Extended formulations in combinatorial optimization: constructions and lower bounds

Samuel Fiorini

Université libre de Bruxelles (ULB, Brussels)

VI IPCO Summer School
VIII Escuela de Verano en Matemáticas Discretas

Day 3

Today

1. Quick recapitulation
2. Geometry matters
3. Polytopes with high complexity
4. Approximate extended formulations
5. A look beyond: Positive semidefinite EFs
6. Conclusion

Today

1. Quick recapitulation
2. Geometry matters
3. Polytopes with high complexity
4. Approximate extended formulations
5. A look beyond: Positive semidefinite EFs
6. Conclusion

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \ldots, v_n\}$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\mathrm{xc}(P) =$ min. size of an EF of P

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \text{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\text{xc}(P) = \text{min. size of an EF of } P$
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \text{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\text{xc}(P) =$ min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:

  $S = TU$ where $T \in \mathbb{R}_+^{m \times r}$ and $U \in \mathbb{R}_+^{r \times n}$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\mathrm{xc}(P) =$ min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:
  $$S = TU \quad \text{where} \quad T \in \mathbb{R}_+^{m \times r} \text{ and } U \in \mathbb{R}_+^{r \times n}$$
- Nonnegative rank of S:
  $$\mathrm{rk}_+(S) := \min\{r \mid \exists \text{ rank-}r \text{ nneg. factorization of } S\}$$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \text{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\text{xc}(P) =$ min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:
  $$S = TU \quad \text{where} \quad T \in \mathbb{R}_+^{m \times r} \text{ and } U \in \mathbb{R}_+^{r \times n}$$
- Nonnegative rank of S:
  $$\text{rk}_+(S) := \min\{r \mid \exists \text{ rank-}r \text{ nneg. factorization of } S\}$$
- Factorization theorem [**Yannakakis**'91]: $\text{xc}(P) = \text{rk}_+(S)$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\mathrm{xc}(P) =$ min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:
  $$S = TU \quad \text{where} \quad T \in \mathbb{R}_+^{m \times r} \text{ and } U \in \mathbb{R}_+^{r \times n}$$
- Nonnegative rank of S:
  $$\mathrm{rk}_+(S) := \min\{r \mid \exists \text{ rank-}r \text{ nneg. factorization of } S\}$$
- Factorization theorem [**Yannakakis**'91]: $\mathrm{xc}(P) = \mathrm{rk}_+(S)$
- Communication complexity: $\log \mathrm{rk}_+(S) =$ min complexity of a protocol computing S in expectation $+O(1)$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \ldots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \ldots, v_n\}$

- Extension complexity of P: $\mathrm{xc}(P) =$ min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:
  $$S = TU \quad \text{where} \quad T \in \mathbb{R}_+^{m \times r} \text{ and } U \in \mathbb{R}_+^{r \times n}$$
- Nonnegative rank of S:
  $$\mathrm{rk}_+(S) := \min\{r \mid \exists \text{ rank-}r \text{ nneg. factorization of } S\}$$
- Factorization theorem [**Yannakakis**'91]: $\mathrm{xc}(P) = \mathrm{rk}_+(S)$
- Communication complexity: $\log \mathrm{rk}_+(S) =$ min complexity of a protocol computing S in expectation $+O(1)$
- Rectangle covering bound: $\mathrm{rk}_+(S) \geqslant \mathrm{rc}(S)$

Definitions, factorization theorem and 1st bounds

Polytope $P = \{x \mid A_1 x \leqslant b_1, \dots, A_m x \leqslant b_m\} = \mathrm{conv}\{v_1, \dots, v_n\}$

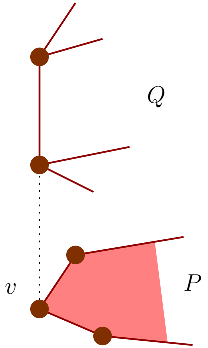
- Extension complexity of P: $\mathrm{xc}(P)$ = min. size of an EF of P
- Slack matrix $S \in \mathbb{R}_+^{m \times n}$ of P: $S_{ij} := b_i - A_i v_j$
- Rank-r nonnegative factorization of S:
  $S = TU$ where $T \in \mathbb{R}_+^{m \times r}$ and $U \in \mathbb{R}_+^{r \times n}$
- Nonnegative rank of S:
  $\mathrm{rk}_+(S) := \min\{r \mid \exists$ rank-r nneg. factorization of $S\}$
- Factorization theorem [**Yannakakis**'91]: $\mathrm{xc}(P) = \mathrm{rk}_+(S)$
- Communication complexity: $\log \mathrm{rk}_+(S)$ = min complexity of a protocol computing S in expectation $+O(1)$
- Rectangle covering bound: $\mathrm{rk}_+(S) \geqslant \mathrm{rc}(S)$
- $\mathrm{xc}(d\text{-cube}) = 2d$, $\mathrm{xc}(\text{regular } n\text{-gon}) = \Theta(\log n)$, ...

Today

Generic n-gons

Theorem (F, Rothvoß & Tiwary'12)

If P is a generic n-gon, then $\quad xc(P) \geqslant \sqrt{2n}$



Theorem (Yannakakis'91)

Every symmetric EF of the traveling salesman polytope TSP(n) := *TSP*(K_n) *has super-polynomial size. This also applies to the perfect matching polytope of* K_n.

Theorem (Yannakakis'91)

Every symmetric EF of the traveling salesman polytope $TSP(n) := TSP(K_n)$ has super-polynomial size. This also applies to the perfect matching polytope of K_n.

Theorem (Kaibel, Pashkovich & Theis'10)

Some polytopes have no poly-size symmetric EF but poly-size non-symmetric EFs

Theorem (**Yannakakis**'91)

Every symmetric EF of the traveling salesman polytope $TSP(n) := TSP(K_n)$ has super-polynomial size. This also applies to the perfect matching polytope of K_n.

Theorem (**Kaibel, Pashkovich & Theis**'10)

Some polytopes have no poly-size symmetric EF but poly-size non-symmetric EFs

Theorem (**Rothvoß'11**)

There are 0/1-polytopes P in $\mathbb{R}^d$ such that every EF has size $2^{(1/2-o(1))d}$

History

Yannakakis's problem

From **Yannakakis**'11, repeating a problem in **Yannakakis**'91:

> I believe in fact that it should be possible to prove that there is no polynomial-size formulation for the TSP polytope or any other NP-hard problem, although of course showing this remains a challenging task.

From **Yannakakis**'11, repeating a problem in **Yannakakis**'91:

> I believe in fact that it should be possible to prove that there is no polynomial-size formulation for the TSP polytope or any other NP-hard problem, although of course showing this remains a challenging task.

F, Massar, Pokutta, Tiwary & de Wolf'12 solve this problem and prove:

Theorem (FMPTW'12)

- $xc(TSP(n)) = 2^{\Omega(n^{1/2})}$

From **Yannakakis**'11, repeating a problem in **Yannakakis**'91:

> I believe in fact that it should be possible to prove that there is no polynomial-size formulation for the TSP polytope or any other NP-hard problem, although of course showing this remains a challenging task.

F, Massar, Pokutta, Tiwary & de Wolf'12 solve this problem and prove:

Theorem (**FMPTW**'12)

- $xc(TSP(n)) = 2^{\Omega(n^{1/2})}$

$$\Uparrow$$

- $xc(CUT(n)) = 2^{\Omega(n)}$

From **Yannakakis**'11, repeating a problem in **Yannakakis**'91:

> I believe in fact that it should be possible to prove that there is no polynomial-size formulation for the TSP polytope or any other NP-hard problem, although of course showing this remains a challenging task.

F, Massar, Pokutta, Tiwary & de Wolf'12 solve this problem and prove:

Theorem (**FMPTW**'12)

- $xc(TSP(n)) = 2^{\Omega(n^{1/2})}$

$$\Uparrow$$

- $xc(CUT(n)) = 2^{\Omega(n)}$

$$\Downarrow$$

- $\exists\ (G_n)$ *s.t.* $xc(STAB(G_n)) = 2^{\Omega(n^{1/2})}$

Proof

Let $M = M(n)$ be the $2^n \times 2^n$ matrix with

$$M_{ab} := (1 - a^\intercal b)^2$$

for $a, b \in \{0,1\}^n$

Proof

Let $M = M(n)$ be the $2^n \times 2^n$ matrix with

$$M_{ab} := (1 - a^\intercal b)^2$$

for $a, b \in \{0,1\}^n$

- M has rank $\Theta(n^2)$

Proof

Let $M = M(n)$ be the $2^n \times 2^n$ matrix with

$$M_{ab} := (1 - a^\intercal b)^2$$

for $a, b \in \{0,1\}^n$

- M has rank $\Theta(n^2)$
- M not a slack matrix, but embeds in a slack matrix

Proof

Let $M = M(n)$ be the $2^n \times 2^n$ matrix with

$$M_{ab} := (1 - a^\intercal b)^2$$

for $a, b \in \{0,1\}^n$

- M has rank $\Theta(n^2)$
- M not a slack matrix, but embeds in a slack matrix
- suppmat(M) appears in **de Wolf**'03 for separating classical vs. quantum nondeterministic complexity

For $b \in \{0,1\}^n$:

$$1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle \quad = \quad 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle$$

For $b \in \{0,1\}^n$:

$$\begin{aligned} 1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b) \rangle + \langle aa^\intercal, bb^\intercal \rangle \end{aligned}$$

For $b \in \{0,1\}^n$:

$$\begin{aligned} 1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b) \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\,a^\intercal b + (a^\intercal b)^2 \end{aligned}$$

For $b \in \{0,1\}^n$:

$$\begin{aligned} 1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b) \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\,a^\intercal b + (a^\intercal b)^2 \\ &= (1 - a^\intercal b)^2 \end{aligned}$$

For $b \in \{0,1\}^n$:

$$\begin{aligned} 1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b) \rangle + \langle aa^\intercal, bb^\intercal \rangle \\ &= 1 - 2\,a^\intercal b + (a^\intercal b)^2 \\ &= (1 - a^\intercal b)^2 \\ &= M_{ab} \end{aligned}$$

For $b \in \{0,1\}^n$:

$$\begin{aligned} 1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal\rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal\rangle + \langle aa^\intercal, bb^\intercal\rangle \\ &= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b)\rangle + \langle aa^\intercal, bb^\intercal\rangle \\ &= 1 - 2\,a^\intercal b + (a^\intercal b)^2 \\ &= (1 - a^\intercal b)^2 \\ &= M_{ab} \end{aligned}$$

$\mathsf{COR}(n) := \mathsf{conv}\{bb^\intercal \in \mathbb{R}^{n\times n} \mid b \in \{0,1\}^n\}$ correlation polytope

For $b \in \{0,1\}^n$:

$$
\begin{aligned}
1 - \langle 2\,\mathsf{diag}(a) - aa^\intercal, bb^\intercal \rangle &= 1 - 2\langle \mathsf{diag}(a), bb^\intercal \rangle + \langle aa^\intercal, bb^\intercal \rangle \\
&= 1 - 2\langle \mathsf{diag}(a), \mathsf{diag}(b) \rangle + \langle aa^\intercal, bb^\intercal \rangle \\
&= 1 - 2\, a^\intercal b + (a^\intercal b)^2 \\
&= (1 - a^\intercal b)^2 \\
&= M_{ab}
\end{aligned}
$$

$\mathsf{COR}(n) := \mathsf{conv}\{bb^\intercal \in \mathbb{R}^{n \times n} \mid b \in \{0,1\}^n\}$ correlation polytope

Lemma (Key lemma)

For every $a \in \{0,1\}^n$, the inequality

$$(\star) \qquad \langle 2\,\mathit{diag}(a) - aa^\intercal, x \rangle \leqslant 1$$

is valid for COR(n). The slack of vertex $bb^\intercal$ w.r.t. $(\star)$ is M_{ab}.

Consider complete linear description for $\mathsf{COR}(n)$ starting with

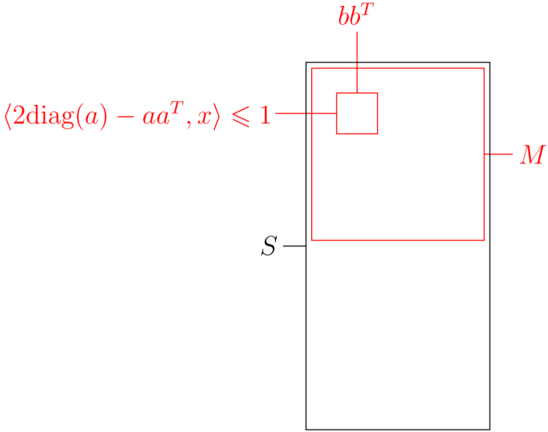
$$\langle 2\,\mathsf{diag}(a) - aa^\intercal, x\rangle \leqslant 1 \quad \forall a \in \{0,1\}^n$$

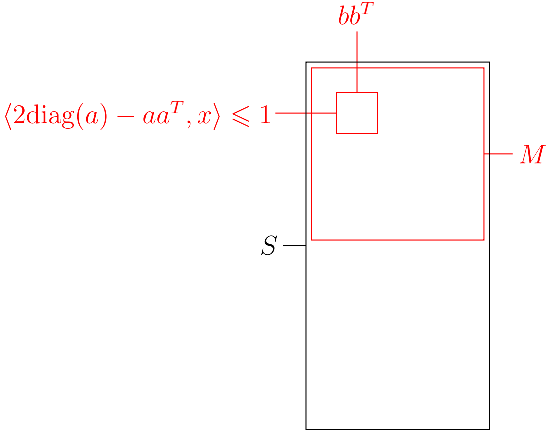
and corresponding slack matrix S

Consider complete linear description for $\mathrm{COR}(n)$ starting with

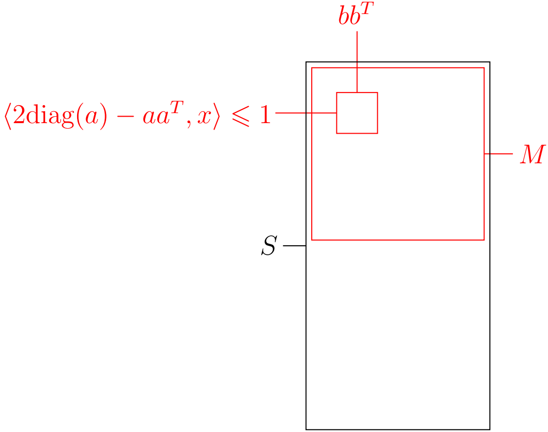
$$\langle 2\,\mathrm{diag}(a) - aa^{\mathsf{T}}, x\rangle \leqslant 1 \quad \forall a \in \{0,1\}^n$$

and corresponding slack matrix S

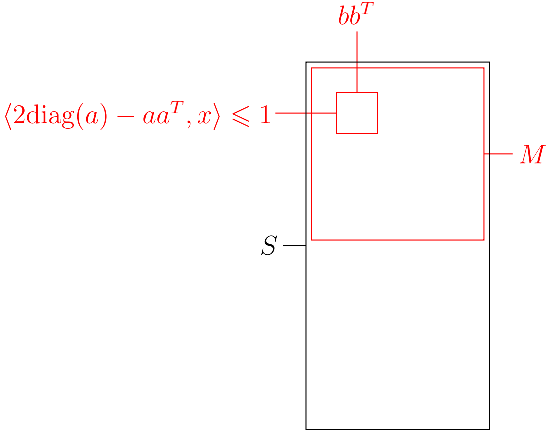




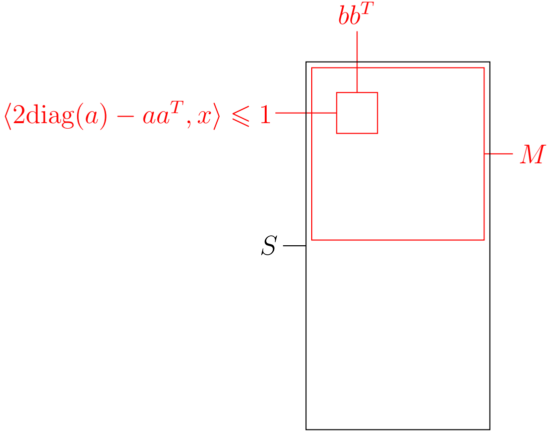
$\mathsf{xc}(\mathsf{COR}(n))$



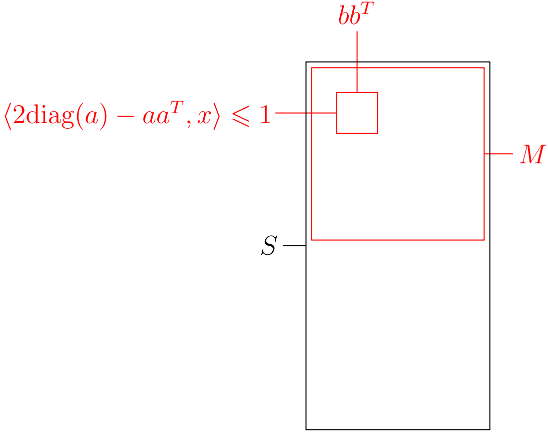
$$\mathsf{xc}(\mathsf{COR}(n)) \quad = \quad \mathsf{rk}_+(S)$$



$$\begin{aligned} \mathsf{xc}(\mathsf{COR}(n)) &= \mathsf{rk}_+(S) \\ &\geqslant \mathsf{rk}_+(M) \end{aligned}$$



$$\begin{aligned}
\mathsf{xc}(\mathsf{COR}(n)) &= \mathsf{rk}_+(S) \\
&\geqslant \mathsf{rk}_+(M) \\
&\geqslant \mathsf{rc}(M)
\end{aligned}$$



$$
\begin{aligned}
\mathsf{xc}(\mathsf{COR}(n)) &= \mathsf{rk}_+(S) \\
&\geqslant \mathsf{rk}_+(M) \\
&\geqslant \mathsf{rc}(M) \\
&= 2^{\Omega(n)} \qquad \text{(\textbf{de Wolf}'03, building on \textbf{Razborov}'92)}
\end{aligned}
$$

Cut polytope. $\boxed{\mathsf{xc}(\mathsf{CUT}(n)) = \mathsf{xc}(\mathsf{COR}(n-1)) = 2^{\Omega(n)}}$

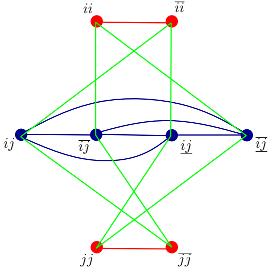
Cut polytope. $\boxed{\mathsf{xc}(\mathsf{CUT}(n)) = \mathsf{xc}(\mathsf{COR}(n-1)) = 2^{\Omega(n)}}$

Lemma (monotonicity)

- Q *is an extension of* $P \implies xc(Q) \geqslant xc(P)$
- P *contains* F *as a face* $\implies xc(P) \geqslant xc(F)$

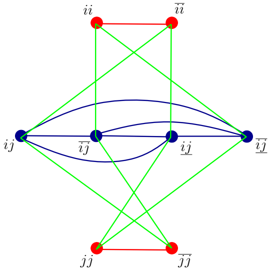
Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).

Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).



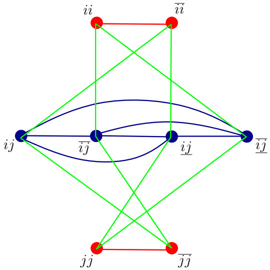
xc(STAB(H_k))

Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).



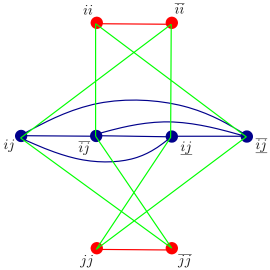
$$\text{xc}(\text{STAB}(H_k)) \quad \geqslant \quad \text{xc}(F(k))$$

Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).



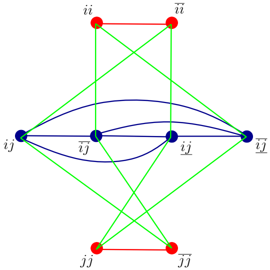
$$\begin{aligned}\text{xc}(\text{STAB}(H_k)) &\geqslant \text{xc}(F(k)) \\ &\geqslant \text{xc}(\text{COR}(k))\end{aligned}$$

Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).



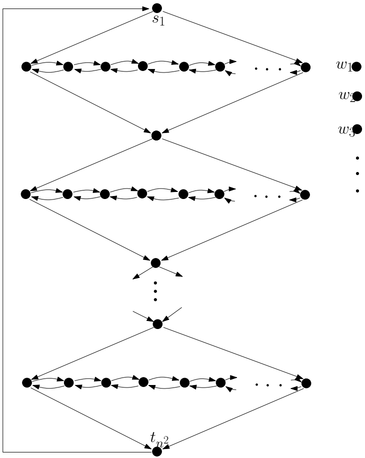
$$
\begin{aligned}
\text{xc}(\text{STAB}(H_k)) &\geqslant \text{xc}(F(k)) \\
&\geqslant \text{xc}(\text{COR}(k)) \\
&= 2^{\Omega(k)}
\end{aligned}
$$

Stable set polytope. $\forall\ k\ \exists\ H_k$ with $O(k^2)$ vertices s.t. STAB(H_k) has a face $F = F(k)$ that is an extension of COR(k).

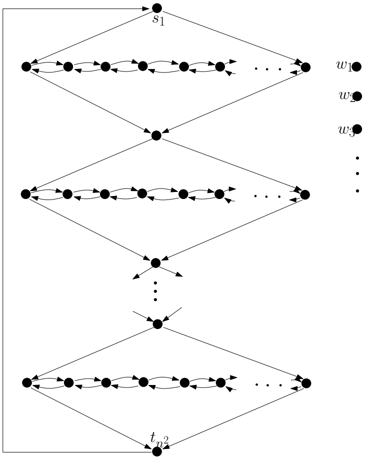


$$\begin{aligned} \text{xc}(\text{STAB}(H_k)) &\geqslant \text{xc}(F(k)) \\ &\geqslant \text{xc}(\text{COR}(k)) \\ &= 2^{\Omega(k)} \end{aligned}$$

$\Longrightarrow \forall\ n\ \exists\ n$-vertex G_n s.t. $\boxed{\text{xc}(\text{STAB}(G_n)) = 2^{\Omega(n^{1/2})}}$



$\forall\, k\ \exists\, H_k$ with $O(k^2)$ vertices s.t. TSP(H_k) is an extension of COR(k).



$\forall\, k\ \exists\, H_k$ with $O(k^2)$ vertices s.t. TSP(H_k) is an extension of COR(k).

Theorem (FMPTW'12)

$xc(TSP(n)) = 2^{\Omega(n^{1/2})}$

Extension/extended formulation of a pair

- $P \subseteq Q \subseteq \mathbb{R}^d$ with P polytope, Q polyhedron
- $L \subseteq \mathbb{R}^e$ polytope

Definition (extension/EF of a pair)

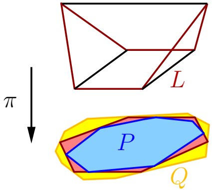
L is an extension of (P, Q) if $\exists$ linear π with $P \subseteq \pi(L) \subseteq Q$

Extension/extended formulation of a pair

- $P \subseteq Q \subseteq \mathbb{R}^d$ with P polytope, Q polyhedron
- $L \subseteq \mathbb{R}^e$ polytope

Definition (extension/EF of a pair)

L is an extension of (P, Q) if $\exists$ linear π with $P \subseteq \pi(L) \subseteq Q$

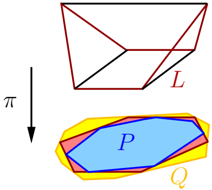


Extension/extended formulation of a pair

- $P \subseteq Q \subseteq \mathbb{R}^d$ with P polytope, Q polyhedron
- $L \subseteq \mathbb{R}^e$ polytope

Definition (extension/EF of a pair)

L is an extension of (P, Q) if $\exists$ linear π with $P \subseteq \pi(L) \subseteq Q$



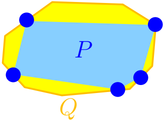
Definition (extension complexity of a pair)

$\text{xc}(P, Q) = \min\{\#\text{facets}(L) \mid L \text{ is an extension of } (P, Q)\}$

Slack matrix of a pair

Let $V = \{v_1, \dots, v_\ell\} \subseteq \mathbb{R}^d$ s.t. $P = \mathrm{conv}(V)$

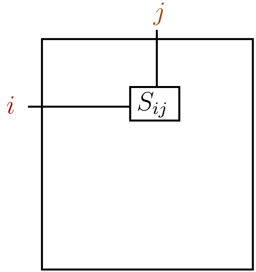
Let $A \in \mathbb{R}^{k \times d}$, $b \in \mathbb{R}^k$ s.t. $Q = \{x \in \mathbb{R}^d \mid Ax \leqslant b\}$

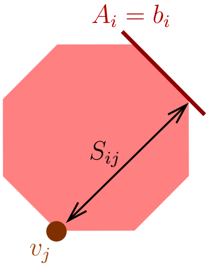


Definition (slack matrix)

Slack matrix $S = S^{P,Q} \in \mathbb{R}_+^{k \times \ell}$ of (P, Q) (w.r.t. $Ax \leqslant b$ and V):

$$S_{ij}^{P,Q} := b_i - A_i v_j$$





Remark. Every nneg matrix is the slack matrix of some pair!

Remark. Every nneg matrix is the slack matrix of some pair!

Theorem (Factorization theorem for pairs)

For every slack matrix $S^{P,Q}$ of (P, Q): $\quad$ $\mathit{xc}(P, Q) = \mathit{rk}_+(S^{P,Q})$

Linear encodings of combinatorial optimization problems

Definition (linear encoding)

A linear encoding of a comb. opt. problem is a pair $(\mathcal{L}, \mathcal{O})$ where

- $\mathcal{L} \subseteq \{0,1\}^*$ feasible solutions
- $\mathcal{O} \subseteq \mathbb{R}^*$ admissible objective functions

An instance is a pair (d, w) where $d \geqslant 1$ and $w \in \mathcal{O} \cap \mathbb{R}^d$

Given (d, w)*, find* $x \in \mathcal{L} \cap \{0,1\}^d$ *such that* $w^\intercal x$ *is max/min*

Linear encodings of combinatorial optimization problems

Definition (linear encoding)

A linear encoding of a comb. opt. problem is a pair $(\mathcal{L}, \mathcal{O})$ where

- $\mathcal{L} \subseteq \{0,1\}^*$ feasible solutions
- $\mathcal{O} \subseteq \mathbb{R}^*$ admissible objective functions

An instance is a pair (d, w) where $d \geqslant 1$ and $w \in \mathcal{O} \cap \mathbb{R}^d$

Given (d, w), find $x \in \mathcal{L} \cap \{0,1\}^d$ such that $w^\intercal x$ is max/min

"Faithfulness" condition:

instances of the problem $\longrightarrow$ instances of the linear encoding

"Constraints do not depend on instance, only on d."
"Instances are encoded in the objective function."

Approximate extended formulations

For a maximization problem

Linear encoding $(\mathcal{L}, \mathcal{O}) \rightsquigarrow$ pair of nested polyhedra $P \subseteq Q$:

- $P := \mathrm{conv}(\{x \in \{0,1\}^d \mid x \in \mathcal{L}\})$
- $Q := \{x \in \mathbb{R}^d \mid \forall w \in \mathcal{O} \cap \mathbb{R}^d : w^\intercal x \leqslant \max\{w^\intercal y \mid y \in P\}\}$

Approximate extended formulations

For a maximization problem

Linear encoding $(\mathcal{L}, \mathcal{O}) \rightsquigarrow$ pair of nested polyhedra $P \subseteq Q$:

- $P := \text{conv}(\{x \in \{0,1\}^d \mid x \in \mathcal{L}\})$
- $Q := \{x \in \mathbb{R}^d \mid \forall w \in \mathcal{O} \cap \mathbb{R}^d : w^\intercal x \leqslant \max\{w^\intercal y \mid y \in P\}\}$

Definition (ρ-approximate extended formulation, $\rho \geqslant 1$)

$Ex + Fy = g$, $y \geqslant \mathbf{0}$ is a ρ-approximate EF w.r.t. $(\mathcal{L}, \mathcal{O})$ if

1. $\forall w \in \mathbb{R}^d$:
$$\max\{w^\intercal x \mid Ex + Fy = g,\ y \geqslant \mathbf{0}\} \geqslant \max\{w^\intercal x \mid x \in P\}$$
2. $\forall w \in \mathcal{O} \cap \mathbb{R}^d$:
$$\max\{w^\intercal x \mid Ex + Fy = g,\ y \geqslant \mathbf{0}\} \leqslant \rho \max\{w^\intercal x \mid x \in P\}$$

Approximate extended formulations

For a maximization problem

Linear encoding $(\mathcal{L}, \mathcal{O}) \rightsquigarrow$ pair of nested polyhedra $P \subseteq Q$:

- $P := \mathrm{conv}(\{x \in \{0,1\}^d \mid x \in \mathcal{L}\})$
- $Q := \{x \in \mathbb{R}^d \mid \forall w \in \mathcal{O} \cap \mathbb{R}^d : w^\intercal x \leqslant \max\{w^\intercal y \mid y \in P\}\}$

Definition (ρ-approximate extended formulation, $\rho \geqslant 1$)

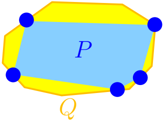
$Ex + Fy = g$, $y \geqslant \mathbf{0}$ is a ρ-approximate EF w.r.t. $(\mathcal{L}, \mathcal{O})$ if

1. $\forall w \in \mathbb{R}^d$:
$$\max\{w^\intercal x \mid Ex + Fy = g,\ y \geqslant \mathbf{0}\} \geqslant \max\{w^\intercal x \mid x \in P\}$$
2. $\forall w \in \mathcal{O} \cap \mathbb{R}^d$:
$$\max\{w^\intercal x \mid Ex + Fy = g,\ y \geqslant \mathbf{0}\} \leqslant \rho \max\{w^\intercal x \mid x \in P\}$$

Geometrically: $P \subseteq \{x \mid \exists y : Ex + Fy = g,\ y \geqslant \mathbf{0}\} \subseteq \rho Q$

Sizes of approximate extended formulations

- $\mathcal{L} \rightsquigarrow P = \text{conv}(V)$
- $\mathcal{O} \rightsquigarrow Q = \{x \in \mathbb{R}^d \mid Ax \leqslant b\}$



Observation:

1. $\rho Q = \{x \in \mathbb{R}^d \mid Ax \leqslant \rho b\}$
2. $S_{ij}^{P,\rho Q} = \rho b_i - A_i v_j = S_{ij}^{P,Q} + (\rho - 1) b_i$

Corollary

Minimum size of a ρ-approximate EF $= \mathrm{rk}_+(S^{P,\rho Q})$

Polyhedral inapproximability of CLIQUE

Theorem

W.r.t. natural (faithful) linear encoding, CLIQUE has

- *a (trivial) poly-size n-approximate EF*
- *no $2^{o(n^{2\epsilon})}$-size $n^{1/2-\epsilon}$-approximate EF, for all $\epsilon \in (0, 1/2)$*

Polyhedral inapproximability of CLIQUE

Theorem

W.r.t. natural (faithful) linear encoding, CLIQUE has

- *a (trivial) poly-size n-approximate EF*
- *no $2^{o(n^{2\epsilon})}$-size $n^{1/2-\epsilon}$-approximate EF, for all $\epsilon \in (0, 1/2)$*

NEWS: Braverman and Moitra improved $n^{1/2-\epsilon}$ to $n^{1-\epsilon}$ (tight)

Polyhedral inapproximability of CLIQUE

Theorem

W.r.t. natural (faithful) linear encoding, CLIQUE has

- *a (trivial) poly-size n-approximate EF*
- *no $2^{o(n^{2\epsilon})}$-size $n^{1/2-\epsilon}$-approximate EF, for all $\epsilon \in (0, 1/2)$*

NEWS: Braverman and Moitra improved $n^{1/2-\epsilon}$ to $n^{1-\epsilon}$ (tight)

The encoding:

- $d = n^2$
- $x \in \{0,1\}^{n \times n}$ is feasible if $x_{ij} = b_i b_j$ for $b \in \{0,1\}^n$
- $w \in \mathbb{R}^{n \times n}$ is admissible if
  - $w_{ii} \in \{0,1\}$ for all i
  - $w_{ij} = w_{ji} \in \{-1, 0\}$ for all i, j
- G graph with $V(G) \subseteq [n] \longrightarrow w^G := I(G) - A(\bar{G})$

Theorem (nonnegative rank of shifted UDISJ)

Let M *be any* $2^n \times 2^n$ *matrix such that*

- $M_{ab} = 1 + (\rho - 1)$ *if* $|a \cap b| = 0$
- $M_{ab} = 0 + (\rho - 1)$ *if* $|a \cap b| = 1$

Then

1. *rk*$_+(M) = 2^{\Omega(n)}$ *if* ρ *is a fixed constant*
2. *rk*$_+(M) = 2^{\Omega(n^{1-2\beta})}$ *if* $\rho = O(n^\beta)$ *for some constant* $\beta < 1/2$

Remark: This is the Unique DISJointness (partial) matrix when $\rho = 1$.

The corruption bound in general

Consider any matrix $S \in \mathbb{R}_{+}^{k \times \ell}$

Assume that **weights** $\mu \in \mathbb{R}^{k \times \ell}$ ($-$'s allowed) satisfy:

$$\langle \mu, X \rangle \leqslant ||X||_{\infty} \quad \forall X \in \mathbb{R}_{+}^{k \times \ell} \text{ that is rank-1}$$

Then if $S = \sum_{i=1}^{r} X_i$ where $X_i \in \mathbb{R}_{+}^{k \times \ell}$ are rank-1:

The corruption bound in general

Consider any matrix $S \in \mathbb{R}_{+}^{k\times\ell}$

Assume that **weights** $\mu \in \mathbb{R}^{k\times\ell}$ ($-$'s allowed) satisfy:

$$\langle \mu, X\rangle \leqslant ||X||_\infty \quad \forall X \in \mathbb{R}_{+}^{k\times\ell} \text{ that is rank-1}$$

Then if $S = \sum_{i=1}^{r} X_i$ where $X_i \in \mathbb{R}_{+}^{k\times\ell}$ are rank-1:

$$\begin{aligned}
\langle \mu, S\rangle &= \langle \mu, \sum_{i=1}^{r} X_i\rangle \\
&= \sum_{i=1}^{r} \langle \mu, X_i\rangle \\
&\leqslant \sum_{i=1}^{r} ||X_i||_\infty \\
&\leqslant r||S||_\infty
\end{aligned}$$

The corruption bound in general

Consider any matrix $S \in \mathbb{R}_{+}^{k \times \ell}$

Assume that **weights** $\mu \in \mathbb{R}^{k \times \ell}$ ($-$'s allowed) satisfy:

$$\langle \mu, X \rangle \leqslant ||X||_\infty \quad \forall X \in \mathbb{R}_{+}^{k \times \ell} \text{ that is rank-1}$$

Then if $S = \sum_{i=1}^{r} X_i$ where $X_i \in \mathbb{R}_{+}^{k \times \ell}$ are rank-1:

$$\begin{aligned}
\langle \mu, S \rangle &= \langle \mu, \sum_{i=1}^{r} X_i \rangle \\
&= \sum_{i=1}^{r} \langle \mu, X_i \rangle \\
&\leqslant \sum_{i=1}^{r} ||X_i||_\infty \\
&\leqslant r||S||_\infty
\end{aligned}$$

$$\implies \boxed{r \geqslant \frac{\langle \mu, S \rangle}{||S||_\infty}}$$

Core tool: Razborov's corruption lemma

Let $1 \leqslant \ell \leqslant (n+1)/4$

Distribution μ **on pairs** $(a, b) \in 2^{[n]} \times 2^{[n]}$ **with** $|a| = |b| = \ell$:

- on $A := \{(a, b) \mid |a \cap b| = 0\}$: $\frac{1}{4} \times$ uniform
- on $B := \{(a, b) \mid |a \cap b| = 1\}$: $\frac{3}{4} \times$ uniform

Core tool: Razborov's corruption lemma

Let $1 \leqslant \ell \leqslant (n+1)/4$

Distribution μ **on pairs** $(a,b) \in 2^{[n]} \times 2^{[n]}$ **with** $|a| = |b| = \ell$:

- on $A := \{(a,b) \mid |a \cap b| = 0\}$: $\frac{1}{4} \times$ uniform
- on $B := \{(a,b) \mid |a \cap b| = 1\}$: $\frac{3}{4} \times$ uniform

Random variable: $X(a,b) := f(a)g(b)$ with $f, g \geqslant 0$

Core tool: Razborov's corruption lemma

Let $1 \leqslant \ell \leqslant (n+1)/4$

Distribution μ **on pairs** $(a,b) \in 2^{[n]} \times 2^{[n]}$ **with** $|a| = |b| = \ell$:

- on $A := \{(a,b) \mid |a \cap b| = 0\}$: $\frac{1}{4} \times$ uniform
- on $B := \{(a,b) \mid |a \cap b| = 1\}$: $\frac{3}{4} \times$ uniform

Random variable: $X(a,b) := f(a)g(b)$ with $f, g \geqslant 0$

Lemma (Razborov's corruption lemma, improved)

Then for every $0 < \epsilon < 1$*:*

$$2^{\frac{\epsilon^2}{4 \ln 2}\ell + O(\log \ell)} \left((1-\epsilon)\,\mathbb{E}\left[X \mid A\right] - \mathbb{E}\left[X \mid B\right]\right) \leqslant \|X\|_\infty$$

Today

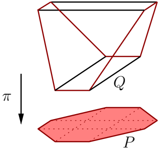
1. Quick recapitulation
2. Geometry matters
3. Polytopes with high complexity
4. Approximate extended formulations
5. A look beyond: Positive semidefinite EFs
6. Conclusion

The threefold way

The following are equivalent [**Yannakakis**'88/91, **FFGT**'11]:

1. A linear system $Ex + Fy = g$, $y \geqslant \mathbf{0}$ with $y \in \mathbb{R}^r$ s.t.

$$P = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^r : Ex + Fy = g,\ y \geqslant \mathbf{0}\}$$

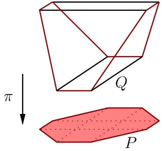


The threefold way

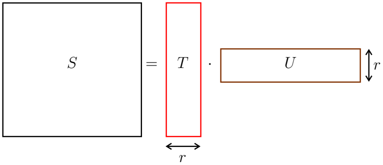
The following are equivalent [**Yannakakis**'88/91, **FFGT**'11]:

1. A linear system $Ex + Fy = g$, $y \geqslant \mathbf{0}$ with $y \in \mathbb{R}^r$ s.t.

$$P = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^r : Ex + Fy = g,\ y \geqslant \mathbf{0}\}$$



2. A rank-r nonnegative factorization $S = TU$ of slack matrix S

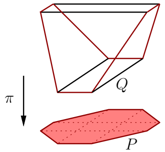


The threefold way

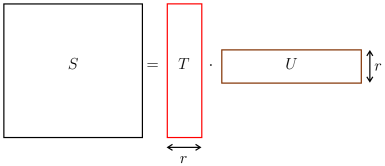
The following are equivalent [**Yannakakis**'88/91, **FFGT**'11]:

1. A linear system $Ex + Fy = g,\ y \geqslant \mathbf{0}$ with $y \in \mathbb{R}^r$ s.t.

$$P = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^r : Ex + Fy = g,\ \ y \geqslant \mathbf{0}\}$$



2. A rank-r nonnegative factorization $S = TU$ of slack matrix S

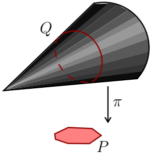


3. A $\log r$-complexity randomized protocol computing S in expectation

The threefold way (revisited)

The following are equivalent [**FMPTW**'12; **Gouveia, Parillo & Thomas**'11]:

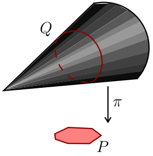
1. A semidefinite EF $Ex + Fy = g$, $y \succeq \mathbf{0}$ with $y \in \mathbb{R}^{r \times r}$



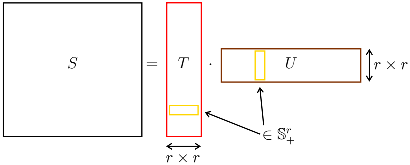
The threefold way (revisited)

The following are equivalent [**FMPTW**'12; **Gouveia, Parillo & Thomas**'11]:

1. A semidefinite EF $Ex + Fy = g$, $y \succeq \mathbf{0}$ with $y \in \mathbb{R}^{r \times r}$



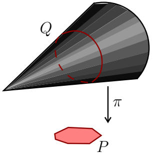
2. A rank-r PSD factorization $S_{ij} = \langle T_i, U^j \rangle$ of slack matrix S



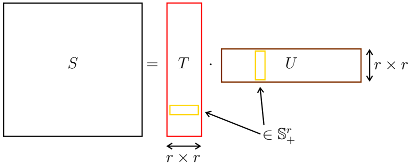
The threefold way (revisited)

The following are equivalent [**FMPTW**'12; **Gouveia, Parillo & Thomas**'11]:

1. A semidefinite EF $Ex + Fy = g$, $y \succeq \mathbf{0}$ with $y \in \mathbb{R}^{r \times r}$



2. A rank-r PSD factorization $S_{ij} = \langle T_i, U^j \rangle$ of slack matrix S



3. A $\log r$-complexity quantum one-way protocol computing S in expectation

Today

Take home messages

- EFs are important and interesting

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP
- We can prove some inapproximability results

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP
- We can prove some inapproximability results
- There are **many** links to other areas

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP
- We can prove some inapproximability results
- There are **many** links to other areas
- There are **many** open problems

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP
- We can prove some inapproximability results
- There are **many** links to other areas
- There are **many** open problems
- Stay tuned (a new survey is coming)

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of $\mathrm{P} \neq \mathrm{NP}$
- We can prove some inapproximability results
- There are **many** links to other areas
- There are **many** open problems
- Stay tuned (a new survey is coming)

Take home messages

- EFs are important and interesting
- EFs give a way to understand the power of LPs (and SDPs)
- We can prove a geometric analogue of P $\neq$ NP
- We can prove some inapproximability results
- There are **many** links to other areas
- There are **many** open problems
- Stay tuned (a new survey is coming)

Thank You! :)