The Permutahedron
The Permutahedron

Let $\mathbf{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\mathbf{t})\mid \pi \text{ is permutation}\}$
The Permutahedron

Let $\overrightarrow{t} = [1, 2, 3, \ldots, n]$, \quad P = \text{conv}\{\pi(\overrightarrow{t}) \mid \pi \text{ is permutation}\}

How many facets of $P$ have?
The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\vec{t}) | \pi \text{ is permutation}\}$

How many facets of $P$ have? exponentially many!
The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\vec{t})\mid \pi \text{ is permutation}\}$

How many facets of $P$ have?

e.g. $S \subseteq [n]$, $\sum_{i \in S} x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

exponentially many!
The Permutahedron

Let $\overrightarrow{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\overrightarrow{t}) \mid \pi \text{ is permutation}\}$

How many facets of $P$ have? exponentially many!

e.g. $S \subset [n]$, $\sum_{i \in S} x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

Let $Q = \{A \mid A \text{ is doubly-stochastic}\}$
The Permutahedron

Let $\mathbf{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\mathbf{t}) | \pi \text{ is permutation}\}$

How many facets of $P$ have? exponentially many!

e.g. $S \subset [n]$, $\sum_{i \in S} x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

Let $Q = \{A | A \text{ is doubly-stochastic}\}$

Then $P$ is the projection of $Q$: $P = \{A^\top A \text{ in } Q \}$
The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots, n]$, $P = \text{conv}\{\pi(\vec{t})| \pi \text{ is permutation}\}$

How many facets of $P$ have?

exponentially many!

e.g. $S \subset [n]$, $\sum_{i \in S} x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

Let $Q = \{A| A \text{ is doubly-stochastic}\}$

Then $P$ is the projection of $Q$: $P = \{A \uparrow \downarrow A \text{ in } Q\}$

Yet $Q$ has only $O(n^2)$ facets
Extended Formulations

The **extension complexity (xc)** of a polytope $P$ is the minimum number of facets of $Q$ so that $P = \text{proj}(Q)$.
Extended Formulations

The **extension complexity (xc)** of a polytope $P$ is the minimum number of facets of $Q$ so that $P = \text{proj}(Q)$

e.g. $xc(P) = \Theta(n \log n)$ for permutahedron
**Extended Formulations**

The **extension complexity** (xc) of a polytope $P$ is the minimum number of facets of $Q$ so that $P = \text{proj}(Q)$.

e.g. $\text{xc}(P) = \Theta(n \log n)$ for permutahedron.

$\text{xc}(P) = \Theta(\log n)$ for a regular $n$-gon, but $\Omega(\sqrt{n})$ for its perturbation.
Extended Formulations

The **extension complexity** \((xc)\) of a polytope \(P\) is the minimum number of facets of \(Q\) so that \(P = \text{proj}(Q)\).

- \(xc(P) = \Theta(n \log n)\) for permutahedron
- \(xc(P) = \Theta(\log n)\) for a regular \(n\)-gon, but \(\Omega(\sqrt{n})\) for its perturbation

In general, \(P = \{ x \mid \exists y, (x,y) \in Q \}\)
Extended Formulations

The **extension complexity (xc)** of a polytope $P$ is the minimum number of facets of $Q$ so that $P = \text{proj}(Q)$

- $\text{xc}(P) = \Theta(n \log n)$ for a permutahedron
- $\text{xc}(P) = \Theta(\log n)$ for a regular $n$-gon, but $\Omega(\sqrt{n})$ for its perturbation

In general, $P = \{ x \mid \exists y, (x,y) \in Q \}$

...analogy with **quantifiers** in Boolean formulae
Applications of EFs

In general, \( P = \{x \mid \exists y, (x,y) \in Q\} \)
Applications of EFs

In general, \( P = \{x \mid \exists y, (x,y) \text{ in } Q\} \)

Through EFs, we can reduce # facets exponentially!
Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm
Applications of EFs

In general, \( P = \{ x \mid \exists y, (x,y) \text{ in } Q \} \)

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm.

EFs often give, or are based on new combinatorial insights.
Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \in Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

EFs often give, or are based on new combinatorial insights

  e.g. Birkhoff-von Neumann Thm and permutahedron
Applications of EFs

In general, \( P = \{ x \mid \exists y, (x,y) \text{ in } Q \} \)

Through EFs, we can reduce \# facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

EFs often give, or are based on new combinatorial insights

  e.g. Birkhoff-von Neumann Thm and permutahedron

  e.g. prove there is low-cost object, through its polytope
Explicit, Hard Polytopes?
Explicit, Hard Polytopes?

Definition: TSP polytope:

\[ P = \text{conv}\{1_F \mid F \text{ is the set of edges on a tour of } K_n\} \]
Explicit, Hard Polytopes?

**Definition:** TSP polytope:

\[ P = \text{conv}\{1_F | F \text{ is the set of edges on a tour of } K_n\} \]

(If we could optimize over this polytope, then \( P = \text{NP} \))
Explicit, Hard Polytopes?

**Definition:** TSP polytope:

\[ P = \text{conv}\{1_F | F \text{ is the set of edges on a tour of } K_n\} \]

(If we could optimize over this polytope, then \( P = \text{NP} \))

Can we prove **unconditionally** there is no small EF?
Explicit, Hard Polytopes?

**Definition:** TSP polytope:

\[ P = \text{conv}\{1_F | F \text{ is the set of edges on a tour of } K_n\} \]

(If we could optimize over this polytope, then P = NP)

**Can we prove unconditionally** there is no small EF?

**Caveat:** this is unrelated to proving complexity l.b.s
Explicit, Hard Polytopes?

Definition: TSP polytope:

\[ P = \text{conv} \{ 1_F \mid F \text{ is the set of edges on a tour of } K_n \} \]

(If we could optimize over this polytope, then P = NP)

Can we prove unconditionally there is no small EF?

Caveat: this is unrelated to proving complexity l.b.s

[Yannakakis ’90]: Yes, through the nonnegative rank
An Abridged History

Theorem [Yannakakis ’90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$.
An Abridged History

**Theorem [Yannakakis ’90]:** Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

...but asymmetric EFs can be more powerful
An Abridged History

**Theorem [Yannakakis ’90]:** Any symmetric EF for TSP or matching has size $2^\Omega(n)$

…but asymmetric EFs can be more powerful
An Abridged History

Theorem [Yannakakis ’90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

...but asymmetric EFs can be more powerful

Theorem [Fiorini et al ’12]: Any EF for TSP has size $2^{\Omega(\sqrt{n})}$ (based on a $2^{\Omega(n)}$ lower bd for clique)

Approach: connections to non-deterministic CC
Theorem [Braun et al ’12]: Any EF that approximates clique within $n^{1/2-\epsilon}$ has size $\exp(n^{\epsilon})$

**Approach:** Razborov’s rectangle corruption lemma
An Abridged History II

Theorem [Braun et al ’12]: Any EF that approximates clique within $n^{1/2-\varepsilon}$ has size $\exp(n^{\varepsilon})$

**Approach:** Razborov’s rectangle corruption lemma

Theorem [Braverman, Moitra ’13]: Any EF that approximates clique within $n^{1-\varepsilon}$ has size $\exp(n^{\varepsilon})$

**Approach:** information complexity
An Abridged History II

**Theorem [Braun et al ’12]:** Any EF that approximates clique within $n^{1/2-\epsilon}$ has size $\exp(n^\epsilon)$

**Approach:** Razborov’s rectangle corruption lemma

**Theorem [Braverman, Moitra ’13]:** Any EF that approximates clique within $n^{1-\epsilon}$ has size $\exp(n^\epsilon)$

**Approach:** information complexity

see also [Braun, Pokutta ’13]: reformulation using common information, applications to avg. case
An Abridged History III

**Theorem [Chan et al ’12]:** Any EF that approximates MAXCUT within 2-eps has size $n^{\Omega(\log n/\log\log n)}$

**Approach:** reduction to Sherali-Adams
An Abridged History III

**Theorem [Chan et al ’12]:** Any EF that approximates MAXCUT within 2-\(\varepsilon\) has size \(n^{\Omega(\log n / \log \log n)}\)

**Approach:** reduction to Sherali-Adams

**Theorem [Rothvoss ’13]:** Any EF for perfect matching has size \(2^{\Omega(n)}\) (same for TSP)

**Approach:** hyperplane separation lower bound
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
Outline

Part I: Tools for Extended Formulations
- Yannakakis’s Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications
- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope
The Factorization Theorem
The Factorization Theorem

How can we prove lower bounds on EFs?
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter \quad \leftrightarrow \quad Algebraic Parameter
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter \[\text{↔} \quad \text{Algebraic Parameter}\]

Definition of the \textbf{slack matrix}…
The Slack Matrix
The Slack Matrix
The Slack Matrix

\[
\text{vertex} \quad \quad \quad \quad S(P) \quad \quad \quad \quad \text{facet}
\]

\[
P
\]
The Slack Matrix

facet

S(P)

vertex

v_j
The Slack Matrix

$S(P)$

$\langle a_i, x \rangle \leq b_i$

vertex

facet
The Slack Matrix

The entry in row $i$, column $j$ is how slack the $j^{th}$ vertex is on the $i^{th}$ constraint.

$$\langle a_i, x \rangle \leq b_i$$

$S(P)$

vertex

facet

$P$

$v_j$
The Slack Matrix

The entry in row $i$, column $j$ is how slack the $j^{th}$ vertex is on the $i^{th}$ constraint.

$$<a_i, x> \leq b_i$$

$$b_i - <a_i, v_j>$$

The entry in row $i$, column $j$ is how slack the $j^{th}$ vertex is on the $i^{th}$ constraint.
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter  Algebraic Parameter

Definition of the slack matrix...
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter \leftrightarrow Algebraic Parameter

Definition of the \textit{slack matrix}…

Definition of the \textit{nonnegative rank}…
Nonnegative Rank

\[ S = \]
Nonnegative Rank

rank one, nonnegative

\[ S = M_1 + \ldots + M_r \]
Nonnegative Rank

\[ S = M_1 + \ldots + M_r \]

**Definition:** \( \text{rank}^+(S) \) is the smallest \( r \) s.t. \( S \) can be written as the sum of \( r \) rank one, nonnegative matrices.
Nonnegative Rank

**Definition:** \( \text{rank}^+(S) \) is the smallest \( r \) s.t. \( S \) can be written as the sum of \( r \) rank one, nonnegative matrices.

\[ S = M_1 + \ldots + M_r \]

**Note:** \( \text{rank}^+(S) \geq \text{rank}(S) \), but can be much larger too!
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter \rightarrow Algebraic Parameter
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]: $xc(P) = \text{rank}^+(S(P))$

Geometric Parameter $\leftrightarrow$ Algebraic Parameter
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]: \( xc(P) = \text{rank}^+ (S(P)) \)

Geometric Parameter \hspace{1cm} \longleftrightarrow \hspace{1cm} Algebraic Parameter

Intuition: the factorization gives a change of variables that preserves the slack matrix!
The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]: $xc(P) = \text{rank}^+(S(P))$

Intuition: the factorization gives a change of variables that preserves the slack matrix!

Next we will give a method to lower bound $\text{rank}^+$ via 
information complexity…
Outline

Part I: Tools for Extended Formulations
• Yannakakis’s Factorization Theorem
• The Rectangle Bound
• A Sampling Argument

Part II: Applications
• Correlation Polytope
• Approximating the Correlation Polytope
• Matching Polytope
Outline

Part I: Tools for Extended Formulations
   • Yannakakis’s Factorization Theorem
   • The Rectangle Bound
   • A Sampling Argument

Part II: Applications
   • Correlation Polytope
   • Approximating the Correlation Polytope
   • Matching Polytope
The Rectangle Bound

$S = \mathbf{M}_1 + \ldots + \mathbf{M}_r$

rank one, nonnegative
The Rectangle Bound

rank one, nonnegative

\[ M = M_1 + \ldots + M_r \]
The Rectangle Bound

rank one, nonnegative

\[ \begin{array}{c}
\text{[Image of rectangle]} \\
= \\
\text{[Image of factorization]} \\
\text{[Image of sum and \(M_r\)]}
\end{array} \]
The Rectangle Bound

rank one, nonnegative

\[ \begin{array}{c}
= \\
\end{array} \ ] + \ldots + M_r \]
The Rectangle Bound

rank one, nonnegative

\[
\begin{array}{c}
= \\
\hline
\end{array}
\]
The Rectangle Bound

rank one, nonnegative

The support of each $M_i$ is a combinatorial rectangle
The Rectangle Bound

rank one, nonnegative

rank^+(S) is at least \# rectangles needed to cover supp of S
The Rectangle Bound

rank one, nonnegative

\[ \text{rank}^+(S) \text{ is at least } \# \text{ rectangles needed to cover supp of } S \]
The Rectangle Bound

rank one, nonnegative

Non-deterministic Comm. Complexity

\[ \text{rank}^+(S) \text{ is at least } \# \text{ rectangles needed to cover supp of } S \]
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • **A Sampling Argument**

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
A Sampling Argument

\[ \begin{align*}
\text{[Image 1]} & = \begin{array}{|c|c|c|}
\hline
\text{[Image 2]} & \text{[Image 3]} & \text{[Image 4]} \\
\hline
\end{array}
\end{align*} \]
A Sampling Argument

\[ T = \{ \text{\color{green}{\square}} \}, \text{ set of entries in } S \text{ with same value} \]
A Sampling Argument

\[ T = \{ \square \}, \text{ set of entries in } S \text{ with same value} \]
A Sampling Argument

$T = \{ \text{\textcolor{green}{\textbullet}} \}$, set of entries in $S$ with same value

Choose $M_i$ proportional to total value on $T$
A Sampling Argument

$T = \{ \square \}$, set of entries in $S$ with same value

Choose $M_i$ proportional to total value on $T$
A Sampling Argument

\[ T = \{ \square \} \text{, set of entries in } S \text{ with same value} \]

Choose \( M_i \) proportional to total value on \( T \)
Choose (\( a, b \)) in \( T \) proportional to relative value in \( M_i \)
A Sampling Argument

\[ T = \{\text{\cellcolor{green}}\}, \text{ set of entries in } S \text{ with same value} \]

Choose \( M_i \) proportional to total value on \( T \)

Choose \((a,b)\) in \( T \) proportional to relative value in \( M_i \)
A Sampling Argument

\[ T = \{ \square \} \text{, set of entries in } S \text{ with same value} \]

Choose \( M_i \) proportional to total value on \( T \)
Choose \( (a,b) \) in \( T \) proportional to relative value in \( M_i \)
A Sampling Argument

$T = \{\square\}$, set of entries in $S$ with same value

Choose $M_i$ proportional to total value on $T$
Choose $(a,b)$ in $T$ proportional to relative value in $M_i$

This outputs a uniformly random sample from $T$
A Sampling Argument

T = {□}, set of entries in S with same value

Choose \( M_i \) proportional to total value on T
Choose \( (a,b) \) in T proportional to relative value in \( M_i \)
A Sampling Argument

\[ T = \{\square\}, \text{ set of entries in } S \text{ with same value} \]

Choose \( M_i \) proportional to total value on \( T \)

Choose \((a,b)\) in \( T \) proportional to relative value in \( M_i \)

If \( r \) is too small, this procedure uses too little entropy!
Outline

Part I: Tools for Extended Formulations
- Yannakakis’s Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications
- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
The Construction of [Fiorini et al] correlation polytope: $P_{\text{corr}} = \text{conv}\{aa^T|a \text{ in } \{0,1\}^n \}$
The Construction of [Fiorini et al]

correlation polytope: $P_{\text{corr}} = \text{conv}\{aa^T|a \text{ in } \{0,1\}^n \}$

vertices:

constraints:

S
The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \in \{0,1\}^n \}$

vertices: $a \in \{0,1\}^n$

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

S
The Construction of [Fiorini et al]

**correlation polytope:** $P_{\text{corr}} = \text{conv}\{aa^T|a \in \{0,1\}^n \}$

**vertices:** $a \in \{0,1\}^n$

**constraints:**

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

$(1-a^Tb)^2$
The Construction of [Fiorini et al]

correlation polytope: \( P_{\text{corr}} = \text{conv}\{aa^T|a \in \{0,1\}^n \} \)

constraints:
\( b \in \{0,1\}^n \)

(1-\( a^Tb \))^2

vertices: \( a \in \{0,1\}^n \)

**UNIQUE DISJ.**
Output ‘YES’ if \( a \) and \( b \) as sets are disjoint, and ‘NO’ if \( a \) and \( b \) have one index in common
The Construction of [Fiorini et al]

correlation polytope: $P_{\text{corr}} = \text{conv}\{aa^T | a \in \{0,1\}^n \}$
The Construction of [Fiorini et al]

correlation polytope: $P_{\text{corr}} = \text{conv}\{aa^T | a \in \{0,1\}^n \}$

Why is that (a sub-matrix of) the slack matrix?
The Construction of [Fiorini et al]

correlation polytope: \( P_{\text{corr}} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n \} \)

Why is that (a sub-matrix of) the slack matrix?

\[(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2\]
The Construction of [Fiorini et al]

**correlation polytope:** $P_{\text{corr}} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n \}$

Why is that (a sub-matrix of) the slack matrix?

\[
(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2
= 1 - 2\langle\text{diag}(b),aa^T\rangle + \langle bb^T,aa^T\rangle
\]
The Construction of [Fiorini et al]

correlation polytope: \( P_{\text{corr}} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n \} \)

Why is that (a sub-matrix of) the slack matrix?

\[
(1-a^T b)^2 = 1 - 2a^T b + (a^T b)^2 \\
= 1 - 2 \langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle \\
1 \geq \langle 2\text{diag}(b) - bb^T, aa^T \rangle
\]
The Construction of [Fiorini et al]

correlation polytope: \( P_{\text{corr}} = \text{conv}\{aa^T|a \text{ in } \{0,1\}^n\} \)

Why is that (a sub-matrix of) the slack matrix?

\[
(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2 \\
= 1 - 2 \langle \text{diag}(b),aa^T \rangle + \langle bb^T,aa^T \rangle \\
\geq \langle 2\text{diag}(b) - bb^T,aa^T \rangle
\]

What is the slack?
The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \in \{0,1\}^n \}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^T b)^2 = 1 - 2a^T b + (a^T b)^2$$

$$= 1 - 2 \langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle$$

$$1 \geq \langle 2\text{diag}(b) - bb^T, aa^T \rangle$$

What is the slack? $$(1-a^T b)^2$$
A Hard Distribution
A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$
A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b} = (1-a^T b)^2$, so $S_{a,b} = 1$ for all pairs in $T$
Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b} = (1-a^T b)^2$, so $S_{a,b} = 1$ for all pairs in $T$

How does the sampling procedure specialize to this case? (Recall it generates $(a,b)$ unif. from $T$)
A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b} = (1-a^T b)^2$, so $S_{a,b} = 1$ for all pairs in $T$

How does the sampling procedure specialize to this case? (Recall it generates $(a,b)$ unif. from $T$)

Sampling Procedure:
A Hard Distribution

Let \( T = \{(a,b) \mid a^T b = 0\}, \) \(|T| = 3^n\)

Recall: \( S_{a,b} = (1-a^T b)^2\), so \( S_{a,b} = 1\) for all pairs in \( T\)

How does the sampling procedure specialize to this case? (Recall it generates \((a,b)\) unif. from \( T\))

**Sampling Procedure:**

- Let \( R_i \) be the sum of \( M_i(a,b) \) over \((a,b)\) in \( T\) and let \( R \) be the sum of \( R_i\)
A Hard Distribution

Let \( T = \{(a,b) \mid a^T b = 0\}, |T| = 3^n \)

Recall: \( S_{a,b} = (1-a^T b)^2 \), so \( S_{a,b} = 1 \) for all pairs in \( T \)

How does the sampling procedure specialize to this case? (Recall it generates \((a,b)\) unif. from \( T\))

**Sampling Procedure:**

- Let \( R_i \) be the sum of \( M_i(a,b) \) over \((a,b)\) in \( T \) and let \( R \) be the sum of \( R_i \)
- Choose \( i \) with probability \( R_i / R \)
A Hard Distribution

Let $T = \{(a,b) | a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b} = (1 - a^T b)^2$, so $S_{a,b} = 1$ for all pairs in $T$

How does the sampling procedure specialize to this case? (Recall it generates $(a,b)$ unif. from $T$)

Sampling Procedure:

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$
- Choose $i$ with probability $R_i/R$
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$
Entropy Accounting 101
Entropy Accounting 101

Sampling Procedure:

• Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$

• Choose $i$ with probability $R_i/R$

• Choose $(a,b)$ with probability $M_i(a,b)/R_i$
Entropy Accounting 101

Sampling Procedure:

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$
- Choose $i$ with probability $R_i/R$
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$

Total Entropy:

$n \log_2 3 \leq$
Entropy Accounting 101

Sampling Procedure:

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$
- Choose $i$ with probability $R_i/R$
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \text{choose } i + \text{choose } (a,b) \text{ conditioned on } i$$
### Entropy Accounting 101

**Sampling Procedure:**

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$.
- Choose $i$ with probability $R_i/R$.
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$.

**Total Entropy:**

$$n \log_2 3 \leq \log_2 r + \text{choose } (a,b) \text{ conditioned on } i$$
Sampling Procedure:

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$
- Choose $i$ with probability $R_i/R$
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \log_2 r + (1-\delta)n \log_2 3$$
Sampling Procedure:

- Let $R_i$ be the sum of $M_i(a,b)$ over $(a,b)$ in $T$ and let $R$ be the sum of $R_i$
- Choose $i$ with probability $R_i/R$
- Choose $(a,b)$ with probability $M_i(a,b)/R_i$

Total Entropy:

\[
\begin{align*}
\text{choose } i & \quad \log_2 r \\
\text{choose } (a,b) \text{ conditioned on } i & \quad (1-\delta)n \log_2 3 \\
\end{align*}
\]

\[
\begin{align*}
n \log_2 3 & \leq \log_2 r + (1-\delta)n \log_2 3 \quad (?)
\end{align*}
\]
Suppose that $a_j$ and $b_j$ are fixed
Suppose that $a_j$ and $b_j$ are fixed

$M_i$ restricted to $(a_j, b_j)$
Suppose that $a_j$ and $b_j$ are fixed

$M_i$ restricted to $(a_j, b_j)$

$\begin{array}{c}
(a_{1..j-1}, a_j = 0, a_{j+1}...n) \\
(a_{1..j-1}, a_j = 1, a_{j+1}...n)
\end{array}$

$\begin{array}{cc}
M_i(a, b) & M_i(a, b) \\
M_i(a, b) & M_i(a, b)
\end{array}$
\( (a_{1..j-1}, a_j = 0, a_{j+1}...n) \quad M_i(a,b) \quad M_i(a,b) \)

\( (a_{1..j-1}, a_j = 1, a_{j+1}...n) \quad M_i(a,b) \quad M_i(a,b) \)

\( (\ldots b_j = 0 \ldots) \quad (\ldots b_j = 1 \ldots) \)
If \( a_j = 1, \ b_j = 1 \) then \( a^\top b = 1 \), hence \( M_i(a, b) = 0 \).
If \( a_j = 1 \), \( b_j = 1 \) then \( a^T b = 1 \), hence \( M_i(a, b) = 0 \)

\[
\begin{array}{c|c}
(a_1..j-1, a_j = 0, a_{j+1}..n) & M_i(a, b) & M_i(a, b) \\
(a_1..j-1, a_j = 1, a_{j+1}..n) & M_i(a, b) & \text{zero}
\end{array}
\]
If $a_j = 1$, $b_j = 1$ then $a^T b = 1$, hence $M_i(a,b) = 0$

But $\text{rank}(M_i) = 1$, hence there must be another zero in either the same row or column
If $a_j=1$, $b_j=1$ then $a^\top b = 1$, hence $M_i(a,b) = 0$

But rank($M_i$)=1, hence there must be another zero in either the same row or column

<table>
<thead>
<tr>
<th>$(a_{1..j-1}, a_j=0, a_{j+1}...n)$</th>
<th>$M_i(a,b)$</th>
<th>$M_i(a,b)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(a_{1..j-1}, a_j=1, a_{j+1}...n)$</td>
<td>zero</td>
<td>zero</td>
</tr>
</tbody>
</table>

(...$b_j=0$...)  (...$b_j=1$...)
If \( a_j=1, \ b_j=1 \) then \( a^\top b = 1 \), hence \( M_i(a,b) = 0 \)

But \( \text{rank}(M_i) = 1 \), hence there must be another zero in either the same row or column

\[ H(a_j, b_j | i, a_{-j}, b_{-j}) \leq 1 < \log_2 3 \]

\[ \begin{array}{c|cc}
(a_1..j-1, a_j=0, a_{j+1}...n) & M_i(a,b) & M_i(a,b) \\
\hline
(a_1..j-1, a_j=1, a_{j+1}...n) & \text{zero} & \text{zero} \\
\end{array} \]
Entropy Accounting 101

Generate uniformly random \((a,b)\) in \(T\):

- Let \(R_i\) be the sum of \(M_i(a,b)\) over \((a,b)\) in \(T\) and let \(R\) be the sum of \(R_i\).
- Choose \(i\) with probability \(R_i/R\).
- Choose \((a,b)\) with probability \(M_i(a,b)/R_i\).

Total Entropy:

\[
n \log_2 3 \leq \log_2 r + \text{choose } (a,b) \text{ conditioned on } i
\]
### Entropy Accounting 101

**Generate uniformly random \((a,b)\) in \(T:\)**

- Let \(R_i\) be the sum of \(M_i(a,b)\) over \((a,b)\) in \(T\) and let \(R\) be the sum of \(R_i\)
- Choose \(i\) with probability \(R_i/R\)
- Choose \((a,b)\) with probability \(M_i(a,b)/R_i\)

**Total Entropy:**

\[
\begin{align*}
\text{choose } i & \quad \text{choose } (a,b) \\
\log_2 3 & \leq \log_2 r + n
\end{align*}
\]
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
Approximate EFs [Braun et al]

vertices: \( a \) in \( \{0, 1\}^n \)

constraints:
\( b \) in \( \{0, 1\}^n \)

\( (1-a^T b)^2 \)
Approximate EFs [Braun et al]

Is there a $K$ (with small $x_c$) s.t. $P_{corr} \subseteq K \subseteq (C+1)P_{corr}$?

vertices: $a$ in $\{0,1\}^n$

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

$(1-a^Tb)^2$
Approximate EFs [Braun et al]

Is there a $K$ (with small $xc$) s.t. $P_{\text{corr}} \subseteq K \subseteq (C+1)P_{\text{corr}}$?

vertices: $a$ in $\{0,1\}^n$

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

$(1 - a^T b)^2 + C$
Approximate EFs [Braun et al]

Is there a $K$ (with small $xc$) s.t. $P_{corr} \subset K \subset (C+1)P_{corr}$?

vertices: $a \in \{0,1\}^n$

constraints:

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

$(1-a^T b)^2 + C$

New Goal:
Output the answer to UDISJ with prob. at least $\frac{1}{2} + \frac{1}{2}(C+1)$
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s.
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

**Claim:** If UDISJ can be computed with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ using $o(n/C^2)$ bits, then UDISJ can be computed with prob. $\frac{3}{4}$ using $o(n)$ bits.

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s.
Is the correlation polytope hard to approximate for large values of \( C \)?

**Analogy:** Is UDISJ hard to compute with prob. \( \frac{1}{2} + \frac{1}{2}(C+1) \) for large values of \( C \)?

There is a natural barrier at \( C = \sqrt{n} \) for proving l.b.s: 

**Claim:** If UDISJ can be computed with prob. \( \frac{1}{2} + \frac{1}{2}(C+1) \) using \( o(n/C^2) \) bits, then UDISJ can be computed with prob. \( \frac{3}{4} \) using \( o(n) \) bits.

**Proof:** Run the protocol \( O(C^2) \) times and take the majority vote.
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s.
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

**Corollary [from K-S]:** Computing UDISJ with probability $\frac{1}{2} + \frac{1}{2}(C+1)$ requires $\Omega(n/C^2)$ bits
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

**Corollary [from K-S]:** Computing UDISJ with probability $\frac{1}{2} + \frac{1}{2}(C+1)$ requires $\Omega(n/C^2)$ bits

**Theorem [B-M]:** Computing UDISJ with probability $\frac{1}{2} + \frac{1}{2}(C+1)$ requires $\Omega(n/C)$ bits
Is the correlation polytope hard to approximate for large values of $C$?

**Analogy:** Is UDISJ hard to compute with prob. $\frac{1}{2} + \frac{1}{2}(C+1)$ for large values of $C$?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

**Theorem [B-M]:** Any EF that approximates clique within $n^{1-\epsilon}$ has size $\exp(n^{\epsilon})$

**Theorem [B-M]:** Computing UDISJ with probability $\frac{1}{2} + \frac{1}{2}(C+1)$ requires $\Omega(n/C)$ bits
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
Outline

Part I: Tools for Extended Formulations
  • Yannakakis’s Factorization Theorem
  • The Rectangle Bound
  • A Sampling Argument

Part II: Applications
  • Correlation Polytope
  • Approximating the Correlation Polytope
  • Matching Polytope
The Matching Polytope [Edmonds]

\[ P_{PM} = \text{conv}\{1_M | M \text{ is a perfect matching in } K_n\} \]
The Matching Polytope [Edmonds]

\[ P_{PM} = \text{conv}\{1_M | M \text{ is a perfect matching in } K_n\} \]
The Matching Polytope [Edmonds]

\[ P_{PM} = \text{conv}\{1_M | M \text{ is a perfect matching in } K_n\} \]

vertices: \(1_M\)

countstraints:
\[ U \subset [n] \]
with \(|U| = \text{odd}\)

\[ |\delta(U) \cap M| - 1 \]
The Matching Polytope [Edmonds]

\[ P_{PM} = \text{conv}\{1_M \mid M \text{ is a perfect matching in } K_n\} \]

- Vertices: \(1_M\)
- Constraints:
  - \(U \subseteq [n]\)
  - \(|U| = \text{odd}\)
  - \(|\delta(U) \cap M| - 1\)

Is there a small rectangle covering?
The Matching Polytope [Edmonds]

$$P_{PM} = \text{conv}\{1_M | M \text{ is a perfect matching in } K_n\}$$

Constraints:

- $U \subseteq [n]$ with $|U| = \text{odd}$
- $|\delta(U) \cap M| - 1$

Vertices: $1_M$

Is there a small rectangle covering?

Yes! Just guess two edges in $M$, crossing the cut
Hyperplane Separation Lemma

[Rothvoss] attributed to [Fiorini]:
Hyperplane Separation Lemma

[Rothvoss] attributed to [Fiorini]:

**Lemma:** For slack matrix $S$, any matrix $W$:

$$\text{rank}^+(S) \geq \frac{\langle S, W \rangle}{\|S\|_\infty \alpha}$$

where $\alpha = \max \langle W, R \rangle$ s.t. $R$ is rank one, entries in $[0,1]$
Hyperplane Separation Lemma

[Rothevoss] attributed to [Fiorini]:

**Lemma:** For slack matrix $S$, any matrix $W$:

$$\text{rank}^+(S) \geq \frac{\langle S, W \rangle}{\|S\|_\infty \alpha}$$

where $\alpha = \max \langle W, R \rangle$ s.t. $R$ is rank one, entries in $[0,1]$

**Proof:**

$$\langle W, S \rangle = \sum \|R_i\|_\infty \langle W, R_i/\|R_i\|_\infty \rangle \leq \alpha \|R\|_\infty \|S\|_\infty$$
Theorem [Rothvoss ’13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)
Theorem [Rothvoss ’13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose $W$?
Theorem [Rothvoss ’13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose $W$?

$$W_{U,M} = \begin{cases} -\infty & \text{if } |\delta(U) \cap M| = 1 \\ 1/Q_3 & \text{if } |\delta(U) \cap M| = 3 \\ -1/Q_k & \text{if } |\delta(U) \cap M| = k \\ 0 & \text{else} \end{cases}$$
Theorem [Rothvoss ’13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose $W$?

$$W_{U,M} = \begin{cases} 
-\infty & \text{if } |\delta(U) \cap M| = 1 \\
1/Q_3 & \text{if } |\delta(U) \cap M| = 3 \\
-1/Q_k & \text{if } |\delta(U) \cap M| = k \\
0 & \text{else}
\end{cases}$$

Proof is a substantial modification to Razborov’s rectangle corruption lemma
Summary:

- Extended formulations and Yannakakis’ factorization theorem
Summary:

- Extended formulations and Yannakakis’ factorization theorem

- **Lower bound techniques**: rectangle bound, information complexity, hyperplane separation
Summary:

- Extended formulations and Yannakakis’ factorization theorem
- **Lower bound techniques**: rectangle bound, information complexity, hyperplane separation
- **Applications**: connections between correlation polytope and disjointness,
Summary:

- Extended formulations and Yannakakis’ factorization theorem
- **Lower bound techniques**: rectangle bound, information complexity, hyperplane separation
- **Applications**: connections between correlation polytope and disjointness,
- **Open question**: Can we prove lower bounds against general SDPs?
Any Questions?

Summary:

- Extended formulations and Yannakakis’ factorization theorem
- **Lower bound techniques**: rectangle bound, information complexity, hyperplane separation
- **Applications**: connections between correlation polytope and disjointness,
- **Open question**: Can we prove lower bounds against general SDPs?
Thanks!