 0 \}$$

  where *x* var.s only are controlled by the **leader**, while *y* var.s are computed by another player (the **follower**) solving a different problem.

- A very very hard problem even in a **convex setting with continuous var.s** only

- **Convergent** solution algorithms are problematic and typically require additional assumptions (binary/integer var.s or alike)

Example: 0-1 ILP

- A generic 0-1 ILP
  can be reformulated as
  the following **linear &**
  continuos bilevel problem

$$\min c^T x$$
$$Ax = b$$
$$x \in \{0,1\}^n$$

$$\min c^T x$$
$$Ax = b$$
$$x \in [0,1]^n$$
$$y = 0$$

$$y \in \arg\min_{y'}\{-\sum_{j=1}^{n} y'_j : \; y'_j \le x_j, \; y'_j \le 1 - x_j \quad \forall j = 1, \ldots, n\}$$

Note that y is fixed to 0 but it cannot be removed from the model!

Reformulation

$$\min_{x\in\mathbb{R}^{n_1}, y\in\mathbb{R}^{n_2}} F(x,y)$$
$$G(x,y) \le 0$$
$$y \in \arg\min_{y'\in\mathbb{R}^{n_2}} \{f(x,y') : g(x,y') \le 0\}$$

- By defining the **value function**

$$\Phi(x) = \min_{y\in\mathbb{R}^{n_2}} \{f(x,y) : g(x,y) \le 0\},$$

  the problem can be restated as

$$\begin{aligned} \min\ & F(x,y) \\ & G(x,y) \le 0 \\ & g(x,y) \le 0 \\ & f(x,y) \le \Phi(x) \\ & (x,y) \in \mathbb{R}^n. \end{aligned}$$

- Dropping the nonconvex condition $f(x,y) \le \Phi(x)$ one gets the so-called **High Point Relaxation** (HPR)

Mixed-Integer Bilevel Linear Problems

- We will focus the **Mixed-Integer Bilevel Linear** case (MIBLP)

$$\begin{aligned} \min F(x,y) & \\ G(x,y) &\le 0 \\ g(x,y) &\le 0 \\ (x,y) &\in \mathbb{R}^n \\ f(x,y) &\le \Phi(x) \\ x_j \text{ integer}, & \quad \forall j \in J_1 \\ y_j \text{ integer}, & \quad \forall j \in J_2, \end{aligned}$$

  where *F, G, f* and *g* are **affine functions**

- Note that $f(x,y) \le \Phi(x)$ **remains highly nonconvex** even when all y var.s are continuous

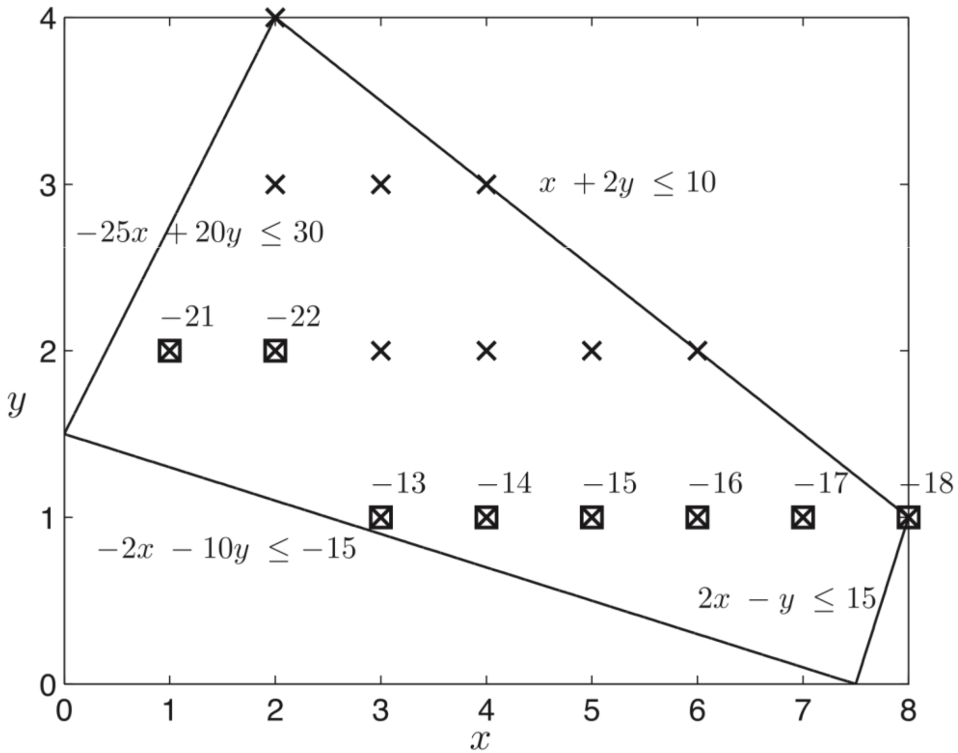
- **HPR is a familiar MILP** → we can apply our whole MILP **bag of tricks**!

Example

- A notorious example from

J. Moore and J. Bard. The mixed integer linear bilevel programming problem. *Operations Research*, 38(5):911–921, 1990.

$$\min_{x \in \mathbb{Z}} \; -x - 10y$$
$$y \in \arg\min_{y' \in \mathbb{Z}} \{ \; y' :$$
$$-25x + 20y' \leq 30$$
$$x + 2y' \leq 10$$
$$2x - y' \leq 15$$
$$2x + 10y' \geq 15 \; \}$$



where *f(x,y) = y*

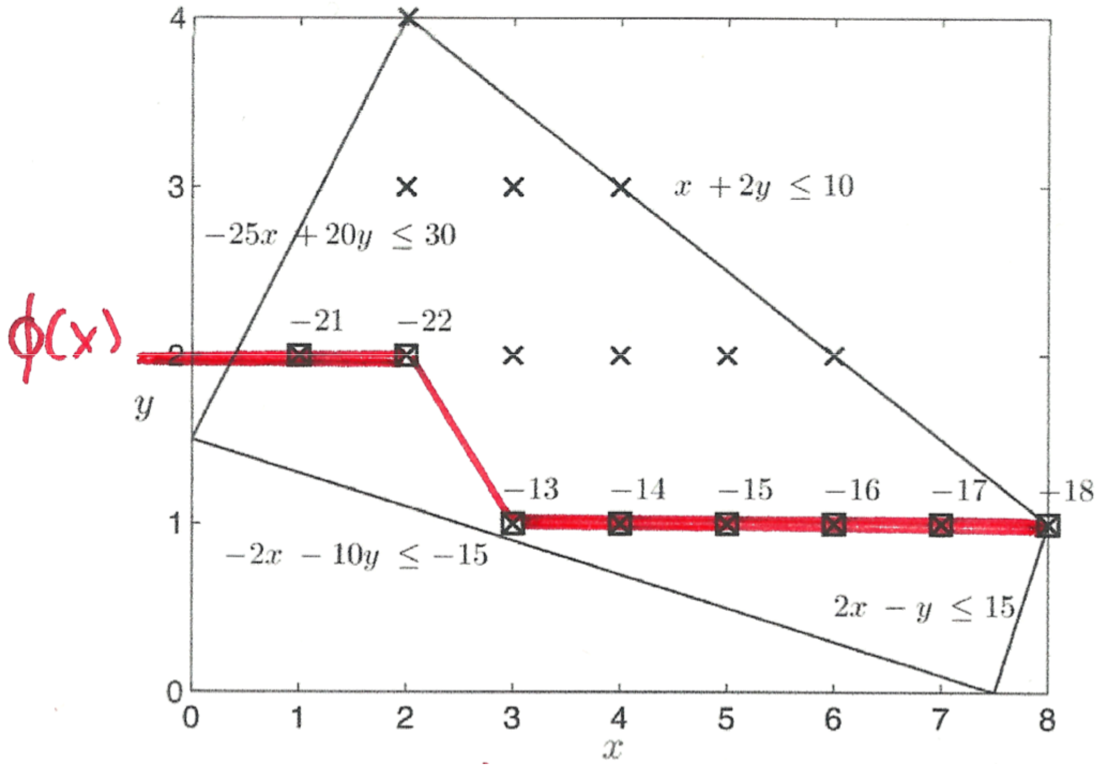
x points of HPR relax.

____ LP relax. of HPR

Example (cont.d)

Value-function reformulation

$$\begin{aligned} \min \; & -x - 10y \\ & -25x + 20y \leq 30 \\ & x + 2y \leq 10 \\ & 2x - y \leq 15 \\ & -2x - 10y \leq -15 \\ & x, y \in \mathbb{Z} \\ & y \leq \Phi(x) \end{aligned}$$



MILP-based solver

- Suppose to apply a **Branch-and-Cut** MILP solver to HPR

- Forget for a moment about internal heuristics, and assume the LP relaxation at each node is solved by the simplex algorithm

- **What is needed to guarantee correctness of the MILP solver?**

- At each node, let ***(x*,y*)*** be the current **LP optimal vertex**:

  ***if (x*,y*)* is fractional → branch as usual**

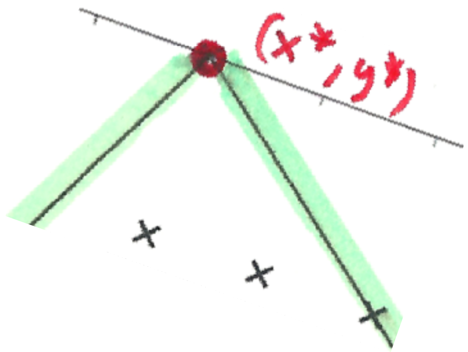
  ***if (x*,y*)* is integer and $f(x^*, y^*) \leq \Phi(x^*)$ → update the incumbent as usual**

The difficult case

- But, what can we do in third possible case, namely

  ***(x*,y*)* is integer but not bilevel-feasible, i.e.**

$$f(x^*, y^*) > \Phi(x^*)$$



  Possible answers from the literature

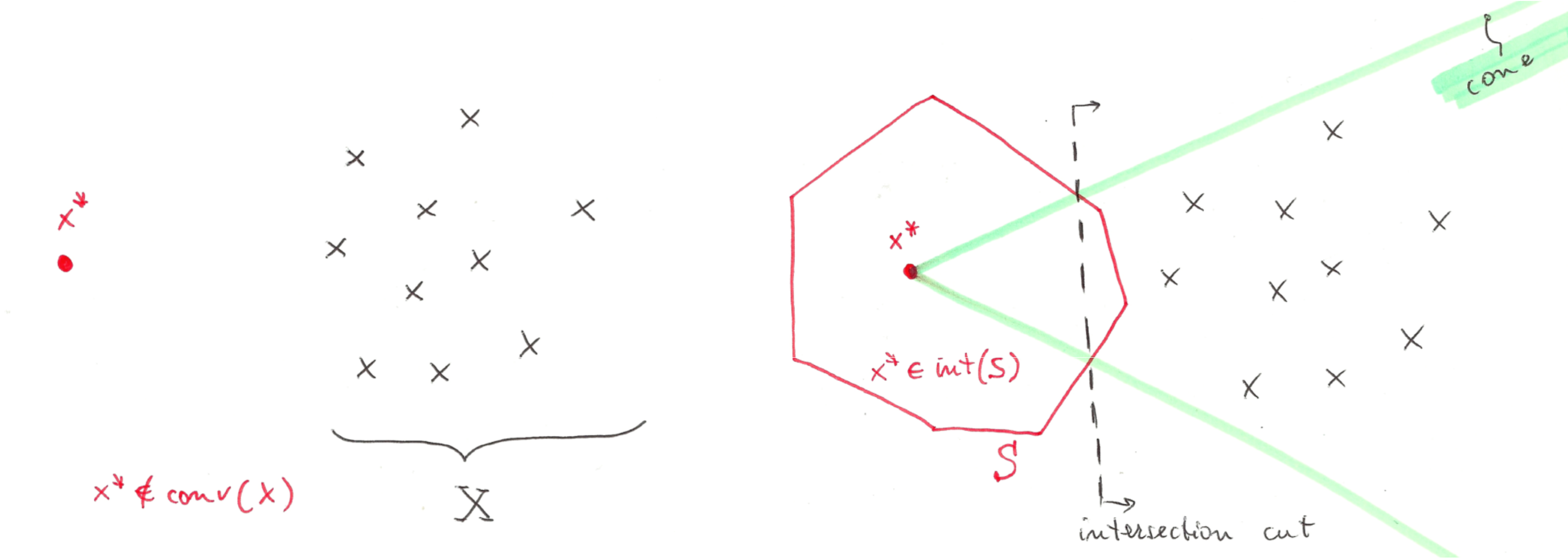
- If *(x,y)* is restricted to be **binary**, add **a no-good cut** requiring to flip at least one variable w.r.t. *(x*,y*)* or w.r.t. x*

- If *(x,y)* is restricted to be **integer** and all MILP coeff.s are integer, add a cut requiring a slack of 1 for the sum of all the inequalities that are tight at *(x*,y*)*

- Weak conditions as they do not addresses the **reason of infeasibility** by trying to enforce $f(x^*, y^*) \leq \Phi(x^*)$ somehow

Intersection Cuts (IC's)

- We propose the use of **intersection cuts** (Balas, 1971) instead
- IC is powerful tool to separate a point **x*** from a set **X** by a liner cut

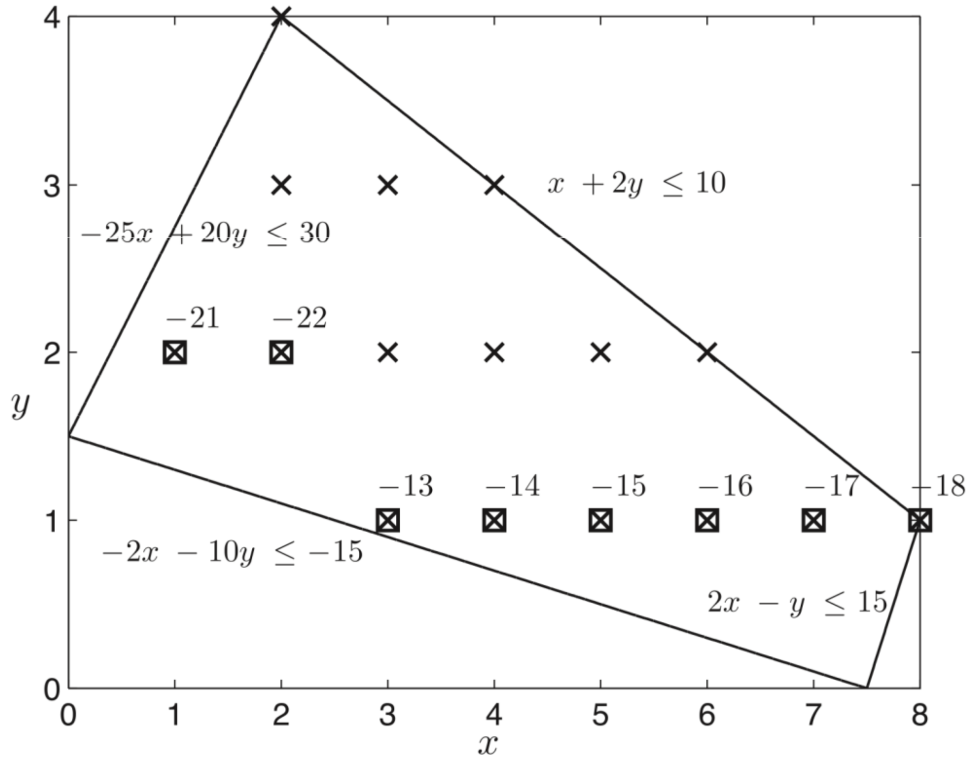


- All you need is […love, but also]
  - a **cone** pointed at **x*** containing all **x ε X**
  - a **convex set S** with x* (but no **x ε X**) in its **<u>interior</u>**
- If x* **vertex** of an LP relaxation, a possible cone comes for **LP basis**

IC's for bilevel problems

- Our idea is first illustrated on the Moore&Bard example

$$\min_{x \in \mathbb{Z}} \; -x - 10y$$
$$y \in \arg\min_{y' \in \mathbb{Z}} \{ \; y' :$$
$$-25x + 20y' \leq 30$$
$$x + 2y' \leq 10$$
$$2x - y' \leq 15$$
$$2x + 10y' \geq 15 \; \}$$

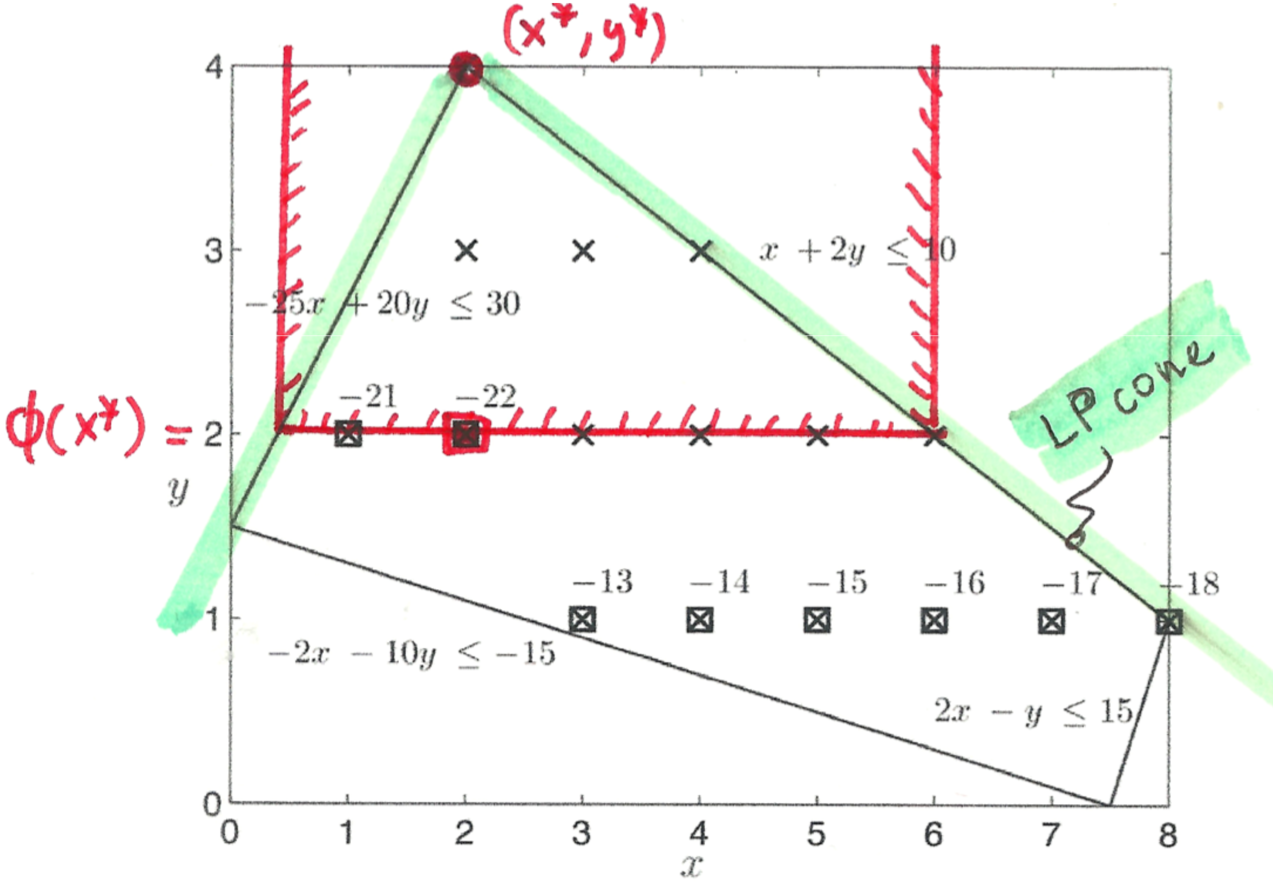


x points of HPR relax.

_____ LP relax. of HPR

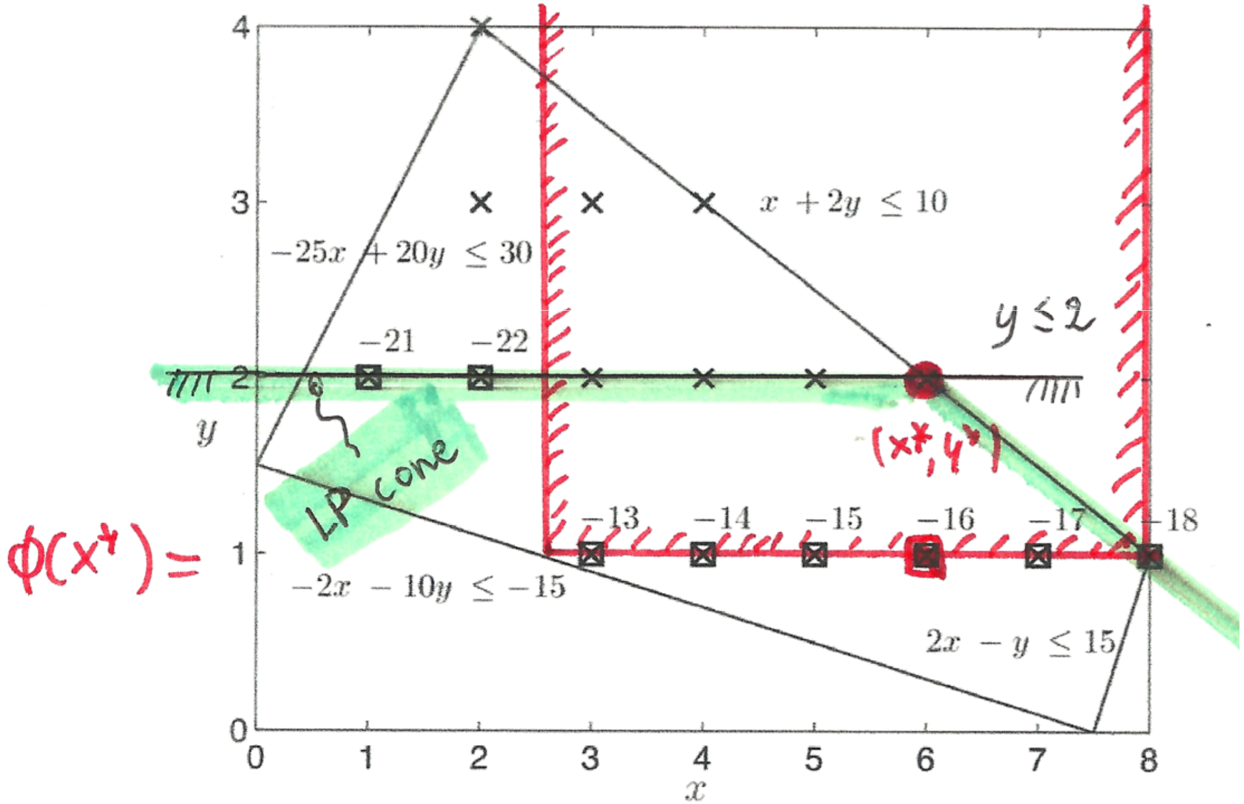
Bilevel-free sets

- Take the LP vertex $(x^*,y^*) = (2,4)$ → $f(x^*,y^*) = y^* = 4 > Phi(x^*) = 2$



Intersection cut

- We can therefore generate the intersection cut *y <= 2* and repeat



A basic bilevel-free set

Lemma 1. *For any feasible solution $\hat{y}$ of the follower, the set*

$$S(\hat{y}) = \{(x,y) \in \mathbb{R}^n : f(x,y) \geq f(x,\hat{y}),\ g(x,\hat{y}) \leq 0\} \tag{10}$$

does not contain any bilevel-feasible point in its interior.

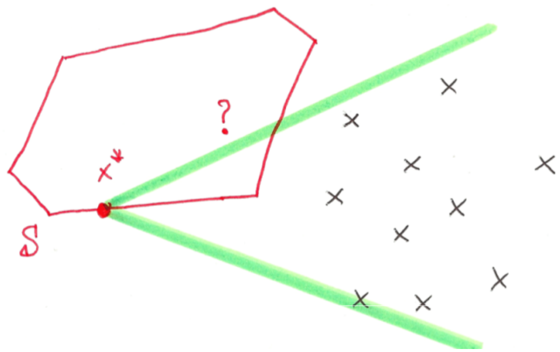
- **Note**: $S(\hat{y})$ is a convex set (actually, a **polyhedron**) when *f* and *g* are affine functions, i.e., in the MIBLP case

- **Separation algorithm**: given an optimal vertex *(x*,y*)* of the LP relaxation of HPR
  - Solve the follower for *x=x** and get an optimal sol., say $\hat{y}$
  - **if** *(x*,y*)* strictly inside $S(\hat{y})$ **then**

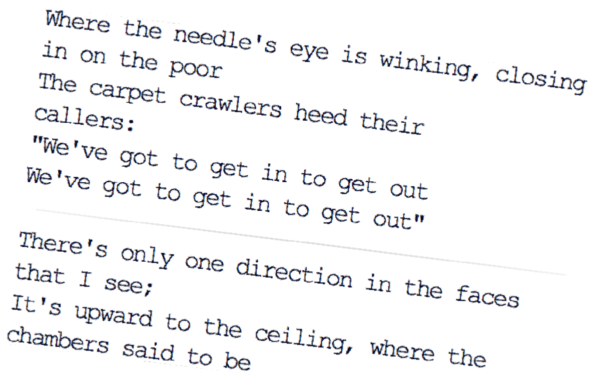
    generate a violated IC using the LP-cone pointed at *(x*,y*)* together with the bilevel-free set $S(\hat{y})$

We’ve got to get in to get out!

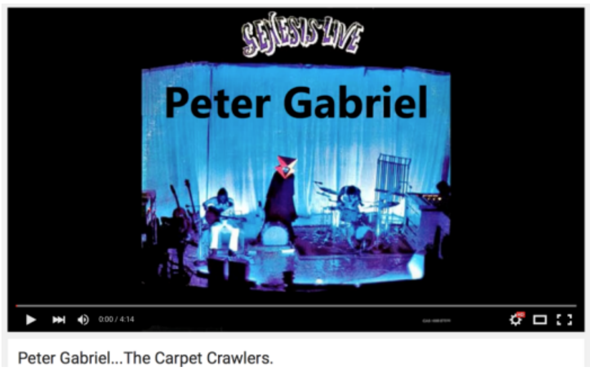
- However, the above does not lead to a convergent MILP algorithm as a **bilevel-infeasible** integer vertex ***(x*,y*)*** can be on the **frontier** of the bilevel-free set *S* so we cannot be sure to cut it by using our IC’s



- Indeed, this is a well-know issue with IC’s already pointed out in the 70th by [GCRBH74]


Where the needle's eye is winking, closing
in on the poor
The carpet crawlers heed their
callers:
"We've got to get in to get out
We've got to get in to get out"

There's only one direction in the faces
that I see;
It's upward to the ceiling, where the
chambers said to be



[GCRBH74] P. Gabriel, P. Collins, M. Rutherford, T. Banks, and S. Hackett, “The Carpet Crawlers”, in *The Lamb Lies Down on Broadway* (Genesis ed.s), 1974

Getting well inside bilevel-free sets

- Assuming *g(x,y)* is integer for all integer HPR solutions, we proved

Theorem 1. *Assume that* $g(x,y)$ *is integer for all HPR solutions* (x,y)*. Then, for any feasible solution* $\hat{y}$ *of the follower, the extended set*

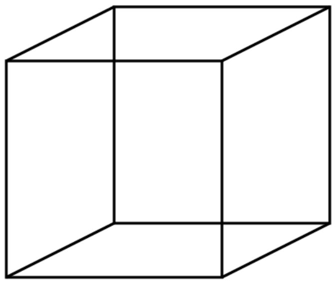
$$S^{+}(\hat{y}) = \{(x,y) \in \mathbb{R}^n : f(x,y) \geq f(x,\hat{y}),\ g(x,\hat{y}) \leq 1\} \qquad (11)$$

does not contain any bilevel-feasible point in its interior, where $\mathbf{1}$ *denotes a vector of all one's.*

- The corresponding intersection cut is always violated and leads to a **convergent MILP-based solver** when, e.g., var.s *x,y* are required to be integer and follower constraint coeff.s are all integer

Informed No-Good (ING) cuts

- IC's using tableaux information (LP cone) become shallow and numerically unstable in the long run **#ThinkOfGomoryCuts**
- Possibly deactivated after root node for fractional sol.s **#TooManyCuts**
- More stable performance if combined with the following new class of **Informed No-Good (ING) cuts** when mathematically correct (e.g. for binary problems)
- No LP cone required, just use the cone associated with **tight** lower/upper var. bounds



- ING cuts **dominate** standard no-good cuts when using an "**informed**" bilevel-free set → ING cuts can play a role in other contexts such as CP where no-goods rule

Preliminary computational results

- First-shot comparison with **MibS**, a state of the art open-source solver developed and maintained by T. Ralphs & S. DeNegre

- Results not directly comparable as MibS is based on SYMPHONY while our B&C is built on top of IBM ILOG CPLEX 12.6.2

- To me more fair: IC's only → no ING cuts, no CPLEX cuts, no heur.s, 1 thread **(good for #JoCM)**

name	Mibs UB	Mibs LB	Mibs %gap	Mibs t.[s]	B&C with IC's UB	B&C with IC's LB	B&C with IC's %gap	B&C with IC's t.[s]
fast0507-0.1	-	173	100.00	TL	12553	173	98.62	TL
fast0507-0.5	-	173	100.00	TL	61503	174	99.72	TL
fast0507-0.9	-	173	100.00	TL	109916	109916	0.00	7
lseu-0.1	1120	1120	0.00	4	1120	1120	0.00	2
lseu-0.5	2400	1205	49.79	TL	2263	1235	45.43	TL
lseu-0.9	5838	1171	79.94	TL	5838	1275	78.75	TL
p0033-0.1	3089	3089	0.00	0	3089	3089	0.00	0
p0033-0.5	3095	3095	0.00	0	3095	3095	0.00	0
p0033-0.9	4679	4679	0.00	90	4679	4679	0.00	3
p0201-0.1	12615	7859	37.70	TL	12465	7931	36.37	TL
p0201-0.5	14220	7832	44.92	TL	13910	7925	43.03	TL
p0201-0.9	15025	7809	48.03	TL	15025	7925	47.25	TL
p0282-0.1	261188	258435	1.05	TL	260781	260067	0.27	TL
p0282-0.5	276338	258432	6.48	TL	272659	259331	4.89	TL
p0282-0.9	724572	258427	64.33	TL	636846	284519	55.32	TL
p0548-0.1	-	317	100.00	TL	10982	8691	20.86	TL
p0548-0.5	-	317	100.00	TL	22450	8620	61.60	TL
p0548-0.9	-	317	100.00	TL	48959	8694	82.24	TL
p2756-0.1	-	2691	100.00	TL	12765	2734	78.58	TL
p2756-0.5	-	2691	100.00	TL	23976	2723	88.64	TL
p2756-0.9	-	2691	100.00	TL	35867	2733	92.38	TL
seymour-0.1	-	407	100.00	TL	480	407	15.21	TL
seymour-0.5	-	407	100.00	TL	823	408	50.43	TL
seymour-0.9	-	407	100.00	TL	1251	1251	0.00	2
stein27-0.1	18	18	0.00	0	18	18	0.00	1
stein27-0.5	19	19	0.00	7	19	19	0.00	3
stein27-0.9	24	20	16.67	TL	24	24	0.00	0
stein45-0.1	30	30	0.00	103	30	30	0.00	32
stein45-0.5	33	31	6.06	TL	32	32	0.00	205
stein45-0.9	40	31	22.50	TL	40	40	0.00	0

- B&C: just few hundred lines (the callback for IC separation) on top of Cplex
- B&C produces better lower and upper bounds (and solves more instances)

Thanks for your attention

Slides available http://www.dei.unipd.it/~fisch/papers/slides/

"The Carpet Crawlers"

There is lambswool under my naked feet.
The wool is soft and warm,
-gives off some kind of heat.
A salamander scurries into flame to be destroyed.
Imaginary creatures are trapped in birth on celluloid.
The fleas cling to the golden fleece,
Hoping they'll find peace.
Each thought and gesture are caught in celluloid.
There's no hiding in my memory.
There's no room to void.

The crawlers cover the floor in the red ochre corridor.
For my second sight of people, they've more lifeblood than before.
They're moving. They're moving in time to a heavy wooden door,
Where the needle's eye is winking, closing in on the poor.
The carpet crawlers heed their callers:
"We've got to get in to get out
We've got to get in to get out."

There's only one direction in the faces that I see;
It's upward to the ceiling, where the chambers said to be.
Like the forest fight for sunlight, that takes root in every tree.
They are pulled up by the magnet, believing that they're free.
The carpet crawlers heed their callers:
"We've got to get in to get out
We've got to get in to get out."

Mild mannered supermen are held in kryptonite,
And the wise and foolish virgins giggle with their bodies glowing bright.
Through a door a harvest feast is lit by candlelight;
It's the bottom of a staircase that spirals out of sight.
The carpet crawlers heed their callers:
"We've got to get in to get out
We've got to get in to get out."

The porcelain mannikin with shattered skin fears attack.
The eager pack lift up their pitchers- the carry all they lack.
The liquid has congealed, which has seeped out through the crack,
And the tickler takes his stickleback.
The carpet crawlers heed their callers:
"We've got to get in to get out
We've got to get in to get out."