Valid Inequalities for Mixed Integer Bilevel Linear Optimization Problems

Ted Ralphs[1]
Joint work with Sahar Tahernejad[1], Suresh Bolusani[1]

[1]COR@L Lab, Department of Industrial and Systems Engineering, Lehigh University

IWOBIP, Lille, France, June 20, 2018

Attributions

Many current and students contributed to development of this work.

Current and Former Ph.D Students

- Suresh Bolusani
- Scott DeNegre
- Menal Gúzelsoy
- Anahita Hassanzadeh
- Ashutosh Mahajan
- Sahar Tahernajad

Thanks!

Mixed Integer Bilevel Linear Optimization Problems

- *First-level variables*: $x \in X$ where $X = \mathbb{Z}_+^{r_1} \times \mathbb{R}_+^{n_1 - r_1}$
- *Second-level variables*: $y \in Y$ where $Y = \mathbb{Z}_+^{r_2} \times \mathbb{R}_+^{n_2 - r_2}$

MIBLP

$$\min_{x,y} \left\{ cx + d^1 y \;\middle|\; x \in X, y \in \mathcal{P}_1(x), y \in \operatorname{argmin}\{ d^2 z \mid z \in \mathcal{P}_2(x) \cap Y \} \right\}, \tag{1}$$

where

$$\mathcal{P}_1(x) = \left\{ y \in \mathbb{R}_+^{n_2} \;\middle|\; A^1 x + G^1 y \geq b^1 \right\},$$

$$\mathcal{P}_2(x) = \left\{ y \in \mathbb{R}_+^{n_2} \;\middle|\; G^2 y \geq b^2 - A^2 x \right\},$$

Value Function Reformulation

- *First-level variables*: $x \in X$ where $X = \mathbb{Z}_+^{r_1} \times \mathbb{R}_+^{n_1 - r_1}$
- *Second-level variables*: $y \in Y$ where $Y = \mathbb{Z}_+^{r_2} \times \mathbb{R}_+^{n_2 - r_2}$

MIBLP

$$\min_{x,y} \left\{ cx + d^1 y \;\middle|\; x \in X, y \in \mathcal{P}_1(x) \cap \mathcal{P}_2(x) \cap Y, d^2 y \leq \phi(b^2 - A^2 x) \right\}, \tag{MIBLP}$$

where

$$\mathcal{P}_1(x) = \left\{ y \in \mathbb{R}_+^{n_2} \;\middle|\; A^1 x + G^1 y \geq b^1 \right\},$$

$$\mathcal{P}_2(x) = \left\{ y \in \mathbb{R}_+^{n_2} \;\middle|\; G^2 y \geq b^2 - A^2 x \right\},$$

$$\phi(\beta) = \min \left\{ d^2 y \;\middle|\; G^2 y \geq \beta, y \in Y \right\} \forall \beta \in \mathbb{R}^{m_2}.$$

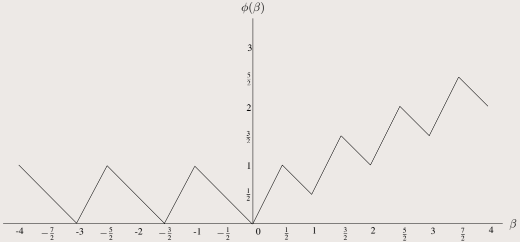
The Value Function

- The second-level *value function* is

MILP Value Function

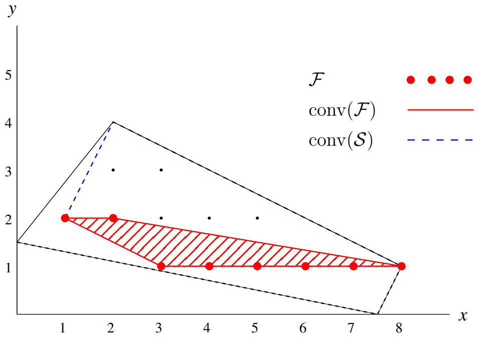
$$\phi(\beta) = \min_{x \in \mathcal{P}_2(\beta) \cap X} c^\top x \qquad \text{(VF)}$$

for $\beta \in \mathbb{R}^m$. We let $\phi(\beta) = \infty$ if $\beta \in \Omega = \{\beta \in \mathbb{R}^m \mid \mathcal{S}(\beta) = \emptyset\}$.



Polyhedral Reformulation

Convexification considers the following conceptual reformulation.



$$\begin{aligned} \min \quad & c^1 x + d^1 y \\ \text{s.t.} \quad & (x, y) \in \text{conv}(\mathcal{F}) \end{aligned}$$

- This reformulation leads to a similar branch-and-cut algorithm similar to that used for solving MILPs.
- To get bounds, we optimize over a relaxed feasible region.
- We iteratively approximate the true feasible region with linear inequalities.

Branch-and-Cut Algorithm

- The algorithm is based on the framework originally described by DeNegre and Ralphs [2009], but with many additional enhancements.
- The algorithm has been implemented in the `MibS` framework, which is open source and available from COIN-OR.
- Details are contained in a forthcoming paper by Tahernejad et al. [2016] (preprint available)

Components

- Bounding
  - **Lower bound** ⇒ An LP relaxation strengthened with **valid inequalities**
  - Upper bound ⇒ Feasible solutions
- Branching ⇒ Several schemes for branching
- Feasibility checking
- Search strategies
- Preprocessing methods
- Primal heuristics

Linking Variables

- The first-level variables with non-zero coefficients in the second-level problem are called the *linking variables* and are indexed by set L.
- **Assumption**: All linking variables are discrete $\Rightarrow$ The optimal value of MIBLP is attainable.

For the vectors $x^1, x^2 \in X$ with $x^1_L = x^2_L$, we have

$$\phi(b^2 - A^2 x^1) = \phi(b^2 - A^2 x^2)$$

- The best bilevel feasible solution (x,y) with $x_L = \gamma \in \mathbb{Z}^L$ can be obtained by solving just one MILP.

$$\min \left\{ cx + d^1 y \;\middle|\; x \in X, y \in \mathcal{P}_1(x) \cap \mathcal{P}_2(x) \cap Y, d^2 y \leq \phi(b^2 - A^2 x), x_L = \gamma \right\} \quad \text{(UB)}$$

Lower Bound

Bilevel Feasible Region

$$\mathcal{F} = \left\{ (x, y) \in \mathbb{R}_{+}^{n_1 \times n_2} \;\middle|\; x \in X, y \in \mathcal{P}_1(x) \cap \mathcal{P}_2(x) \cap Y, d^2 y \leq \phi(b^2 - A^2 x) \right\}$$

Two possible relaxations

1. Removing the *optimality constraint of the second-level problem*

$$\mathcal{S} = \left\{ (x, y) \in \mathbb{R}_{+}^{n_1 \times n_2} \;\middle|\; x \in X, y \in \mathcal{P}_1(x) \cap \mathcal{P}_2(x) \cap Y \right\}$$

2. Removing the *optimality constraint of the second-level problem* and the *integrality constraints*

$$\mathcal{P} = \left\{ (x, y) \in \mathbb{R}_{+}^{n_1 \times n_2} \;\middle|\; y \in \mathcal{P}_1(x) \cap \mathcal{P}_2(x) \right\}$$

Valid Inequalities

Valid inequality: The triple $(\alpha^x, \alpha^y, \beta) \in \mathbb{R}^{n_1+n_2+1}$ is a *valid inequality* for (MIBLP) if

$$\mathcal{F} \subseteq \left\{ (x, y) \in \mathbb{R}^{n_1 \times n_2} \;\middle|\; \alpha^x x + \alpha^y y \geq \beta \right\}.$$

Valid improving inequality: The triple $(\alpha^x, \alpha^y, \beta) \in \mathbb{R}^{n_1+n_2+1}$ is a *valid improving inequality* for (MIBLP) with respect to $(\bar{x}, \bar{y}) \in \mathcal{F}$ if

$$\left\{ (x, y) \in \mathcal{F} \;\middle|\; cx + d^1 y < c\bar{x} + d^1 \bar{y} \right\} \subseteq \left\{ (x, y) \in \mathbb{R}^{n_1 \times n_2} \;\middle|\; \alpha^x x + \alpha^y y \geq \beta \right\}.$$

Cut Generation

- As usual, a *cutting plane* (cut) refers to a valid inequality violated by a given (infeasible) solution to the current relaxation.
- Unlike in MILP, we have several distinctly different classes of solutions with respect to which we might want to generate a cut.
  1. $(\bar{x}, \bar{y}) \notin X \times Y$
  2. $(\bar{x}, \bar{y}) \in X \times Y$, but for which $d^2\bar{y} > \phi(b^2 - A^2\bar{x})$
  3. All $(x, y) \in \mathcal{P}$ with $x_L = \lambda \in \mathbb{Z}^L$
- The set of valid inequalities for MIBLPs can be classified based on the types of the solutions they remove.

Very Rough Classification

- In general, cuts may be generated by separating solutions from a polyhedron obtained by *convexification*.
- Inequalities valid for the following polyhedra correspond roughly to the classes defined on the previous slide.

Classes of inequalities

- Inequalities valid for $\text{conv}(\mathcal{S}) \Rightarrow$ Feasibility Cuts
- Inequalities valid for $\text{conv}(\mathcal{F}) \Rightarrow$ Optimality Cuts
- Inequalities valid for $\text{conv}\left(\{(x \in \mathcal{F}_1 \mid cx + \Xi(x) < U\}\right) \Rightarrow$ Projected Optimality Cuts

where

$$\mathcal{F}_1 = \text{proj}_x(\mathcal{F})$$

and Ξ is the so-called “risk” function that captures the impact of the second level decision.

Cut Generation Recipes

Generalized Chvátal Cuts

Theorem 1 *If* $(\pi^x, \pi^y) \in X \times Y$ *is such that* $\pi^x_i = 0$ *for* $i > r_1$ *and* $\pi^y_j = 0$ *for* $j > r_2$, $\beta \in \mathbb{Z}$, *and* $\{(x, y) \in \mathcal{P} \mid \pi^x x + \pi^y y \leq \beta - 1\}$ *contains no bilevel feasible solutions, then* $\pi^x x + \pi^y y \geq \beta$ *is a valid inequality.*

Disjunctive Cuts

Theorem 2 *Let* $X_1, X_2, \ldots, X_k$ *be a set representing a disjunction valid for* $\mathcal{F}$. *Then if* $(\alpha^x, \alpha^y, \beta)$ *is valid for* $\mathrm{conv}(\cup_{1 \leq i \leq k} X_i)$, *it is valid for* $\mathcal{F}$

Intersection Cuts

Theorem 3 *Let* C *be a convex set containing no improving solutions and let* (x, y) *be an extreme point of* $\mathcal{P}$ *in the interior of* C. *Then the intersection cut with respect to* C *and* (x, y) *is valid for* $\mathcal{F}$.

Known Classes

- **Feasibility cuts:** This set includes inequalities valid for MILPs.
- **Optimality cuts:**
  - Integer no-good cut [DeNegre and Ralphs, 2009]
  - Intersection cuts [Fischetti et al., 2017]
  - Bound cuts
    - Benders cut [Ralphs et al., 2015, Caprara et al., 2014]
    - Increasing objective cut [DeNegre, 2011]
    - Dual function cuts
- **Projected optimality cuts:**
  - Generalized no-good cut

Integer No-good Cut

Assumptions:

- $r_1 = n_1$ and $r_2 = n_2$.
- Vectors b^1 and b^2 and all matrices A^1, A^2, G^1 and G^2 are discrete.

Theorem 4 *Let $(\bar{x}, \bar{y}) \in \mathcal{S} \setminus \mathcal{F}$ be the optimal solution of relaxation problem and H_1 and H_2 denote the set of indices of the first- and second-level constraints, respectively, binding at $(\bar{x}, \bar{y})$. Then, under the stated assumptions, we have*

$$\alpha^x x + \alpha^y y \geq \beta \quad \forall (x, y) \in \mathcal{F},$$

where

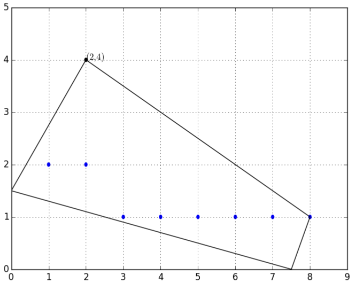
$$\alpha^x = \sum_{i \in H_1} a_i^1 + \sum_{i \in H_2} a_i^2, \alpha^y = \sum_{i \in H_1} g_i^1 + \sum_{i \in H_2} g_i^2, \beta = \sum_{i \in H_1} b_i^1 + \sum_{i \in H_2} b_i^2 + 1$$

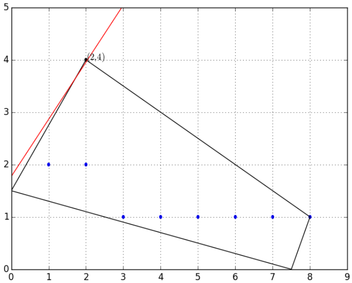
and $a_i^1, a_i^2, g_i^1, g_i^2$, b_i^1 and b_i^2 represent the i^{th} rows of A^1, A^2, G^1, G^2, b^1 and b^2, respectively. Furthermore, we have

$$\alpha^x \bar{x} + \alpha^y \bar{y} = \beta - 1,$$

so the inequality is violated by $(\bar{x}, \bar{y})$.

Integer No-good Cut





Benders Cut

Assumptions:

- $x_L \subseteq \mathbb{B}^L$.
- There exists a second-level variable y_i corresponding to each linking variable x_i so that $x_i = 1$ results $y_i = 0$.
- $G^2 \leq 0$.

Theorem 5 *Let $(\hat{x}, \hat{y}) \in \mathcal{F}$. Then, under the above assumptions, we have*

$$d^2 y \leq \sum_{i=1}^{n_1} d_i^2 \hat{y}_i (1 - x_i) \quad \forall (x, y) \in \mathcal{F}.$$

Furthermore, this inequality is violated by all $(x, y) \in \mathcal{S} \setminus \mathcal{F}$ with $x = \hat{x}$.

Increasing Objective Cut

Assumptions:

- $x_L \subseteq \mathbb{B}^L$.
- $A^2 \leq 0$.

Theorem 6 *Let* $\hat{x} \in X$ *and* $\hat{y} \in \mathcal{R}(\hat{x})$. *Then, under the above assumptions, we have*

$$d^2 y \leq d^2 \hat{y} + M \left(\sum_{i \in L : x_i = 0} \hat{x}_i \right) \quad \forall (x, y) \in \mathcal{F}.$$

Furthermore, this inequality is violated by all $(x, y) \in \mathcal{S} \setminus \mathcal{F}$ *with* $x_i = 0$ *for* $\{i \in L \mid \hat{x}_i = 0\}$.

Intersection Cut

Assumption:

- $A^2x + G^2y - b^2 \in \mathbb{Z}$ for all $(x, y) \in \mathcal{S}$.
- $d^2y \in \mathbb{Z}$ for all $(x, y) \in \mathcal{S}$.

Theorem 7 *Let* $(\bar{x}, \bar{y}) \in \mathcal{S} \setminus \mathcal{F}$ *be the optimal solution of the relaxation problem and* $\hat{y} \in Y$ *satisfy these conditions:*

- $d^2\hat{y} < d^2\bar{y}$.
- $G^2\hat{y} \geq b^2 - A^2\bar{x}$.

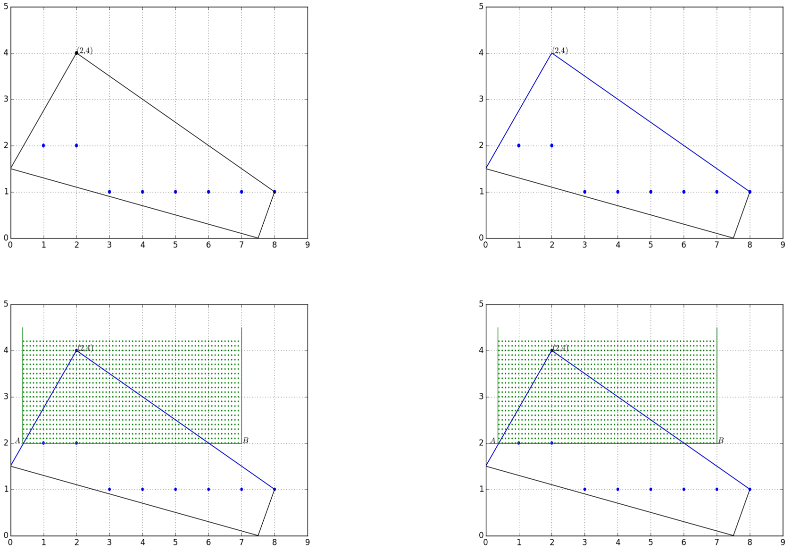
Then, under the stated assumptions, we have

$$\alpha^x x + \alpha^y y \geq \beta \quad \forall (x, y) \in \mathcal{F}, \tag{2}$$

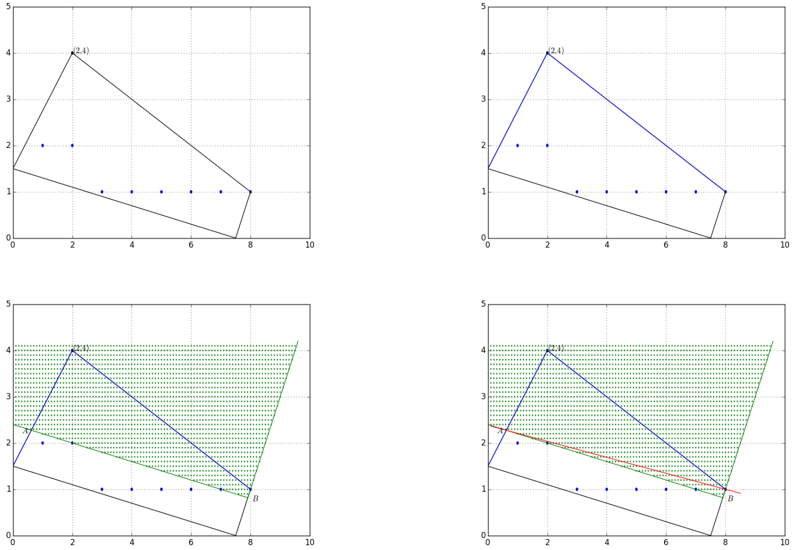
where the inequality (2) *is the intersection cut generated associated with* $(\bar{x}, \bar{y})$ *and the set*

$$\mathcal{C} = \left\{ (x, y) \in \mathbb{R}^{n_1 \times n_2} \;\middle|\; d^2y \geq d^2\hat{y}, G^2\hat{y} \geq b^2 - A^2x - 1 \right\}. \tag{3}$$

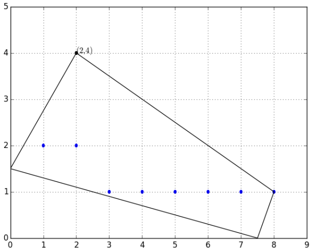
Intersection Cut

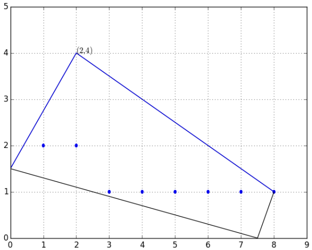


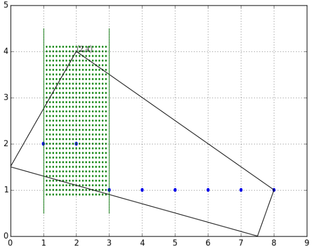
Watermelon Intersection Cut

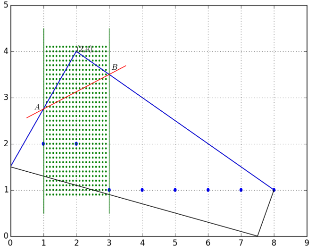


Hypercube Intersection Cut









Generalized No-good Cut

Assumption:

- $x_L \subseteq \mathbb{B}^L$.

Theorem 8 *Let $\gamma \in \mathbb{B}^L$. Then, under the desired assumption, we have*

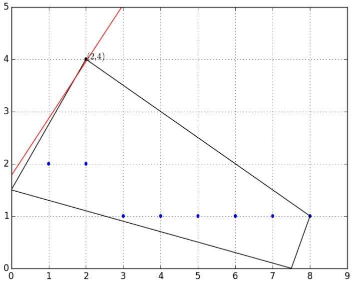
$$\sum_{i \in L: \gamma_i = 0} x_i + \sum_{i \in L: \gamma_i = 1} (1 - x_i) \geq 1 \quad \forall (x, y) \in \mathcal{F} \text{ such that } x_L \neq \gamma. \tag{4}$$

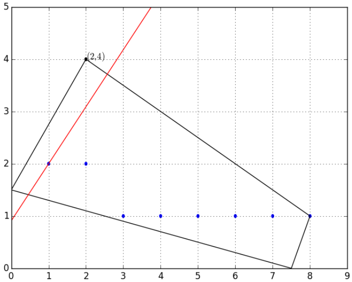
Furthermore, the inequality (4) *is violated by all $(x, y) \in \mathcal{P}$ with $x_L = \gamma$.*

- The problem (UB) should be solved with $\gamma = \bar{x}_L$ prior to generating this cut.

Strengthening the Valid Inequalities

- **Idea:** A given valid inequality can be strengthened by increasing the value of the RHS.
- Ideally, we can do this until the inequality supports $\text{conv}(\mathcal{F})$.





- However, this need not be done optimally.

The Gauge Function

- The *gauge function* is a function that returns the largest valid right-hand side for an inequality with a given left-hand side vector.

Gauge Function

$$\Gamma(\zeta^x, \zeta^y) = \min_{(x,y)\in\mathcal{F}} \zeta^x x + \zeta^y y \tag{5}$$

- Note that the gauge function is closely related to the value function.
- We have that

$$\alpha^x x + \alpha^y y \geq \Gamma(\alpha^x, \alpha^y) \;\; \forall (x, y) \in \mathcal{F}$$

Valid Inequality

An *inequality* defined by $(\alpha^x, \alpha^y, \beta) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}$ is *valid* for $\mathcal{F}$ if $\beta \leq \Gamma(\alpha^x, \alpha^y)$.

Dual Function Cuts

- The gauge function need not be computed exactly, we only need a bound.
- If $F : \mathbb{R}^{m_1+m_2} \to \mathbb{R}$ is a *dual function* for the problem (5) (bounds its value function from below), then

$$\alpha^x x + \alpha^y y \geq F(\alpha^1, \alpha^2) \ \ \forall (x, y) \in \mathcal{F},$$

where $(\alpha^x, \alpha^y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}$.

Strengthened Integer No-Goods

Assumption:

- $x_L \subseteq \mathbb{Z}^L$.

Theorem 9 *Let* $(\bar{x}, \bar{y}) \in \mathcal{S} \setminus \mathcal{F}$ *be the optimal solution of relaxation problem and* H_1 *and* H_2 *denote the set of indices of the first- and second-level constraints, respectively, binding at* $(\bar{x}, \bar{y})$*. Then, under the stated assumptions, we have*

$$\alpha^x x + \alpha^y y \geq \beta \quad \forall (x, y) \in \mathcal{F},$$

where

$$\alpha^x = \sum_{i \in H_1} a_i^1 + \sum_{i \in H_2} a_i^2, \alpha^y = \sum_{i \in H_1} g_i^1 + \sum_{i \in H_2} g_i^2, \beta = \min_{(x,y) \in \mathcal{F}} \alpha x$$

and $a_i^1, a_i^2, g_i^1, g_i^2$, b_i^1 *and* b_i^2 *represent the* i^{th} *rows of* A^1, A^2, G^1, G^2, b^1 *and* b^2, *respectively. Furthermore, we have*

$$\alpha^x \bar{x} + \alpha^y \bar{y} < \beta,$$

so the inequality is violated by $(\bar{x}, \bar{y})$.

Software Framework

MibS software

- is an open-source solver for MIBLPs.
- works based on our branch-and-cut algorithm for MIBLPs.
- is implemented in C++.
- is built on top of the BLIS solver [Xu et al., 2009].
- employs different software available from the *Computational Infrastructure for Operations Research (COIN-OR)* repository
    - *COIN High Performance Parallel Search (CHiPPS)*: To manage the global branch-and-bound
    - *SYMPHONY*: To solve the required MIPs
    - *COIN LP Solver (CLP)*: To solve the LPs arising in the branch and cut
    - *Cut Generation Library (CGL)*: To generate cutting planes within both SYMPHONY and MibS
    - *Open Solver Interface (OSI)*: To interface with other solvers

Computational Results

Table: The summary of data sets

Data Set	First-level Vars Type	Second-level Vars Type	Size	Interdiction
INTERD-DEN	binary	binary	300	yes
IBLP-DEN	discrete	discrete	50	no
IBLP-XU	discrete	discrete	100	no

Computational Results

- Performance of different cuts on the instances of `INTERD-DEN`
- 199 instances that can be solved by at least one method in 3600 seconds and whose solution time exceeds 5 seconds for at least one method.

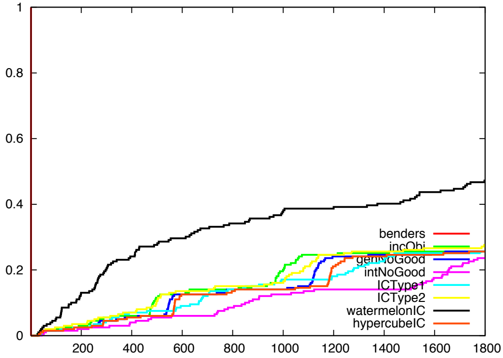


Figure: Performance profile with *solution time* as the performance measure

Computational Results

- Performance of different cuts on the instances of `INTERD-DEN`
- 199 instances that can be solved by at least one method in 3600 seconds and whose solution time exceeds 5 seconds for at least one method.
- Benders cut was not considered.

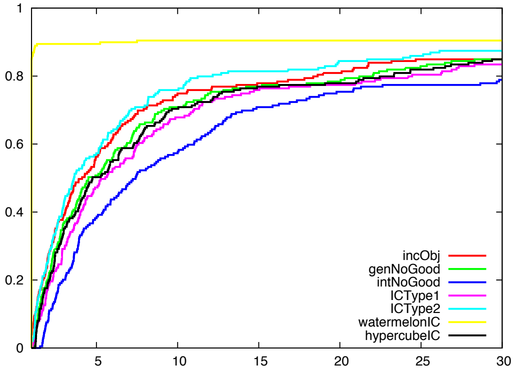


Figure: Performance profile with *solution time* as the performance measure

Computational Results

- Performance of different cuts on the instances of `IBLP-DEN`
- 23 instances that can be solved by at least one method in 3600 seconds and whose solution time exceeds 5 seconds for at least one method.

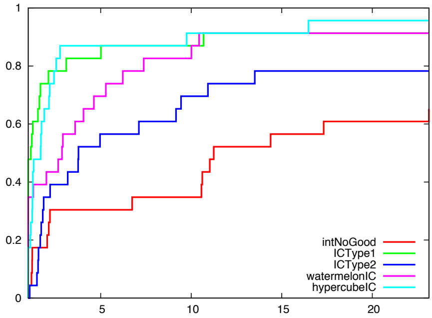


Figure: Performance profile with *solution time* as the performance measure

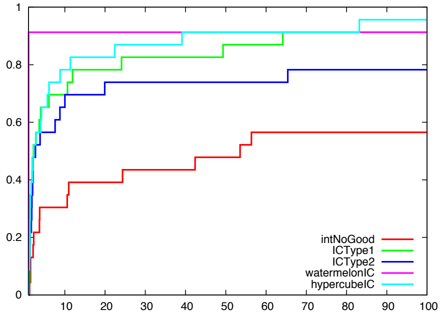


Figure: Performance profile with *number of nodes* as the performance measure

Computational Results

- Performance of different cuts on the instances of `IBLP-XU`
- 77 instances that can be solved by at least one method in 3600 seconds and whose solution time exceeds 5 seconds for at least one method.

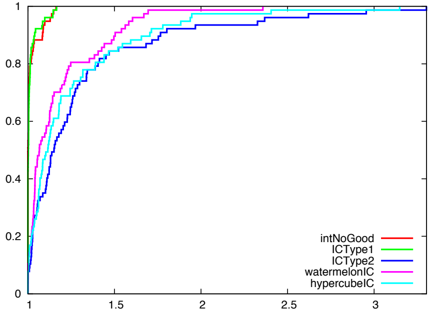


Figure: Performance profile with *solution time* as the performance measure

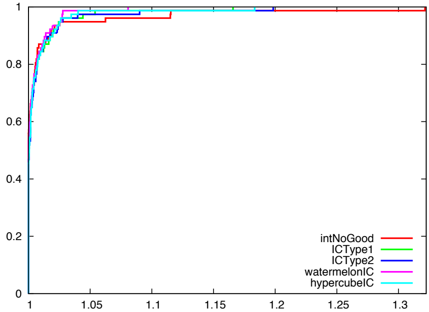


Figure: Performance profile with *number of nodes* as the performance measure

Computational Results

- Comparison of the rhs of different cuts with the best rhs
- 5 instances of INTERD-DEN with 20 variables were considered.

	Benders Cut		Watermelon IC		Integer No-good Cut	
	Orig RHS	Best RHS	Orig RHS	Best RHS	Orig RHS	Best RHS
1	-4520	-4520	-1	-1.65	-1	-7
2	-3917	-3917	-1	-1.27	-1	-6
3	-3334	-3334	-1	-1	-1	-6
4	-3915	-3915	-1	-1.01	-1	-6
5	-4924	-4924	-1	-1.01	-1	-6

Computational Results

- Impact of improving the rhs of integer no-good cut on the instances of `IBLP-DEN`
- 180 instances whose solution time is more than 5 seconds by using the integer no-good cut.

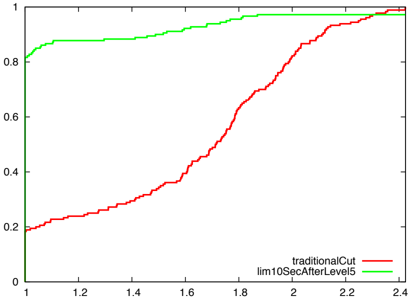


Figure: Performance profile with *solution time* as the performance measure

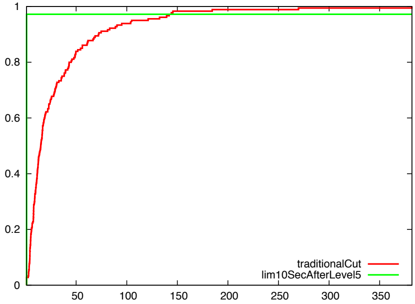


Figure: Performance profile with *number of nodes* as the performance measure

Computational Results

- Impact of improving the rhs of integer no-good cut on the instances of `IBLP-DEN`
- 14 instances whose solution time is more than 5 seconds by using the integer no-good cut.

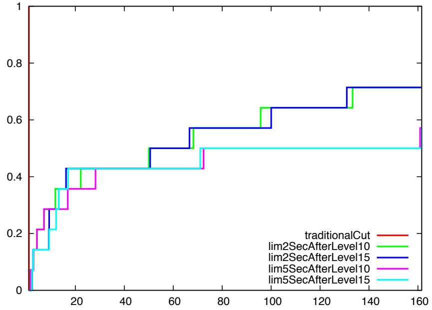


Figure: Performance profile with *solution time* as the performance measure

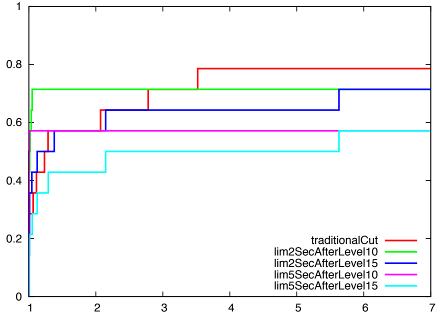


Figure: Performance profile with *number of nodes* as the performance measure

References I

A. Caprara, M. Carvalho, A. Lodi, and G.J. Woeginger. Bilevel knapsack with interdiction constraints. Technical Report OR-14-4, University of Bologna, 2014.

Scott T DeNegre. *Interdiction and Discrete Bilevel Linear Programming*. Phd, Lehigh University, 2011. URL `http://coral.ie.lehigh.edu/{~}ted/files/papers/ScottDeNegreDissertation11.pdf`.

S.T. DeNegre and T.K. Ralphs. A branch-and-cut algorithm for bilevel integer programming. In *Proceedings of the Eleventh INFORMS Computing Society Meeting*, pages 65–78, 2009. doi: 10.1007/978-0-387-88843-9_4. URL `http://coral.ie.lehigh.edu/~ted/files/papers/BILEVEL08.pdf`.

Matteo Fischetti, Ivana Ljubić, Michele Monaci, and Markus Sinnl. A new general-purpose algorithm for mixed-integer bilevel linear programs. *Operations Research*, 65(6):1615–1637, 2017.

T.K. Ralphs, S. Tahernejad, S. DeNegre, M. Guzelsoy, and A. Hassanzadeh. Bilevel integer optimization: Theory and algorithms. Technical report, COR@L Laboratory, Lehigh University, 2015.

References II

S.A. Tahernejad, T.K. Ralphs, and S.T. DeNegre. A branch-and-cut algorithm for mixed integer bilevel linear optimization problems and its implementation. Technical report, COR@L Laboratory Technical Report 16T-015-R3, Lehigh University, 2016. URL `http://coral.ie.lehigh.edu/~ted/files/papers/MIBLP16.pdf`.

Y. Xu, T.K. Ralphs, L. Ladányi, and M.J. Saltzman. Computational experience with a software framework for parallel integer programming. *The INFORMS Journal on Computing*, 21:383–397, 2009. doi: 10.1287/ijoc.1090.0347. URL `http://coral.ie.lehigh.edu/~ted/files/papers/CHiPPS-Rev.pdf`.

Thank You!