A Strengthened SDP Relaxation for an Extended Trust Region Subproblem

Sam Burer

University of Iowa

MIP Workshop (feat. DANniversary)
DIMACS, Rutgers University
May 26, 2022
Joint Work With

Anders Eltved
Outline

1 Introduction
2 A New Class of Cuts
3 A Bit of Computation
4 Applications
5 Conclusions
Non-Convex Mixed-Integer Quadratic Programming

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad x^T H_i x + 2 g_i^T x + f_i \leq 0 \quad \forall \ i \in I \\
& \quad l \leq x \leq u \quad x_J \in \mathbb{Z}^{|J|}
\end{align*}
\]
**Trust Region Subproblem (TRS)**

\[
\min_{x \in \mathbb{R}^n} \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} \quad \|x\| \leq 1
\]
Trust Region Subproblem (TRS)

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad \|x\| \leq 1
\end{align*}
\]

**Fact:** TRS is polynomial-time solvable. In particular, its Shor SDP relaxation is exact because

\[
\begin{align*}
\text{CH} := \text{conv} \left\{ \left( x, x^T x \right) : \|x\| \leq 1 \right\} = \\
\left\{ \left( x, X \right) : \text{trace}(X) \leq 1, (1 x^T x X) \succeq 0 \right\}
\end{align*}
\]
Trust Region Subproblem (TRS)

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad \|x\| \leq 1
\end{align*}
\]

**Fact:** TRS is polynomial-time solvable. In particular, its Shor SDP relaxation is exact because

\[
CH := \text{conv} \left\{ (x, xx^T) : \|x\| \leq 1 \right\} = \left\{ (x, X) : \text{trace}(X) \leq 1, \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \right\}
\]
Extended Trust Region Subproblem (ETRS)

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad \|x\| \leq 1 \\
& \quad \text{(additional quadratic constraints)}
\end{align*}
\]
To describe our main result suppose that for $0 \leq i \leq p$, $g_i(x)$ is a quadratic, over $x \in \mathbb{R}^n$. We consider the problem

$$\min \{ g_0(x) : g_i(x) \leq 0, \ 1 \leq i \leq p \}$$

and prove the following.

**Theorem 1.2.** For each fixed integer $p \geq 1$ there is an algorithm with the following properties. Given a problem of the form (1.2) where at least one of the $g_i(x)$ with $i \geq 1$ is strictly convex, and $0 < \epsilon < 1$, the algorithm either

1. proves that problem (1.2) is infeasible,

or

2. computes an $\epsilon$-feasible vector $\hat{x}$ such that there exists no feasible $x$ with $g_0(x) < g(\hat{x}) - \epsilon$.

Under Assumption 1 the algorithm runs in polynomial time.
Sakaue et al (2016):

“Unfortunately, however, Bienstock’s polynomial-time algorithm does not appear to be very practical, because [Barvinok’s] polynomial-time feasibility algorithm [which Bienstock uses] looks difficult to implement.”
Does ETRS Have an Exact SDP Relaxation?

\[ \text{CH} := \text{conv} \left\{ (x, xx^T) : \|x\| \leq 1 \right\} \text{ (add’l quad constraints)} \]
Does ETRS Have an Exact SDP Relaxation?

\[ \text{CH} := \text{conv} \left\{ (x, xx^T) : \|x\| \leq 1 \right\} \]

(Prefer to work in the space \((x, xx^T)\))
The Swiss-Cheese Result

The Swiss-Cheese Result (Yang-B-Anstreicher 2018)

Starting from the unit ball, remove finitely many ellipsoids and half-spaces. If these deletions are "nonintersecting," then $\mathbb{CH}$ has a polynomially sized representation using SDP, RLT, and SOCRLT constraints. (Will define these constraint types later)
The Swiss-Cheese Result

**Theorem (Yang-B-Anstreicher 2018)**

Starting from the unit ball, remove finitely many ellipsoids and half-spaces. If these deletions are “nonintersecting,” then $CH$ has a polynomially sized representation using SDP, RLT, and SOCRLT constraints.
The Swiss-Cheese Result

**Theorem (Yang-B-Anstreicher 2018)**

Starting from the unit ball, remove finitely many ellipsoids and half-spaces. If these deletions are “nonintersecting,” then CH has a polynomially sized representation using SDP, RLT, and SOCRLT constraints.

(Will define these constraint types later)
On the Other Hand: An Elusive Result

\[ \text{CH} := \text{conv} \left\{ (x, xx^T) : \begin{array}{l}
\|x\| \leq 1 \\
\|Hx - c\| \leq \rho
\end{array} \right\} \]

- This is the CDT (Celis-Dennis-Tapia) problem
- Also called TTRS for “two TRS”
- In this case, CH does not have a known representation (even when either \( H = I \) or \( c = 0 \))
Our ETRS

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad r \leq \|x\| \leq R \\
& \quad \|x\| \leq b^T x \\
& \quad l \leq s^T x \leq u
\end{align*}
\]
Our ETRS

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T H_0 x + 2 g_0^T x + f_0 \\
\text{s. t.} & \quad r \leq \|x\| \leq R \\
& \quad \|x\| \leq b^T x \\
& \quad l \leq s^T x \leq u
\end{align*}
\]

(∃ extensions to \(\|Hx - c\| \leq b^T x - a\))
Motivation: Optimal Power Flow

- Chen et al. (2017) studied an important substructure in the ACOPF (alternating current optimal power flow) problem:

\[ J_C := \text{conv} \left\{ \begin{pmatrix} W_{11} \\ W_{22} \\ W_{12} \\ T_{12} \end{pmatrix} \in \mathbb{R}^4 : \begin{aligned} & L_{jj} \leq W_{jj} \leq U_{jj} \quad \forall \ j = 1, 2 \\ & L_{12}W_{12} \leq T_{12} \leq U_{12}W_{12} \\ & W_{12} \geq 0 \\ & W_{11}W_{22} = W_{12}^2 + T_{12}^2 \end{aligned} \right\} \]

- Data \( L = (L_{11}, L_{22}, L_{12}) \) and \( U = (U_{11}, U_{22}, U_{12}) \) satisfy
  - \( L \leq U \) and \( L_{jj} \geq 0 \) for \( j = 1, 2 \)
  - \( L_{22} > 0 \) and \( U_{12} > L_{12} \)
Motivation: Optimal Power Flow

**Theorem (Chen et al. 2017)**

$\mathcal{J}_C$ is obtained by relaxing

\[ W_{11}W_{22} = W_{12}^2 + T_{12}^2 \quad \rightarrow \quad W_{11}W_{22} \geq W_{12}^2 + T_{12}^2 \]

and enforcing linear cuts

\[
\begin{align*}
\pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} & \geq U_{22}W_{11} + U_{11}W_{22} - U_{11}U_{22} \\
\pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} & \geq L_{22}W_{11} + L_{11}W_{22} - L_{11}L_{22}
\end{align*}
\]

where $\pi_0, \pi_1, \pi_2, \pi_3, \pi_4$ are constants depending on $(L, U)$
Motivation: Optimal Power Flow

\[
\pi_0 := -\sqrt{L_{11}L_{22}U_{11}U_{22}}
\]

\[
\pi_1 := -\sqrt{L_{22}U_{22}}
\]

\[
\pi_2 := -\sqrt{L_{11}U_{11}}
\]

\[
\pi_3 := \left(\sqrt{L_{11}} + \sqrt{U_{11}}\right) \left(\sqrt{L_{22}} + \sqrt{U_{22}}\right) \frac{1 - f(L_{12})f(U_{12})}{1 + f(L_{12})f(U_{12})}
\]

\[
\pi_4 := \left(\sqrt{L_{11}} + \sqrt{U_{11}}\right) \left(\sqrt{L_{22}} + \sqrt{U_{22}}\right) \frac{f(L_{12}) + f(U_{12})}{1 + f(L_{12})f(U_{12})}
\]

where \( f(x) := (\sqrt{1 + x^2} - 1)/x \) when \( x > 0 \) and \( f(0) := 0 \)
Motivation: Optimal Power Flow

**Proposition (Eltved-B 2020)**

\[ J_C \] is a simple projection of

\[
CH := \text{conv} \left\{ (x, xx^T) : \sqrt{L_{11}} \leq \|(x_1, x_2)\| \leq \sqrt{U_{11}} \right\}
\]

where \( b_1 \) and \( b_2 \) uniquely solve

\[
\begin{pmatrix} 1 & L_{12} \\ 1 & U_{12} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \sqrt{1 + \frac{L_{12}^2}{U_{12}}} \\ \sqrt{1 + \frac{U_{12}^2}{U_{12}}} \end{pmatrix}
\]
Outline

1. Introduction
2. A New Class of Cuts
3. A Bit of Computation
4. Applications
5. Conclusions
Our ETRS

\[ \text{CH} := \text{conv} \left\{ (x, xx^T) : \begin{array}{l} r \leq \|x\| \leq R \\ \|x\| \leq b^T x \\ l \leq s^T x \leq u \end{array} \right\} \]
Valid Constraints: The Shor Relaxation

\[ r^2 \leq \text{trace}(X) \leq R^2 \]
\[ \text{trace}(X) \leq bb^T \cdot X, \quad b^T x \geq 0 \]
\[ l \leq s^T x \leq u \]
\[ \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \]
Valid Constraints: RLT

\[(u - s^T x)(s^T x - l) \geq 0 \quad \implies \quad (l + u)s^T x \geq s^T X s + lu\]
Valid Constraints: SOCRLT

\[
\| (u - s^T x)x \| \leq R(u - s^T x)
\]
Valid Constraints: SOCRLT

$$\|(u - s^T x)x\| \leq R(u - s^T x) \quad \rightarrow \quad \|ux - Xs\| \leq R(u - s^T x)$$

Note: The SOCRLT idea first appeared in Sturm and Zhang (2003)
Valid Constraints: SOCRLT

\[ \| (u - s^T x) x \| \leq R(u - s^T x) \quad \rightarrow \quad \| ux - Xs \| \leq R(u - s^T x) \]

\[ \| (s^T x - l) x \| \leq R(s^T x - l) \quad \rightarrow \quad \| Xs - lx \| \leq R(s^T x - l) \]
Valid Constraints: SOCRLT

\[
\| (u - s^T x)x \| \leq R(u - s^T x) \quad \rightarrow \quad \| ux - Xs \| \leq R(u - s^T x) \\
\| (s^T x - l)x \| \leq R(s^T x - l) \quad \rightarrow \quad \| Xs - lx \| \leq R(s^T x - l)
\]

Note: The SOCRLT idea first appeared in Sturm and Zhang (2003)
Valid Constraints: Kronecker SOC

Fact: \( \| v \| \leq v_0 \) iff \( \begin{pmatrix} v_0 & v^T \\ v & v_0I \end{pmatrix} \succeq 0 \)

Fact: The Kronecker product of PSD matrices is PSD

Note: The Kronecker SOC idea first appeared in Anstreicher (2017)
Valid Constraints: Kronecker SOC

Fact: \( \|v\| \leq v_0 \) iff \( \begin{pmatrix} v_0 & v^T \\ v & v_0I \end{pmatrix} \succeq 0 \)

Fact: The Kronecker product of PSD matrices is PSD

\[
\begin{pmatrix} R & x^T \\ x & RI \end{pmatrix} \otimes \begin{pmatrix} b^T x & x^T \\ x & (b^T x)I \end{pmatrix} \succeq 0
\]
Valid Constraints: Kronecker SOC

**Fact:** $\|v\| \le v_0$ iff $\begin{pmatrix} v_0 & v^T \\ v & v_0I \end{pmatrix} \succeq 0$

**Fact:** The Kronecker product of PSD matrices is PSD

$$\begin{pmatrix} R & x^T \\ x & RI \end{pmatrix} \otimes \begin{pmatrix} b^Tx & x^T \\ x & (b^Tx)I \end{pmatrix} \succeq 0$$

**Note:** The Kronecker SOC idea first appeared in Anstreicher (2017)
The Kitchen Sink

Define the *KitchenSink* to be the relaxation that includes:
- Basic SDP relaxation
- RLT cut
- SOCRLT constraints
- Kronecker SOC constraint

**Obs:** *KitchenSink* does not capture CH even for $n = 2$
New Class of Linear Cuts

\[ r \leq \|x\| \leq R \]

\[ \|x\| \leq b^T x \]

\[ \lambda \in \mathcal{S}_x \quad \mathcal{S}_x \leq u \]
New Class of Linear Cuts

\[ r \leq \|x\| \quad \|x\| \leq R \quad \|x\| \leq b^T x \quad \lambda_\in \overline{S}_x \quad \overline{S}_x \in \mathcal{U} \]

\[ (\overline{S}_x - \lambda)(\overline{R}_x) + (u - \overline{S}_x)(\overline{b^T x}) \in \mathcal{SOC} \]
New Class of Linear Cuts

\[ \begin{align*}
    r & \leq \|x\| \\
    \|x\| & \leq R \\
    \|x\| & \leq b^T x \\
    \lambda & \leq s^T x \\
    s^T x & \leq u
\end{align*} \]

\[ \begin{align*}
    (\|x\| - r)(R - \|x\|) & \geq 0 \\
    & \iff \\
    \left( \frac{r + R}{1 + rR \|x\|^2} \right) x & \in \text{SOC} \\
    (s^T - \lambda)(R_x) + (u - s^T x)(b^T x) & \in \text{SOC}
\end{align*} \]
New Class of Linear Cuts

\[ r \leq \|x\| \leq R \quad \|x\| \leq \frac{b^T x}{\bar{s}_x} \quad \lambda \leq \bar{s}_x \quad \bar{s}_x \leq u \]

\[ \left(\|x\| - r\right)\left(R - \|x\|\right) \geq 0 \]

\[ \left(\frac{r + R}{1 + rR \|x\|^{-2}}\right) x \in \text{SOC} \]

\[ \left(\bar{s}_x - \lambda\right)\left(R\right) + \left(u - \bar{s}_x\right)\left(\frac{b^T x}{\bar{s}_x}\right) \in \text{SOC} \]

\[ \left(\bar{s}_x - \lambda\right)R\left(r + R\right) + \left(u - \bar{s}_x\right)b^T x\left(r + R\right) \geq \left(u - \lambda\right)\left(x^T x + rR\right) \]
New Class of Linear Cuts

- Can replace $s^T x - l \geq 0$ and $u - s^T x \geq 0$ by quadratic $q(x) \geq 0$ and linear $l(x) \geq 0$, resp.
- Moreover, since the dual feasible set of KitchenSink encodes nonnegative quadratics, we can use KitchenSink to “bootstrap” $q(x)$ and $l(x)$
- This leads to a separation routine for our cuts, which chooses the best $q(x)$ and $l(x)$ automatically
New Class of Linear Cuts

Theorem (Eltved-B 2020)

Bootstrapping from KitchenSink, our class of linear cuts can be $\epsilon$-separated in polynomial time, and the cuts strengthen KitchenSink
New Class of Linear Cuts

**Theorem (Eltved-B 2020)**

*Bootstrapping from KitchenSink, our class of linear cuts can be $\epsilon$-separated in polynomial time, and the cuts strengthen KitchenSink*

(But we still don’t have a full representation of CH)
Outline

1. Introduction
2. A New Class of Cuts
3. A Bit of Computation
4. Applications
5. Conclusions
A Bit of Computation

- For each $n = 2, \ldots, 10$, we generated 15,000 random feas instances (with $r = 0$)
- “Solved” means we get a rank-1 optimal solution
- We first solve KitchenSink and then, if necessary, add our cuts
- Note: For $n = 10$, solving KitchenSink took 50 seconds, and generating a single cut took approximately 64 seconds
A Bit of Computation

<table>
<thead>
<tr>
<th>$n$</th>
<th># unsolved by KitchenSink</th>
<th># solved by new cuts</th>
<th>avg cuts added</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>15</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>
Outline

1. Introduction
2. A New Class of Cuts
3. A Bit of Computation
4. Applications
5. Conclusions
Application: Optimal Power Flow

Proposition (Eltved-B 2020)

The ACOPF cuts

\[ \pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} \geq U_{22} W_{11} + U_{11} W_{22} - U_{11} U_{22} \]
\[ \pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} \geq L_{22} W_{11} + L_{11} W_{22} - L_{11} L_{22} \]

are members of our class
Application: The Nonnegative Ball

\[ \text{CH} := \text{conv} \{(x, xx^T) : \|x\| \leq 1, x \geq 0\} \]
Application: The Nonnegative Ball

\[ \text{CH} := \text{conv} \left\{ (x, xx^T) : \|x\| \leq 1, x \geq 0 \right\} \]

\[ = \text{conv} \left\{ (x, xx^T) : \|x\| \leq 1, x \geq 0, 0 \leq s^T x \leq 1 \quad \forall \|s\| = 1 \right\} \]
Application: The Nonnegative Ball

\[ \text{CH} := \text{conv} \left\{ (x, xxx^T) : \|x\| \leq 1, x \geq 0 \right\} \]

\[ = \text{conv} \left\{ (x, xxx^T) : \|x\| \leq 1, x \geq 0, 0 \leq s^T x \leq 1 \quad \forall \|s\| = 1 \right\} \]

\[ \subseteq \text{conv} \left\{ (x, xxx^T) : \|x\| \leq 1, \|x\| \leq e^T x, 0 \leq s^T x \leq 1 \quad \forall \|s\| = 1 \right\} \]
Application: The Nonnegative Ball

\[ \text{CH} := \text{conv}\left\{ (x, xx^T) : \|x\| \leq 1, x \geq 0 \right\} \]

\[ = \text{conv}\left\{ (x, xx^T) : \|x\| \leq 1, x \geq 0, \quad 0 \leq s^T x \leq 1 \quad \forall \|s\| = 1 \right\} \]

\[ \subseteq \text{conv}\left\{ (x, xx^T) : \|x\| \leq 1, \|x\| \leq e^T x, \quad 0 \leq s^T x \leq 1 \quad \forall \|s\| = 1 \right\} \]

(Equality holds for \( n = 2 \))
Application: The Nonnegative Ball

Theorem (Eltved-B 2020)

Let \((I, J)\) be a partition of \(\{1, \ldots, n\}\), and define the domain

\[
D_{IJ} := \left\{ (x, X) : \begin{array}{c}
[Xe - x]_I \geq 0 \\
[Xe - x]_J \leq 0
\end{array} \right\}.
\]

Then the following SOC constraints are locally valid on \(D_{IJ}\):

\[
\| [Xe - x]_J \| \leq 1 - \text{trace}(X)
\]

\[
\| [Xe - x]_I \| \leq e^T x - \text{trace}(X)
\]
Application: The Nonnegative Ball

Conjecture

*For* $n = 2$, *KitchenSink* plus these locally valid cuts capture CH
Outline

1. Introduction
2. A New Class of Cuts
3. A Bit of Computation
4. Applications
5. Conclusions
Conclusions

THANKS TO THE MIP ORGANIZERS!