# 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

# Non-Convex Mixed-Integer Quadratic Programming

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\ \text{s. t.} \quad & x^T H_i x + 2\, g_i^T x + f_i \le 0 \;\; \forall\, i \in I \\ & l \le x \le u \qquad x_J \in \mathbb{Z}^{|J|} \end{aligned}$$

# Trust Region Subproblem (TRS)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\ \text{s. t.} \quad & \|x\| \leq 1 \end{aligned}$$

# Trust Region Subproblem (TRS)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\ \text{s. t.} \quad & \|x\| \le 1 \end{aligned}$$

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

# Trust Region Subproblem (TRS)

$$\min_{x \in \mathbb{R}^n} \quad x^T H_0 x + 2\, g_0^T x + f_0$$
$$\text{s. t.} \quad \|x\| \le 1$$

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

$$\begin{aligned} \text{CH} &:= \text{conv}\left\{(x, xx^T) : \|x\| \le 1\right\} \\ &= \left\{(x, X) : \text{trace}(X) \le 1, \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0\right\} \end{aligned}$$

# Extended Trust Region Subproblem (ETRS)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\ \text{s. t.} \quad & \|x\| \le 1 \\ & \text{(additional quadratic constraints)} \end{aligned}$$

# Dan's ETRS Result (2016)

To describe our main result suppose that for $0 \le i \le p$, $g_i(x)$ is a quadratic, over $x \in \mathbb{R}^n$. We consider the problem

$$\min \{ g_0(x) \; : \; g_i(x) \le 0, \; 1 \le i \le p \} \tag{1.2}$$

and prove the following.

THEOREM 1.2. *For each fixed integer $p \ge 1$ 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 $i \ge 1$ 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.*

# Dan's ETRS Result (2016)

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?

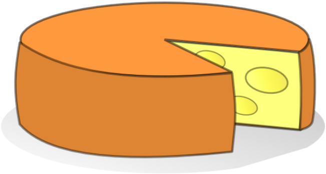
$$\text{CH} := \text{conv} \left\{ (x, xx^T) : \begin{array}{l} \|x\| \le 1 \\ \text{(add'l quad constraints)} \end{array} \right\}$$

# Does ETRS Have an Exact SDP Relaxation?

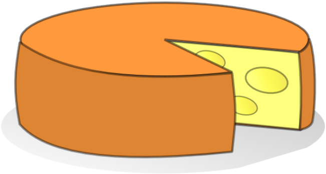
$$\text{CH} := \text{conv}\left\{ (x, xx^T) : \begin{array}{l} \|x\| \leq 1 \\ \text{(add'l quad constraints)} \end{array} \right\}$$

(Prefer to work in the space $(x, xx^T)$)

# The Swiss-Cheese Result



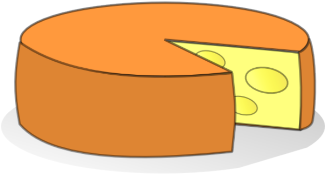
# The Swiss-Cheese Result



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

*Starting from the unit ball, remove finitely many ellipsoids and half-spaces. If these deletions are "nonintersecting," then* $\mathrm{CH}$ *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* $\mathrm{CH}$ *has a polynomially sized representation using SDP, RLT, and SOCRLT constraints.*

(Will define these constraint types later)

# On the Other Hand: An Elusive Result

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

- This is the CDT (Celis-Dennis-Tapia) problem
- Also called TTRS for "two TRS"
- In this case, $\text{CH}$ does *not* have a known representation (even when either $H = I$ or $c = 0$)

# Our ETRS

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\ \text{s. t.} \quad & r \le \|x\| \le R \\ & \|x\| \le b^T x \\ & l \le s^T x \le u \end{aligned}$$

# Our ETRS

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & x^T H_0 x + 2\, g_0^T x + f_0 \\
\text{s.t.} \quad & r \le \|x\| \le R \\
& \|x\| \le b^T x \\
& l \le s^T x \le u
\end{aligned}
$$

($\exists$ extensions to $\|Hx - c\| \le b^T x - a$)

# Motivation: Optimal Power Flow

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

$$\mathcal{J}_C := \text{conv}\left\{ \begin{pmatrix} W_{11} \\ W_{22} \\ W_{12} \\ T_{12} \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{l} L_{jj} \le W_{jj} \le U_{jj} \quad \forall\, j = 1,2 \\ L_{12} W_{12} \le T_{12} \le U_{12} W_{12} \\ W_{12} \ge 0 \\ W_{11} W_{22} = W_{12}^2 + T_{12}^2 \end{array} \right\}$$

- Data $L = (L_{11}, L_{22}, L_{12})$ and $U = (U_{11}, U_{22}, U_{12})$ satisfy
  - $L \le U$ and $L_{jj} \ge 0$ 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*

$$W_{11}W_{22} = W_{12}^2 + T_{12}^2 \quad \longrightarrow \quad W_{11}W_{22} \geq W_{12}^2 + T_{12}^2$$

*and enforcing linear cuts*

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

*where $\pi_0, \pi_1, \pi_2, \pi_3, \pi_4$ are constants depending on $(L, U)$*

# Motivation: Optimal Power Flow

$$
\begin{aligned}
\pi_0 &:= -\sqrt{L_{11}L_{22}U_{11}U_{22}} \\
\pi_1 &:= -\sqrt{L_{22}U_{22}} \\
\pi_2 &:= -\sqrt{L_{11}U_{11}} \\
\pi_3 &:= \left(\sqrt{L_{11}} + \sqrt{U_{11}}\right)\left(\sqrt{L_{22}} + \sqrt{U_{22}}\right)\frac{1 - f(L_{12})f(U_{12})}{1 + f(L_{12})f(U_{12})} \\
\pi_4 &:= \left(\sqrt{L_{11}} + \sqrt{U_{11}}\right)\left(\sqrt{L_{22}} + \sqrt{U_{22}}\right)\frac{f(L_{12}) + f(U_{12})}{1 + f(L_{12})f(U_{12})}
\end{aligned}
$$

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

# Motivation: Optimal Power Flow

## Proposition (Eltved-B 2020)

*$\mathcal{J}_C$ is a simple projection of*

$$\text{CH} := \text{conv}\left\{ (x, xx^T) : \begin{array}{c} \sqrt{L_{11}} \le \|(x_1, x_2)\| \le \sqrt{U_{11}} \\ \|(x_1, x_2)\| \le b_1 x_1 + b_2 x_2 \\ \sqrt{L_{22}} \le x_3 \le \sqrt{U_{22}} \end{array} \right\}$$

*where $b_1$ and $b_2$ uniquely solve*

$$\begin{pmatrix} 1 & L_{12} \\ 1 & U_{12} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \sqrt{1 + L_{12}^2} \\ \sqrt{1 + U_{12}^2} \end{pmatrix}$$

# Outline

# Our ETRS

$$\text{CH} := \text{conv}\left\{ (x, xx^T) : \begin{array}{r} r \le \|x\| \le R \\ \|x\| \le b^T x \\ l \le s^T x \le u \end{array} \right\}$$

# Valid Constraints: The Shor Relaxation

$$
\begin{aligned}
&r^2 \le \operatorname{trace}(X) \le R^2 \\
&\operatorname{trace}(X) \le bb^T \bullet X, \quad b^T x \ge 0 \\
&l \le s^T x \le u \\
&\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0
\end{aligned}
$$

# Valid Constraints: RLT

$$(u - s^T x)(s^T x - l) \geq 0 \quad \longrightarrow \quad (l + u)s^T x \geq s^T X s + lu$$

# Valid Constraints: SOCRLT

$$\|(u - s^T x)x\| \leq R(u - s^T x)$$

# Valid Constraints: SOCRLT

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

# Valid Constraints: SOCRLT

$$
\begin{aligned}
\|(u - s^T x)x\| \le R(u - s^T x) \quad &\longrightarrow \quad \|ux - Xs\| \le R(u - s^T x) \\
\|(s^T x - l)x\| \le R(s^T x - l) \quad &\longrightarrow \quad \|Xs - lx\| \le R(s^T x - l)
\end{aligned}
$$

# Valid Constraints: SOCRLT

$$\|(u - s^T x)x\| \le R(u - s^T x) \quad \longrightarrow \quad \|ux - Xs\| \le R(u - s^T x)$$
$$\|(s^T x - l)x\| \le R(s^T x - l) \quad \longrightarrow \quad \|Xs - lx\| \le R(s^T x - l)$$

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

# Valid Constraints: Kronecker SOC

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

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

# Valid Constraints: Kronecker SOC

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

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

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

# Valid Constraints: Kronecker SOC

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

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

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

**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 $\mathrm{CH}$ even for $n = 2$

# New Class of Linear Cuts

$r \le \|x\|$ $\quad \|x\| \le R$ $\quad \|x\| \le b^T x$ $\quad \ell \le s^T x$ $\quad s^T x \le u$

# New Class of Linear Cuts

$r \le \|x\|$  $\|x\| \le R$  $\|x\| \le b^T x$  $\ell \le s^T x$  $s^T x \le u$

$$(s^T x - \ell)\begin{pmatrix} R \\ x \end{pmatrix} + (u - s^T x)\begin{pmatrix} b^T x \\ x \end{pmatrix} \in SOC$$

# New Class of Linear Cuts

$$r \le \|x\| \qquad \|x\| \le R \qquad \|x\| \le b^T x \qquad \ell \le s^T x \qquad s^T x \le u$$

$$(\|x\| - r)(R - \|x\|) \ge 0$$

$$\Updownarrow$$

$$\begin{pmatrix} r + R \\ (1 + rR\|x\|^{-2})\, x \end{pmatrix} \in SOC$$

$$(s^T x - \ell)\begin{pmatrix} R \\ x \end{pmatrix} + (u - s^T x)\begin{pmatrix} b^T x \\ x \end{pmatrix} \in SOC$$

# New Class of Linear Cuts

$$r \le \|x\| \qquad \|x\| \le R \qquad \|x\| \le b^T x \qquad \ell \le s^T x \qquad s^T x \le u$$

$$(\|x\| - r)(R - \|x\|) \ge 0$$

$$\Updownarrow$$

$$\begin{pmatrix} r + R \\ (1 + rR\|x\|^{-2})\, x \end{pmatrix} \in SOC$$

$$(s^T x - \ell)\begin{pmatrix} R \\ x \end{pmatrix} + (u - s^T x)\begin{pmatrix} b^T x \\ x \end{pmatrix} \in SOC$$

$$(s^T x - \ell) R (r + R) + (u - s^T x) b^T x (r + R) \ge (u - \ell)(x^T x + rR)$$

# 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 $\mathrm{CH}$)

# Outline

# 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

| $n$ | # unsolved by KitchenSink | # solved by new cuts | avg cuts added |
|---|---|---|---|
| 2 | 15 | 15 | 2 |
| 3 | 50 | 30 | 2 |
| 4 | 36 | 28 | 2 |
| 5 | 29 | 27 | 3 |
| 6 | 15 | 12 | 3 |
| 7 | 13 | 11 | 2 |
| 8 | 12 | 12 | 2 |
| 9 | 6 | 5 | 1 |
| 10 | 6 | 5 | 3 |

# Outline

# Application: Optimal Power Flow

## Proposition (Eltved-B 2020)

*The ACOPF cuts*

$$\pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} \ge U_{22} W_{11} + U_{11} W_{22} - U_{11} U_{22}$$
$$\pi_0 + \pi_1 W_{11} + \pi_2 W_{22} + \pi_3 W_{12} + \pi_4 T_{12} \ge L_{22} W_{11} + L_{11} W_{22} - L_{11} L_{22}$$

*are members of our class*

# Application: The Nonnegative Ball

$$\text{CH} := \text{conv}\left\{(x, xx^T) : \|x\| \le 1, x \ge 0\right\}$$

# Application: The Nonnegative Ball

$$\text{CH} := \text{conv}\left\{(x, xx^T) : \|x\| \le 1, x \ge 0\right\}$$

$$= \text{conv}\left\{(x, xx^T) : \begin{array}{l} \|x\| \le 1, x \ge 0 \\ 0 \le s^T x \le 1 \quad \forall \; \|s\| = 1 \end{array} \right\}$$

# Application: The Nonnegative Ball

$$\text{CH} := \text{conv}\left\{(x, xx^T) : \|x\| \le 1, x \ge 0\right\}$$

$$= \text{conv}\left\{(x, xx^T) : \begin{array}{l} \|x\| \le 1, x \ge 0 \\ 0 \le s^T x \le 1 \quad \forall\ \|s\| = 1 \end{array}\right\}$$

$$\subseteq \text{conv}\left\{(x, xx^T) : \begin{array}{l} \|x\| \le 1, \|x\| \le e^T x \\ 0 \le s^T x \le 1 \quad \forall\ \|s\| = 1 \end{array}\right\}$$

# Application: The Nonnegative Ball

$$\text{CH} := \text{conv}\left\{(x, xx^T) : \|x\| \le 1, x \ge 0\right\}$$

$$= \text{conv}\left\{(x, xx^T) : \begin{array}{l} \|x\| \le 1, x \ge 0 \\ 0 \le s^T x \le 1 \quad \forall\ \|s\| = 1 \end{array}\right\}$$

$$\subseteq \text{conv}\left\{(x, xx^T) : \begin{array}{l} \|x\| \le 1, \|x\| \le e^T x \\ 0 \le s^T x \le 1 \quad \forall\ \|s\| = 1 \end{array}\right\}$$

(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*

$$\mathcal{D}_{IJ} := \left\{ (x, X) : \begin{array}{l} [Xe - x]_I \geq 0 \\ [Xe - x]_J \leq 0 \end{array} \right\}.$$

*Then the following SOC constraints are locally valid on* $\mathcal{D}_{IJ}$*:*

$$\|[Xe - x]_J\| \leq 1 - \operatorname{trace}(X)$$
$$\|[Xe - x]_I\| \leq e^T x - \operatorname{trace}(X)$$

# Application: The Nonnegative Ball

**Conjecture**

*For $n = 2$, KitchenSink plus these locally valid cuts capture* $\mathrm{CH}$

# Outline

# Conclusions

THANKS TO THE MIP ORGANIZERS!