Perspective Cuts for Generalized On/Off Constraints

Ksenia Bestuzheva[1], Ambros Gleixner[1,2], Stefan Vigerske[3]

[1]Zuse Institute Berlin, [2]HTW Berlin, [3]GAMS

bestuzheva@zib.de, gleixner@zib.de, svigerske@gams.com

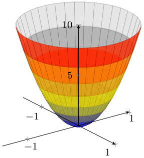
Mixed Integer Programming Workshop
May 24, 2023

SPONSORED BY THE Federal Ministry of Education and Research

Mixed-Integer Nonlinear Programming

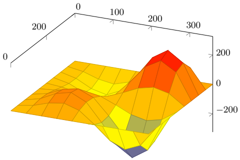
$$
\begin{aligned}
\min\ & c^{\mathsf{T}} x \\
\text{s.t. } & g_k(\mathbf{x},\mathbf{y},\mathbf{z}) \le 0\ \forall k \in \mathcal{C}, \\
& (\underline{x}_i^1 - x_i^0) z_k \le x_i - x_i^0 \le (\overline{x}_i^1 - x_i^0) z_k,\ \ \forall i \in \mathcal{S}_k,\ \ \forall k \in \mathcal{I}, \\
& \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}],\ \ \mathbf{y} \in [\underline{\mathbf{y}}, \overline{\mathbf{y}}], \\
& \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^p, \mathbf{z} \in \{0,1\}^q.
\end{aligned}
$$

- The functions $g_k : [\underline{\mathbf{x}}, \overline{\mathbf{x}}] \times [\underline{\mathbf{y}}, \overline{\mathbf{y}}] \times \{0,1\}^q \to \mathbb{R}$ can be



convex

or



nonconvex

and are given in **algebraic form**.

- Our approaches are aimed to be applied within an **LP-based spatial branch & bound** algorithm.

Semicontinuous Variables

SC variables $\mathbf{x}$ are defined by the following relations:

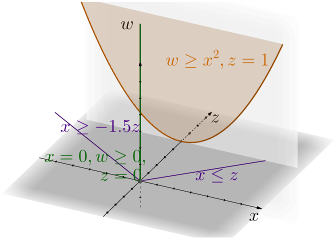
$$(\underline{\mathbf{x}}^1 - \mathbf{x}^0)z \leq \mathbf{x} - \mathbf{x}^0 \leq (\overline{\mathbf{x}}^1 - \mathbf{x}^0)z,$$
$$z \in \{0, 1\},$$

where z - **indicator variable**. This implies:

$$\mathbf{x} = \mathbf{x}^0 \text{ if } z = 0,$$
$$\mathbf{x} \in \left[\underline{\mathbf{x}}^1, \overline{\mathbf{x}}^1\right] \text{ if } z = 1.$$

- The implication may be present in the problem implicitly
- SC variables can be used for describing “on” and “off” states

Constraints with SC Variables



Consider the epigraph set:

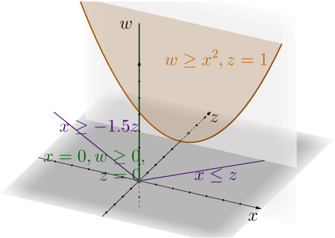
$$g(\mathbf{x}) \leq w,$$
$$(\underline{\mathbf{x}}^1 - \mathbf{x}^0)z \leq \mathbf{x} - \mathbf{x}^0 \leq (\overline{\mathbf{x}}^1 - \mathbf{x}^0)z,$$
$$z \in \{0, 1\}$$

Example:

$$g(x) = x^2 \leq w, \quad -1.5z \leq x \leq z$$

Disjunctive Formulation

- Consider continuous relaxations ($z \in [0, 1]$) of an SC constraint
- Taking into account the semi-continuity of $\mathbf{x}$ is crucial for constructing tight relaxations
- Represent the feasible set of the SC constraint via a disjunctive formulation:



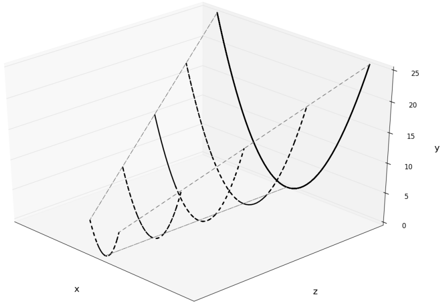
$$S^0 = \{(\mathbf{x}, w, z) \mid \mathbf{x} = \mathbf{x}^0,\ g(\mathbf{x}^0) \leq w,\ \mathbf{z = 0}\},$$
$$S^1 = \{(\mathbf{x}, w, z) \mid \mathbf{x} \in [\underline{\mathbf{x}}^1, \overline{\mathbf{x}}^1],\ g(\mathbf{x}) \leq w,\ \mathbf{z = 1}\},$$
$$S = S^0 \cup S^1.$$

- We are interested in finding the convex hull of S

The Perspective Function

$$\tilde{g}(\mathbf{x}, z) = \begin{cases} zg(\frac{\mathbf{x}}{z}) & \text{if } z > 0, \\ +\infty & \text{otherwise} \end{cases}$$

- epi($\tilde{g}$) is a cone generated by epi(g)
- the perspective operator preserves convexity
- $\tilde{g}$ is not well-defined at $z = 0$, but usually this can be circumvented



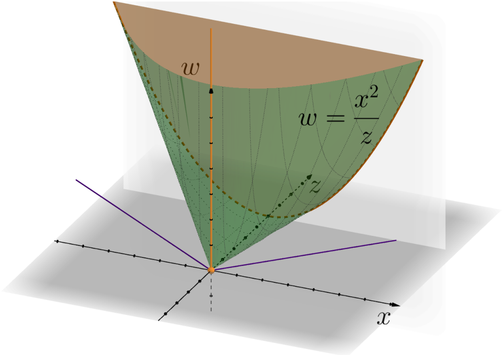
Several dilations of the function $y = x^2$

Perspective Reformulation

If g is **convex**, then $conv(S)$ can be described with the use of the perspective function:

$$cl\{\tilde{g}(\mathbf{x}, z) \leq w\},$$
$$(\underline{\mathbf{x}}^1 - \mathbf{x}^0)z \leq \mathbf{x} - \mathbf{x}^0 \leq (\overline{\mathbf{x}}^1 - \mathbf{x}^0)z,$$

[Günlük, Linderoth'10]



- The closure is necessary since $\tilde{g}$ is not well-defined at 0
- Linearize the perspective formulation at $(\mathbf{x}^*, z^*)$ $\Rightarrow$ perspective cuts [Frangioni, Gentile'06]

Our Contributions

- A **cut strengthening procedure** for SC constraints:
  - requires a valid linear inequality,
  - can be applied to convex and **nonconvex** constraints;
- a **further generalisation** for a broader class of constraints:
  - valid for constraints that become redundant when $z = 0$,
  - i.e. can be used to strengthen outer approximations of big-M constraints;
- a **computational study** of perspective cuts:
  - the cuts were implemented within a general-purpose solver (SCIP),
  - we used a large heterogeneous test set (MINLPLib).

Existing Results for More General Sets

- Lifted space formulation for a union of a finite number of convex sets [Ceria, Soares'99]
- Original space formulation for a union of a finite number of orthogonal convex sets [Tawarmalani, Richard, Chung'10]
- Original space formulation for a union of a box and a convex set given by an isotone function [Hijazi et al.'10]
  - Number of constraints exponential in number of variables
  - A compact relaxation works well in practice

Perspective-Based Cut Strengthening for Nonconvex Constraints

Given **any valid** linear inequality $a\mathrm{x} + b \le w$ for the set $\{(\mathrm{x}, w) \mid g(\mathrm{x}) \le w,\ \mathrm{x} \in [\underline{\mathrm{x}}^1, \overline{\mathrm{x}}^1]\}$, where x is semicontinuous, the inequality

$$a\mathrm{x} + b + (a\mathrm{x}^0 + b - g(\mathrm{x}^0))(z - 1) \le w \qquad (*)$$

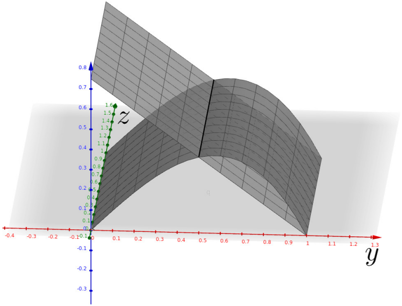
is valid for the disjunctive set

$$\{\mathrm{x} = \mathrm{x}^0,\ g(\mathrm{x}^0) \le w,\ z = 0\} \cup \{\mathrm{x} \in [\underline{\mathrm{x}}^1, \overline{\mathrm{x}}^1],\ g(\mathrm{x}) \le w,\ z = 1\}.$$

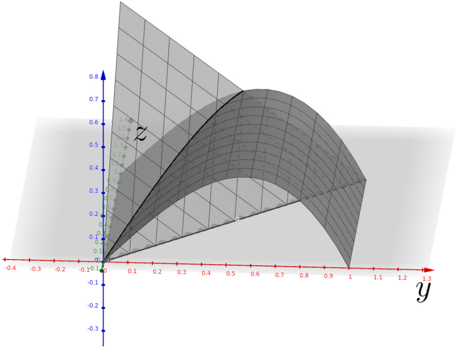
- The strengthening procedure **does not depend on convexity** of g
- $a\mathrm{x} + b \le w$ only needs to be valid when $\mathrm{x} \in [\underline{\mathrm{x}}^1, \overline{\mathrm{x}}^1]$ (also if $\mathrm{x}^0 \notin [\underline{\mathrm{x}}^1, \overline{\mathrm{x}}^1]$)
  - Can set $z = 1$ and perform bound propagation to find tighter bounds; tighter bounds $\rightarrow$ tighter cut
- If g is convex, cut $(*)$ is equivalent to the perspective cut from [Frangioni, Gentile, 2006]

Example Of Cut Strengthening

Cut valid for $z = 1$



Cut valid for $z \in \{0, 1\}$



Further Generalisation: Union of a Nonlinear Set and a Box

- Allow **non-semicontinuous** variables in the constraint
- Require that the constraint becomes **redundant** when $z = 0$

That is, consider sets of the form:

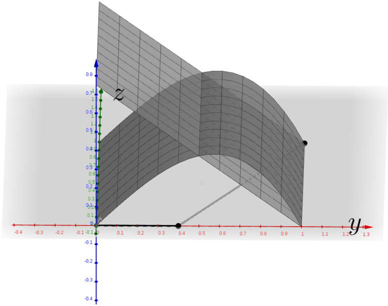
$$S^0 = \{(w, \mathbf{x}, z) \mid w \in [\underline{w}^0, \overline{w}^0],\ \mathbf{x} \in [\underline{\mathbf{x}}^0, \overline{\mathbf{x}}^0],\ z = 0\},$$
$$S^1 = \{(w, \mathbf{x}, z) \mid g(\mathbf{x}) \leq w,\ \mathbf{x} \in [\underline{\mathbf{x}}^1, \overline{\mathbf{x}}^1],\ z = 1\},$$

where $g(\mathbf{x}) \leq w$ can be viewed as an **on/off constraint** controlled by indicator z.

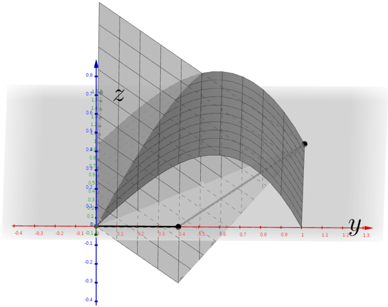
The definition focuses on the properties of the disjunctive set rather than the algebraic formulation $\rightarrow$ detection less dependent on formulation.

Example Of Cut Strengthening - Box Case

Cut valid for $z = 1$



Cut valid for $z \in \{0, 1\}$



Cut Strengthening for Non-SC On/Off Constraints

The procedure is an extension of the procedure described earlier:

Given **any valid** linear inequality $a\mathbf{x} + b \leq w$ for the set $\{(\mathbf{x}, w) \mid g(\mathbf{x}) \leq w,\ \mathbf{x} \in [\underline{\mathbf{x}}^1, \overline{\mathbf{x}}^1]\}$, we look for the tightest inequality of the form

$$a\mathbf{x} + b + \alpha(z-1) \leq w \qquad (*)$$

that is valid for the disjunctive set

$$\{\mathbf{w} \in [\underline{\mathbf{w}}^{\mathbf{0}}, \overline{\mathbf{w}}^{\mathbf{0}}],\ \mathbf{x} \in [\underline{\mathbf{x}}^{\mathbf{0}}, \overline{\mathbf{x}}^{\mathbf{0}}],\ z = 0\} \cup \{\mathbf{x} \in [\underline{\mathbf{x}}^1, \overline{\mathbf{x}}^1],\ g(\mathbf{x}) \leq w,\ z = 1\}.$$

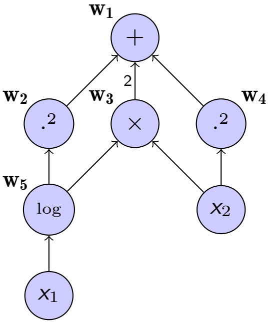
The largest α that maintains validity is:

$$\alpha^* = \min(w - a\mathbf{x} - b) \text{ s.t. } (\mathbf{x}, w) \in [\underline{w}^0, \overline{w}^0] \times [\underline{\mathbf{x}}^0, \overline{\mathbf{x}}^0].$$

Nonlinear Constraints in SCIP

- Expressions are represented as **expression graphs**,
- **Auxiliary variables** are introduced for subexpressions, used in **relaxations only**
- The **original formulation** is kept
- **Nonlinear handler** plugins implement specialized algorithms for specific structures
- The extended formulation has the form:

$$h_i(\mathbf{x}, \mathbf{y}, w_1, \ldots, w_{i-1}, \mathbf{z}) \lesseqqgtr w_i, \qquad i = 1, \ldots, m',$$
$$\underline{\mathbf{w}} \leq \mathbf{w} \leq \overline{\mathbf{w}},$$



Implementation of Strengthening for SC Constraints

- Detect SC variables:
  - analyse linear relations directly,
  - detect implications of $z_k = 0$ and $z_k = 1$ by fixing z_k and propagating all constraints.
- Detect constraints of the extended formulation:

$$h_i(\mathbf{x}, \mathbf{y}, w_1, \ldots, w_{i-1}, \mathbf{z}) = h^{sc}_{i,k}(\mathbf{x}, w_1, \ldots, w_r, z_k) + h^{nsc}_{i,k}(\mathbf{y}, w_{r+1}, \ldots, w_{i-1}, \mathbf{z}_{\setminus k}) \lesseqqgtr w_i,$$

  where h^{nsc} is linear, and all variables in h^{sc} are semi-continuous with respect to the same indicator variable.
- Dynamically separate cuts:
  - set $z_k = 1$ and propagate bounds,
  - request valid cuts from other nonlinear handlers,
  - apply the strengthening procedure to the parts of the cuts that depend on $(\mathbf{x}, w_1, \ldots, w_r, z_k)$.

Implementation of Strengthening for Non-SC On/Off Constraints

- Detect constraints of the extended formulation

$$h_i(\mathbf{x}, \mathbf{y}, w_1, \ldots, w_{i-1}, \mathbf{z}) \overset{\leq}{\underset{\geq}{=}} w_i,$$

  which become redundant when some indicator variable $z_k = 0$. E.g., for a '$\leq$' constraint:
  - use interval arithmetic to compute an upper bound $\overline{h}^0$ on h over the domain corresponding to $z_k = 0$
  - if $\overline{h}^0 \leq \underline{w}_i^0$, then h_i is an on/off constraint.
- Dynamically separate cuts by applying the generalized strengthening formula.

Evaluation of Generalisation 1: Computational Setup

- Selected 186 MINLPLib instances that contain suitable structures for applying Generalisation 1
- 4 permutations of each instance + default
- Time limit one hour

Computational Results: Summary for Generalisation 1

Instances with SC structures

All	Convex	Both	Nonconvex
186	89	53	44

Solved and failed instances

	Off	Convex	Full
Solved	741	**764**	759
Limit	175	**154**	**154**
Fails	14	**12**	17

Geometric means of time and nodes

	Off	Convex	Full
Time	13.79	**11.23**	11.27
Relative time	1.00	**0.81**	0.82
Nodes	620	479	**472**
Relative nodes	1.00	0.77	**0.76**

Instances with improvement in root node dual bound

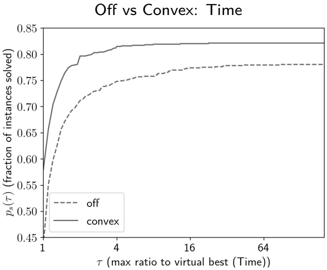
	Off	Convex	Convex	Full
Better by $> 50\%$	16	**46**	0	**31**
Better by $5 - 50\%$	25	**39**	**14**	11
Same within 5%	584		429	

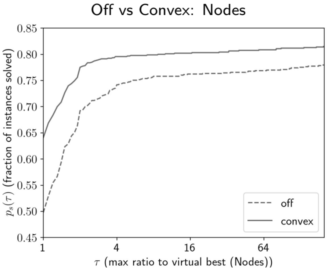
Detailed Evaluation for Generalisation 1

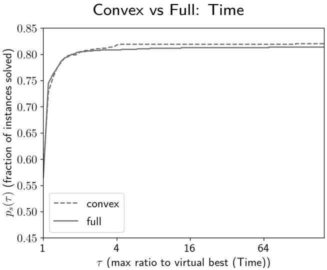
Table: Time on subsets of affected instances

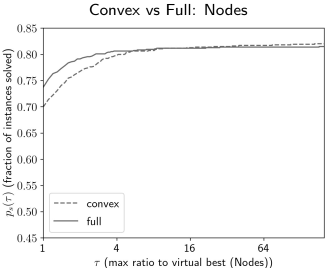
	Off	Convex	Convex	Full
Instances in $[0, 3600]$:	544		205	
Time	12.53	**9.70**	**24.30**	24.82
Relative time	**1.00**	0.77	**1.00**	1.02
Faster	95	**193**	43	**51**
Instances in $[10, 3600]$:	276		149	
Time	70.96	**45.27**	**57.47**	59.12
Relative time	1.00	**0.64**	**1.00**	1.03
Faster	50	**122**	29	**35**
Instances in $[100, 3600]$:	100		49	
Time	506.17	**183.90**	**263.57**	285.85
Relative time	**1.00**	0.36	**1.00**	1.08
Faster	18	**64**	13	**15**
Instances in $[1000, 3600]$:	45		14	
Time	1444.28	**425.60**	**814.18**	1034.83
Relative time	1.00	**0.29**	**1.00**	1.27
Faster	10	**32**	5	5

Computational Results: Performance Profiles





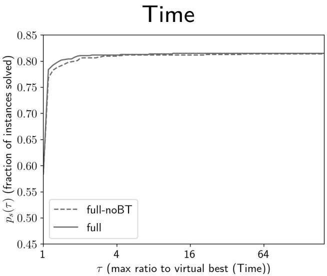


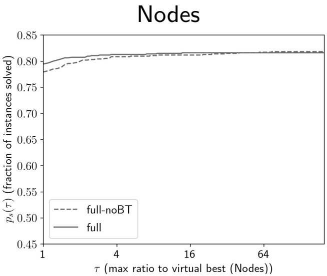


Effect of Tighter Bounds

Comparison between Full-noBT and Full

	Fails	Limit	Solved	RootImpr $> 50\%$	Time	Nodes
Full-noBT	16	153	**761**	4	34.45	2910
Full	17	154	759	**25**	**33.68**	**2618**





Cuts vs Reformulations: Setup

- Tested on instances where a perspective reformulation is available
- squfl* instances: second order cone formulations [Günlük, Linderoth'10]
- rsyn* and syn* instances: ϵ-formulations [Furman, Sawaya, Grossman'18]:
  - replace $\tilde{g}$ with $(\epsilon + (1-\epsilon)z)g(\frac{\mathrm{x}}{\epsilon+(1-\epsilon)z}) - \epsilon g(0)(1-z)$
  - under mild conditions, this is an equivalent reformulation
  - reformulated instances were not recognized as convex by SCIP $\rightarrow$ forced convexity recognition

Cuts vs Reformulations: Results

Table: Comparison of perspective cuts and perspective reformulations on 310 convex instances

	Full	Reformulated	Reformulated-convex
Solved	**308**	253	305
Time	7.08	18.02	**2.80**
Relative time	1.00	2.55	**0.40**
Nodes	261	797	**4.5**
Relative nodes	1.00	3.05	**0.02**

- The benefits of reformulations are not reduced to the benefits of the resulting cuts, even in the context of LP-based BB
- ϵ-reformulations work very well
- Reformulations must be treated carefully so that not to introduce numerical inaccuracies

Evaluation of Generalisation 2

- Selected 173 MINLPLib instances that contain suitable structures for applying Generalisation 2 that do not fit Generalisation 1
- Time limit one hour

Table: Comparison between Full and Full+Box

	Fails	Limit	Solved	Time	Nodes
Full	3	60	110	**16.79**	**816**
Full+Box	3	58	**112**	17.85	880

Conclusions

- We have proposed two generalisations of perspective cuts:
  - for nonconvex constraints defined by semi-continuous variables,
  - for constraints describing a union of a nonlinear set and a box
- Extensive testing on a large heterogeneous test set with a general-puspose solver confirms the effectiveness of perspective cuts
- Cuts for generalized structures have less impact than cuts for convex SC constraints
- Nonlinear reformulations can be stronger than just cuts
- Detection can be further improved