MINLP and Stronger Relaxation of Bilinear Terms

Benjamin Müller, Ambros Gleixner, and Felipe Serrano

Zuse Institute Berlin · serrano@zib.de
SCIP Optimization Suite · http://scip.zib.de

SCIP Workshop

Aachen · March 7, 2018







Outline

Introduction: LP-based Branch and Bound

McCormick Relaxation Spatial Branch and bound Bound tightening

Improving over McCormick

Computational results

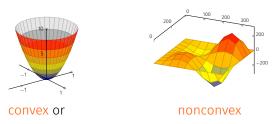
Conclusion



Mixed-Integer Nonlinear Programs (MINLPs)

$$\begin{aligned} & \text{min } c^{\mathsf{T}} X \\ & \text{s.t. } g_k(x) \leq 0 & \forall k \in [m] \\ & x_i \in \mathbb{Z} & \forall i \in \mathcal{I} \subseteq [n] \\ & x_i \in [\ell_i, u_i] & \forall i \in [n] \end{aligned}$$

The functions $g_k \in C^1([\ell, u], \mathbb{R})$ can be



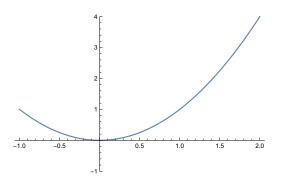
To solve to global optimality, SCIP uses LP-based spatial branch and bound

LP based spatial Branch & Bound

- \cdot Try to build a polyhedral relaxation ${\cal R}$
- Solve \mathcal{R} and get solution x^*
- If x^* is feasible we are done. If not,
- Strengthen \mathcal{R} by separating x^*
- · When not possible, branch spatially (i.e., on continuous variables)

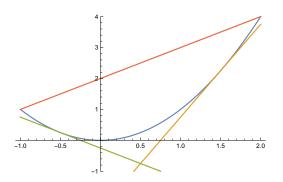


Consider a constraint $y = x^2, x \in [-1, 2]$



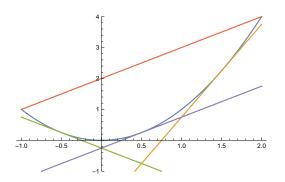


Consider a constraint $y = x^2, x \in [-1, 2]$



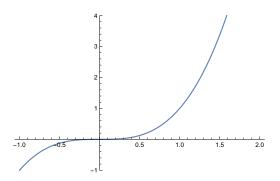


Consider a constraint $y = x^2, x \in [-1, 2]$



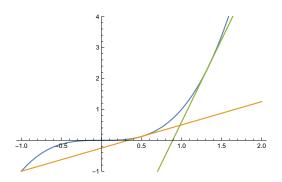


Another example: $y = x^3, x \in [-1, 2]$.



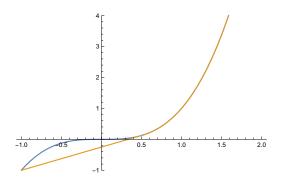


Another example: $y = x^3, x \in [-1, 2]$.





Another example: $y = x^3, x \in [-1, 2]$.



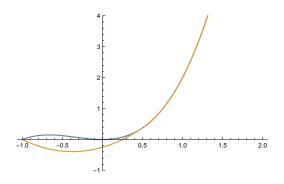


- we can handle simple functions
- · we can build more complicated functions by adding or multiplying



Example: Addition

- consider the sum, $y = x^2 + x^3, x \in [-1, 2]$.
- we can decompose the constraint as $w = x^2, z = x^3, y = w + z$





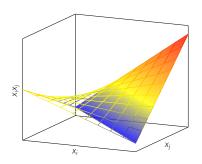
Example: Product

- consider the product $z = x^2y^3$.
- we can, again, decompose the constraint as z = uv, $u = x^2$, $v = y^3$
- what to do with z = uv?



Theorem [McCormick '76, Al-Khayyal and Falk '83]

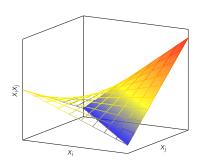
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \mid \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

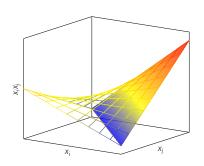
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

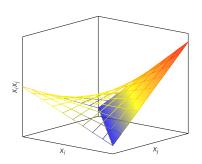
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

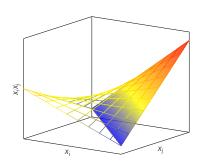
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

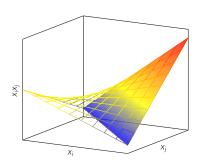
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

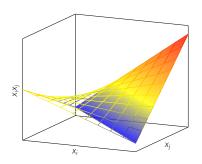
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

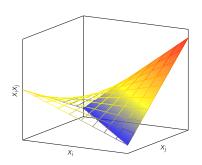
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \mid \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

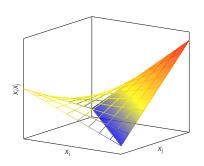
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

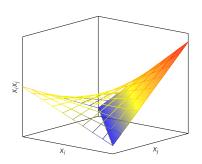
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

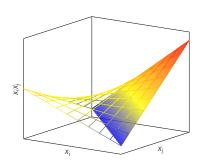
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

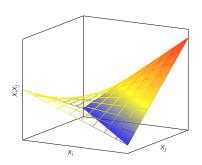
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

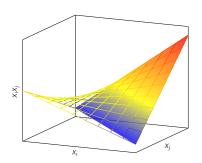
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

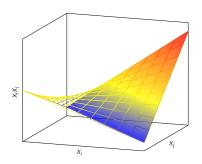
$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \mid \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Theorem [McCormick '76, Al-Khayyal and Falk '83]

$$conv(\{(X_{ij}, x_i, x_j) \mid X_{ij} = x_i x_j\}) = \left\{ (X_{ij}, x_i, x_j) \middle| \begin{array}{l} \ell_j x_i + \ell_i x_j - \ell_i \ell_j & \leq X_{ij} \\ u_j x_i + u_i x_j - u_i u_j & \leq X_{ij} \\ \ell_j x_i + u_i x_j - u_i \ell_j & \geq X_{ij} \\ u_j x_i + \ell_i x_j - \ell_i u_j & \geq X_{ij} \end{array} \right\}.$$





Comments

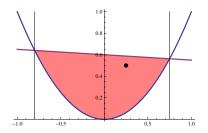
- · procedure can be generalized
- polyhedral relaxation depends on the bounds of variables
- · bounds are very important



Spatial Branch and bound

The variable bounds determine the convex relaxation, e.g.,

$$x^2 \le \ell^2 + \frac{u^2 - \ell^2}{u - \ell}(x - \ell) \quad \forall x \in [\ell, u].$$



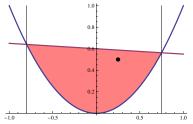


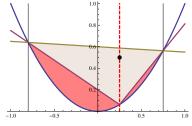
Spatial Branch and bound

The variable bounds determine the convex relaxation, e.g.,

$$x^2 \le \ell^2 + \frac{u^2 - \ell^2}{u - \ell} (x - \ell) \quad \forall x \in [\ell, u].$$

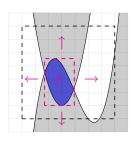
Thus, branching on a nonlinear variable in a nonconvex term allows for tighter relaxations:





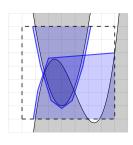


- min/max X_k s.t. $X \in \mathcal{R}, c^T X \leq z^*$
- · simple and effective



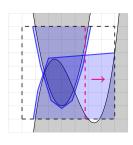


- min/max X_k s.t. $X \in \mathcal{R}, c^T X \leq Z^*$
- · simple and effective



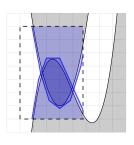


- · min/max X_k s.t. $X \in \mathcal{R}, c^T X \leq Z^*$
- · simple and effective



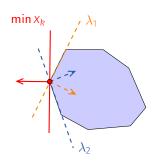


- · min/max X_k s.t. $X \in \mathcal{R}, c^T X \leq Z^*$
- · simple and effective





- · min/max X_k s.t. $X \in \mathcal{R}, c^T X \leq z^*$
- simple and effective
- · careful, might be expensive



Advanced implementation in SCIP

- fast propagation of duality certificates $x_k \ge \sum_i r_i x_i + \mu z^* + \lambda^T b$
- greedy ordering for faster LP warmstarts
- filtering of provably tight bounds
- 16% faster (24% on instances ≥ 100 seconds) and less time outs [Gleixner, Berthold, Mueller, Weltge, 2017]



Outline

Introduction: LP-based Branch and Bound

McCormick Relaxation

Spatial Branch and bound

Bound tightening

Improving over McCormick

Computational results

Conclusion



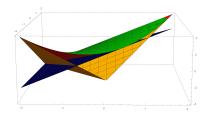
We can do better

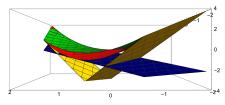
- · we know how to build a polyhedral relaxation
- quality depends on the bounds and we have a procedure for tightening bounds
- but, how good is the McCormick relaxation?



We can do better

- · we know how to build a polyhedral relaxation
- · quality depends on the bounds and we have a procedure for tightening bounds
- · but, how good is the McCormick relaxation?
- · Linderoth '04 suggests to branch along the diagonals of $[\ell_i, u_i] \times [\ell_j, u_j]$





green — graph of $X_{ij} = x_i x_j$ yellow — McCormick relaxation

red — $conv_P(x_ix_i)$: convex envelope over $P \sim significantly tighter$



Idea

Try to find a valid $P \subsetneq [\ell_i, u_i] \times [\ell_j, u_j]$ and study $conv\{(x_i, x_j, X_{ij}) \mid X_{ij} = x_i x_j, (x_i, x_j) \in P\}$

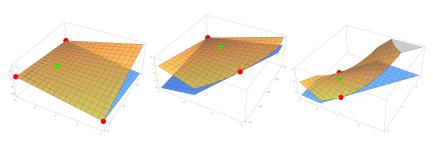
Two ingredients:

- given P, convex envelope of $X_{ij} = x_i x_j$ over P?
- how to find P?



Convex envelope: Locatelli '16

• Three cases when computing tangent inequality of $conv_P(x_ix_j)$ at reference point:



 select three vertices of P

- select a vertex and a point on a facet of P
- select two points on two facets of P

Easy to compute tangent inequalities of $conv_P(x_ix_j)$ if vertices and facets of P are explicitly given.



Before computing P...

- · Some problems readily give us P, e.g., pointpack08 from MINLPLib2
- 56 bilinear terms
- 8 linear constraints $x_i \leq x_i$
- · How does exploiting P perform?



Before computing P...

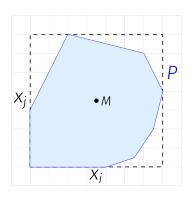
- · Some problems readily give us P, e.g., pointpack08 from MINLPLib2
- 56 bilinear terms
- 8 linear constraints $x_i \leq x_i$
- How does exploiting P perform?

	SCIP de	efault	SCIF	SCIP ⁺		
perm.	nodes	time	nodes	time		
1	1844k	1279	71k	64		
2	1780k	1186	69k	80		
3	1486k	943	87k	72		
4	744k	483	193k	168		
5	1372k	1054	89k	93		
6	370k	277	45k	57		
7	2285k	1522	60k	66		
8	168k	136	53k	56		
9	1771k	1174	45k	55		
10	190k	145	48k	50		



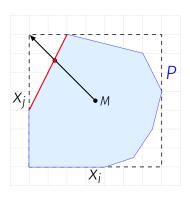
on average 10x less time and 15x less nodes

- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := \left(\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}\right) \in P$
- facets intersecting segment joining M with each corner



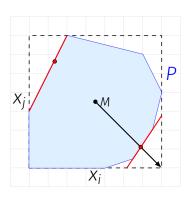


- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := (\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}) \in P$
- facets intersecting segment joining M with each corner



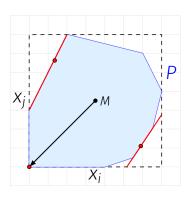


- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := (\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}) \in P$
- facets intersecting segment joining M with each corner



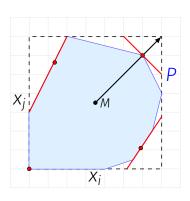


- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := (\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}) \in P$
- facets intersecting segment joining M with each corner





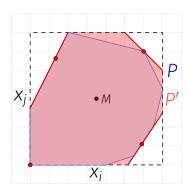
- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := (\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}) \in P$
- facets intersecting segment joining M with each corner





- $P = proj_{(x_i, x_j)}(\mathcal{R})$ is "best" possible, but impractical
- after OBBT, $M := \left(\frac{u_i + \ell_i}{2}, \frac{u_j + \ell_j}{2}\right) \in P$
- facets intersecting segment joining M with each corner

Optimize along directions from M to each corner



After 4 LPs

 $P'\supseteq P=proj_{(x_i,x_j)}(\mathcal{R})$ described by at most

- 4 nontrivial inequalities
- 4 axis-parallel inequalities



Details

E.g. from
$$M=(M_i,M_j):=(\frac{\ell_i+u_i}{2},\frac{\ell_j+u_j}{2})$$
 to (u_i,u_j) via the LP

$$\max \left\{ \alpha : \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} M_i \\ M_j \end{pmatrix} + \alpha \begin{pmatrix} u_i - M_i \\ u_j - M_j \end{pmatrix}, x \in \mathcal{R}, \alpha \in [0, 1] \right\}$$



Details

E.g. from $M = (M_i, M_j) := (\frac{\ell_i + u_i}{2}, \frac{\ell_j + u_j}{2})$ to (u_i, u_j) via the LP

$$\max \left\{ \alpha : \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} M_i \\ M_j \end{pmatrix} + \alpha \begin{pmatrix} u_i - M_i \\ u_j - M_j \end{pmatrix}, x \in \mathcal{R}, \alpha \in [0, 1] \right\}$$
$$\max \left\{ x_i : \frac{x_i - M_i}{u_i - M_i} = \frac{x_j - M_j}{u_j - M_j}, x \in \mathcal{R} \right\}$$

Essentially OBBT with one additional constraint



Connection to OBBT

Can use all OBBT tricks from Gleixner, Berthold, Mueller, Weltge '17:

LP filtering

- E.g., if $\exists (x^*, X^*) \in \mathcal{R} : X_i^* = \ell_i \land X_i^* = u_j \Rightarrow \text{filter } M \to (\ell_i, u_j)$
- · already applicable during standard OBBT

LP ordering

- exploit simplex warmstart
- e.g., after $\max\{x_i \mid (x,X) \in \mathcal{R}\}$ solve all $M \to (u_i,\cdot)$ (dual feasible LP basis)

Updating facets with new incumbent information

- new incumbent solution tightens P if $\lambda \neq 0$
- · similar to Lagrangian Variable Bounds from Gleixner and Weltge '13



Outline

Introduction: LP-based Branch and Bound

McCormick Relaxation

Spatial Branch and bound

Bound tightening

Improving over McCormicl

Computational results

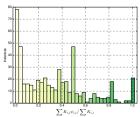
Conclusion



Computational results: affected instances

- 1367 instances of MINLPLib2 after presolving
- filter instances
 - with no bilinear terms
 - solved before finding facets
 - · with no facets found
- → 422 potentially affected instances

Frequency of nontrivial facets:



- $K_{i,j} := |\{k \in [m] \mid (Q_k)_{i,j} \neq 0\}|$ occurrences of a bilinear term $x_i x_j$
- $\psi_{i,i} := 1$ (found facet for $x_i x_i$)
- final measure: $\sum_{i,j} K_{i,j} \psi_{i,j} / \sum_{i,j} K_{i,j} \in [0,1]$

Computational results: gap closed

SCIP settings

- propagating/obbt/itlimitfactor = -1
- provide best known solution
- limits/restart = 0 and limits/totalnodes = 1
- separation/emphasis/aggressive = TRUE



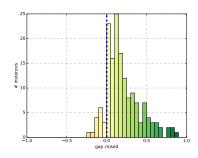
Computational results: gap closed

SCIP settings

- propagating/obbt/itlimitfactor = -1
- · provide best known solution
- limits/restart = 0 and limits/totalnodes = 1
- separation/emphasis/aggressive = TRUE

Root node results

	instances	gap closed
all	422	16%
>1% change	171	40%
>1% better	155	44%
>1% worse	16	-15%





Computational results: performance

Setup¹

- · limits/time = 1800s
- · limits/gap = 1e-4

- 6 different permutations
- · one instance per cluster node

		default	no propagation		no prop. + no sepa.	
	n	# solved	# solved	time	# solved	time
ALL	422	213	205	1.07	199	1.12
[1,tlim] [10,tlim] [100,tlim]	143 96 45	138 91 40	131 84 34	1.17 1.26 1.36	128 81 34	1.30 1.52 1.76

Results

- +14 more solved instances
- 12% speed-up on ALL
- 76% speed-up on [100,tlim] Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz

Outline

Introduction: LP-based Branch and Bound

McCormick Relaxation

Spatial Branch and bound

Bound tightening

Improving over McCormic

Computational results

Conclusion



Conclusion

Summary

- · very brief introduction to LP-based branch and bound
- tighter convexification over projections of the feasible region



Conclusion

Summary

- very brief introduction to LP-based branch and bound
- · tighter convexification over projections of the feasible region

Impact

- 40% root gap closed on affected instances
- · +14 more instances solved; 12% speedup on all and 76% on hard instances
- more drastic speedups on specific instances



Conclusion

Summary

- very brief introduction to LP-based branch and bound
- · tighter convexification over projections of the feasible region

Impact

- 40% root gap closed on affected instances
- +14 more instances solved; 12% speedup on all and 76% on hard instances
- more drastic speedups on specific instances

Thank you very much for your attention!



MINLP and Stronger Relaxation of Bilinear Terms

Benjamin Müller, Ambros Gleixner, and Felipe Serrano

Zuse Institute Berlin · serrano@zib.de
SCIP Optimization Suite · http://scip.zib.de

SCIP Workshop

Aachen · March 7, 2018





