On the existence of 0/1 polytopes with high semidefinite extension complexity

Daniel Dadush

Centrum Wiskunde & Informatica (CWI)

Joint Work with...

My coauthors:

- Sebastian Pokutta (Georgia Tech)
- Jop Briët (CWI)

What is the expressive power of linear / semidefinite programming?

What is the expressive power of linear / semidefinite programming?

For convex hulls such as
Matchings, Hamiltonian cycles, Graph Cuts, ...
what is the *smallest* linear / semidefinite program whose feasible region captures them (even approximately)?

What is the expressive power of linear / semidefinite programming?

For convex hulls such as
Matchings, Hamiltonian cycles, Graph Cuts, ...
what is the *smallest* linear / semidefinite program whose feasible region captures them (even approximately)?

Alternative measure of complexity independent from P vs NP.

Lower bounds on the size of *Linear Programs*:

1. Any symmetric LP that captures the TSP or Matching polytope must have size $2^{\Omega(n)}$ [**Yannakakis** '91]
2. There exists a convex hull of $0/1$ points that cannot be captured by an LP of size less than $2^{\Omega(n)}$ [**Rothvoss** '11]
3. Any LP that captures the TSP polytope must have size $2^{\Omega(n^{1/2})}$ [**Fiorini, Massar, Pokutta, Tiwary, de Wolf** '12]
4. Any LP that ρ-approximates the Correlation polytope must have size $2^{\Omega(n/\rho)}$ [**Braun, Fiornini, Pokutta, Steurer** '12, **Braverman, Moitra** '13, **Pokutta, Braun** '13]
5. Any LP of relaxation of size n^r for the Correlation polytope has integrality gap at least as large as $O(r)$ levels of Sherali-Adams [**Chan, Lee, Raghavendra, Steurer** '13]

Lower bounds on the size of *Linear Programs*:

1. Any symmetric LP that captures the TSP or Matching polytope must have size $2^{\Omega(n)}$ [**Yannakakis** '91]
2. There exists a convex hull of $0/1$ points that cannot be captured by an LP of size less than $2^{\Omega(n)}$ [**Rothvoss** '11]
3. Any LP that captures the TSP polytope must have size $2^{\Omega(n^{1/2})}$ [**Fiorini, Massar, Pokutta, Tiwary, de Wolf** '12]
4. Any LP that ρ-approximates the Correlation polytope must have size $2^{\Omega(n/\rho)}$ [**Braun, Fiornini, Pokutta, Steurer** '12, **Braverman, Moitra** '13, **Pokutta, Braun** '13]
5. Any LP of relaxation of size n^r for the Correlation polytope has integrality gap at least as large as $O(r)$ levels of Sherali-Adams [**Chan, Lee, Raghavendra, Steurer** '13]

Lower bounds on the size of *Linear Programs*:

1. Any symmetric LP that captures the TSP or Matching polytope must have size $2^{\Omega(n)}$ [**Yannakakis** '91]
2. There exists a convex hull of $0/1$ points that cannot be captured by an LP of size less than $2^{\Omega(n)}$ [**Rothvoss** '11]
3. Any LP that captures the TSP polytope must have size $2^{\Omega(n^{1/2})}$ [**Fiorini, Massar, Pokutta, Tiwary, de Wolf** '12]
4. Any LP that ρ-approximates the Correlation polytope must have size $2^{\Omega(n/\rho)}$ [**Braun, Fiornini, Pokutta, Steurer** '12, **Braverman, Moitra** '13, **Pokutta, Braun** '13]
5. Any LP of relaxation of size n^r for the Correlation polytope has integrality gap at least as large as $O(r)$ levels of Sherali-Adams [**Chan, Lee, Raghavendra, Steurer** '13]

Lower bounds on the size of *Linear Programs*:

1. Any symmetric LP that captures the TSP or Matching polytope must have size $2^{\Omega(n)}$ [**Yannakakis** '91]
2. There exists a convex hull of $0/1$ points that cannot be captured by an LP of size less than $2^{\Omega(n)}$ [**Rothvoss** '11]
3. Any LP that captures the TSP polytope must have size $2^{\Omega(n^{1/2})}$ [**Fiorini, Massar, Pokutta, Tiwary, de Wolf** '12]
4. Any LP that ρ-approximates the Correlation polytope must have size $2^{\Omega(n/\rho)}$ [**Braun, Fiornini, Pokutta, Steurer** '12, **Braverman, Moitra** '13, **Pokutta, Braun** '13]
5. Any LP of relaxation of size n^r for the Correlation polytope has integrality gap at least as large as $O(r)$ levels of Sherali-Adams [**Chan, Lee, Raghavendra, Steurer** '13]

Lower bounds on the size of *Linear Programs*:

1. Any symmetric LP that captures the TSP or Matching polytope must have size $2^{\Omega(n)}$ [**Yannakakis** '91]
2. There exists a convex hull of $0/1$ points that cannot be captured by an LP of size less than $2^{\Omega(n)}$ [**Rothvoss** '11]
3. Any LP that captures the TSP polytope must have size $2^{\Omega(n^{1/2})}$ [**Fiorini, Massar, Pokutta, Tiwary, de Wolf** '12]
4. Any LP that ρ-approximates the Correlation polytope must have size $2^{\Omega(n/\rho)}$ [**Braun, Fiornini, Pokutta, Steurer** '12, **Braverman, Moitra** '13, **Pokutta, Braun** '13]
5. Any LP of relaxation of size n^r for the Correlation polytope has integrality gap at least as large as $O(r)$ levels of Sherali-Adams [**Chan, Lee, Raghavendra, Steurer** '13]

What about *Semidefinite Programs*?

Theorem ([Briët, D., Pokutta '13])

There exists a convex hull of $0/1$ points that cannot be captured by an SDP of size less than $2^{\Omega(n)}$.

Extend framework of Rothvoss to SDP setting.
Baby step towards understanding lower bounds for SDPs.

What about *Semidefinite Programs*?

Theorem ([Briët, D., Pokutta '13])

There exists a convex hull of $0/1$ *points that cannot be captured by an SDP of size less than* $2^{\Omega(n)}$.

Extend framework of Rothvoss to SDP setting.
Baby step towards understanding lower bounds for SDPs.

What about *Semidefinite Programs*?

Theorem ([Briët, D., Pokutta '13])

There exists a convex hull of $0/1$ points that cannot be captured by an SDP of size less than $2^{\Omega(n)}$.

Extend framework of Rothvoss to SDP setting.

Baby step towards understanding lower bounds for SDPs.

What about *Semidefinite Programs*?

Theorem ([Briët, D., Pokutta '13])

There exists a convex hull of $0/1$ *points that cannot be captured by an SDP of size less than* $2^{\Omega(n)}$*.*

Extend framework of Rothvoss to SDP setting.
Baby step towards understanding lower bounds for SDPs.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Question

Can we reduce the number of inequalities needed to define P by adding auxilliary variables?

Definition (Linear Extension)

$Q = \{(z, y) : Cz + Dy = d, y \geq 0, y \in \mathbb{R}^r, z \in \mathbb{R}^l\}$, is a linear extension of P of size r if $\exists\ \pi : \mathbb{R}^{l+r} \to \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (Linear Extension Complexity)

$\mathrm{xc}(P) :=$ minimum size of any linear extension of P.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Question

Can we reduce the number of inequalities needed to define P by adding auxilliary variables?

Definition (Linear Extension)

$Q = \{(z, y) : Cz + Dy = d, y \geq 0, y \in \mathbb{R}^r, z \in \mathbb{R}^l\}$, is a linear extension of P of size r if $\exists\ \pi : \mathbb{R}^{l+r} \rightarrow \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (Linear Extension Complexity)

$\mathrm{xc}(P) :=$ minimum size of any linear extension of P.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Question

Can we reduce the number of inequalities needed to define P by adding auxilliary variables?

Definition (Linear Extension)

$Q = \{(z, y) : Cz + Dy = d, y \geq 0, y \in \mathbb{R}^r, z \in \mathbb{R}^l\}$, is a linear extension of P of size r if $\exists\ \pi : \mathbb{R}^{l+r} \rightarrow \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (Linear Extension Complexity)

$\mathrm{xc}(P) :=$ minimum size of any linear extension of P.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Question

Can we reduce the number of inequalities needed to define P by adding auxilliary variables?

Definition (Linear Extension)

$Q = \{(z, y) : Cz + Dy = d, y \geq 0, y \in \mathbb{R}^r, z \in \mathbb{R}^l\}$, is a linear extension of P of size r if $\exists\ \pi : \mathbb{R}^{l+r} \to \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (Linear Extension Complexity)

$\mathrm{xc}(P) :=$ minimum size of any linear extension of P.

Extension Complexity.

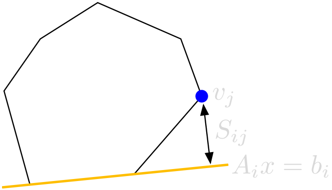
Some known results (constructions & lower bounds):

- $\text{xc}($regular n-gon$) = \Theta(\log n)$ [Ben-Tal, Nemirovski'01]
- $\text{xc}($generic n-gon$) = \Omega(\sqrt{n})$ [Fironi, Rothvoss, Tiwary'11]
- $\text{xc}(n$-permutahedron$) = \Theta(n \log n)$ [Goemans'09]
- $\text{xc}($spanning tree polytope of $K_n) = O(n^3)$ [Kipp-Martin'87]
- $\text{xc}($spanning tree polytope of planar graph $G) = \Theta(n)$ [Williams'01]
- $\text{xc}($stable set polytope of perfect graph $G) = n^{O(\log n)}$ [Yannakakis'91]
- ...

Slack Matrices.

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\mathcal{V} = \{v_1, \ldots, v_N\} \subseteq \mathbb{R}^n$ s.t.

$$P = \{x \in \mathbb{R}^n \mid Ax \leqslant b\} = \text{conv}(\mathcal{V})$$



Definition (slack matrix)

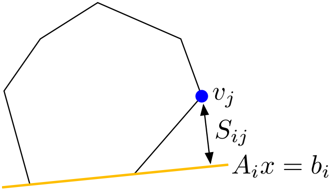
Slack matrix $S \in \mathbb{R}_+^{m \times N}$ of P (w.r.t. $Ax \leqslant b$ and $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\ i \in [m], j \in [N]$$

Slack Matrices.

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\mathcal{V} = \{v_1, \ldots, v_N\} \subseteq \mathbb{R}^n$ s.t.

$$P = \{x \in \mathbb{R}^n \mid Ax \leqslant b\} = \mathrm{conv}(\mathcal{V})$$



Definition (slack matrix)

Slack matrix $S \in \mathbb{R}_+^{m\times N}$ of P (w.r.t. $Ax \leqslant b$ and $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\ i \in [m], j \in [N]$$

Nonnegative Factorizations and Extensions.

Definition (slack matrix)

Slack matrix $S \in \mathbb{R}_{+}^{m \times N}$ of P (w.r.t. $Ax \leqslant b$ and $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\ i \in [m], j \in [N]$$

Definition

A rank-r nonnegative factorization of $S \in \mathbb{R}_{+}^{m \times n}$ is

$$S = UV \quad \text{where} \quad U \in \mathbb{R}_{+}^{m \times r} \quad \text{and} \quad V \in \mathbb{R}_{+}^{r \times n}$$

Proposition (Extensions from Factorizations)

$Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$
is a linear extension of P of size r.

Nonnegative Factorizations and Extensions.

Definition (slack matrix)

Slack matrix $S \in \mathbb{R}_{+}^{m \times N}$ of P (w.r.t. $Ax \leqslant b$ and $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\ i \in [m], j \in [N]$$

Definition

A rank-r nonnegative factorization of $S \in \mathbb{R}_{+}^{m \times n}$ is

$$S = UV \quad \text{where} \quad U \in \mathbb{R}_{+}^{m \times r} \quad \text{and} \quad V \in \mathbb{R}_{+}^{r \times n}$$

Proposition (Extensions from Factorizations)

$Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$
is a linear extension of P of size r.

Nonnegative Factorizations and Extensions.

Definition (Nonnegative Rank)

$$\mathrm{rk}_+(S) := \min\{r \mid \exists \text{ rank-}r \text{ nonnegative factorization of } S\}$$

Theorem (Factorization Theorem [Yannakakis'91])

For every slack matrix S of P:

$$\mathrm{xc}(P) = \mathrm{rk}_+(S)$$

PSD Matrices.

Definition (PSD matrix)

A matrix $U \in \mathbb{R}^{r \times r}$ is PSD if U is symmetric and

$$x^\mathsf{T} U x \geq 0 \quad \forall x \in \mathbb{R}^r.$$

Let $\mathbb{S}_+^r$ denote the set of $r \times r$ PSD matrices.

Definition (Spectral Decomposition)

U is $r \times r$ PSD iff U admits a *Spectral Decomposition*

$$U = \sum_{i=1}^{r} \lambda_i u_i u_i^\mathsf{T},$$

$\lambda_1, \ldots, \lambda_r \geq 0$, $u_1, \ldots, u_r$ an orthonormal basis.

PSD Matrices.

Definition (PSD matrix)

A matrix $U \in \mathbb{R}^{r \times r}$ is PSD if U is symmetric and

$$x^\mathsf{T} U x \geq 0 \quad \forall x \in \mathbb{R}^r.$$

Let $\mathbb{S}^r_+$ denote the set of $r \times r$ PSD matrices.

Definition (Spectral Decomposition)

U is $r \times r$ PSD iff U admits a *Spectral Decomposition*

$$U = \sum_{i=1}^{r} \lambda_i u_i u_i^\mathsf{T},$$

$\lambda_1, \ldots, \lambda_r \geq 0$, $u_1, \ldots, u_r$ an orthonormal basis.

PSD Matrices.

Definition (Operator norm)

For a matrix $T \in \mathbb{R}^{r \times r}$ the operator norm of T is

$$\|T\|_{\text{op}} = \max_{\|x\|_2 = 1} \|Tx\|_2$$

For a PSD matrix $U \in \mathbb{R}^{r \times r}$

$$\|U\|_{\text{op}} = \max_{\|x\|_2 = 1} x^{\mathsf{T}} U x = \text{ largest eigenvalue of } U.$$

PSD Matrices.

Definition (Trace)

For a matrix $T \in \mathbb{R}^{r \times r}$, we define $\mathrm{Tr}[T] = \sum_{i=1}^{r} T_{ii}$.

Remark (Trace Inner Product)

For $A, B \in \mathbb{R}^{r \times r}$ symmetric, $\mathrm{Tr}[AB] = \sum_{i,j \in [r]} A_{ij} B_{ij}$.

Proposition

For PSD matrices $U, V \in \mathbb{S}_+^r$,

$$\mathrm{Tr}[UV] = \sum_{i,j \in [r]} \lambda_i \gamma_j \langle u_i, v_j \rangle^2 \geq 0 \ ,$$

where $U = \sum_{i=1}^{r} \lambda_i u_i u_i^\mathsf{T}$ and $V = \sum_{j=1}^{r} \gamma_j v_j v_j^\mathsf{T}$ are the respective spectral decompositions.

PSD Matrices.

Definition (Trace)

For a matrix $T \in \mathbb{R}^{r\times r}$, we define $\mathrm{Tr}\left[T\right] = \sum_{i=1}^{r} T_{ii}$.

Remark (Trace Inner Product)

For $A, B \in \mathbb{R}^{r\times r}$ *symmetric,* $\mathrm{Tr}\left[AB\right] = \sum_{i,j\in[r]} A_{ij}B_{ij}$.

Proposition

For PSD matrices $U, V \in \mathbb{S}_+^r$,

$$\mathrm{Tr}\left[UV\right] = \sum_{i,j\in[r]} \lambda_i\gamma_j \langle u_i, v_j\rangle^2 \geq 0 \ ,$$

where $U = \sum_{i=1}^{r} \lambda_i u_i u_i^{\mathsf{T}}$ *and* $V = \sum_{j=1}^{r} \gamma_j v_j v_j^{\mathsf{T}}$ *are the respective spectral decompositions.*

PSD Matrices.

Definition (Trace)

For a matrix $T \in \mathbb{R}^{r \times r}$, we define $\mathrm{Tr}[T] = \sum_{i=1}^{r} T_{ii}$.

Remark (Trace Inner Product)

For $A, B \in \mathbb{R}^{r \times r}$ *symmetric,* $\mathrm{Tr}[AB] = \sum_{i,j \in [r]} A_{ij} B_{ij}$.

Proposition

For PSD matrices $U, V \in \mathbb{S}_{+}^{r}$,

$$\mathrm{Tr}[UV] = \sum_{i,j \in [r]} \lambda_i \gamma_j \langle u_i, v_j \rangle^2 \geq 0 \ ,$$

where $U = \sum_{i=1}^{r} \lambda_i u_i u_i^{\mathsf{T}}$ *and* $V = \sum_{j=1}^{r} \gamma_j v_j v_j^{\mathsf{T}}$ *are the respective spectral decompositions.*

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Definition (PSD Extension)

$Q = \{(z, Y) : C_i z + \mathrm{Tr}[D_i Y] = d_i, \forall\, i \in [l], Y \in \mathbb{S}^r_+, z \in \mathbb{R}^l\}$, is a PSD extension of P of size r if $\exists\, \pi : \mathbb{R}^l \times \mathbb{S}^r_+ \to \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (PSD Extension Complexity)

$\mathrm{xc}_{psd}(P) :=$ minimum size of any PSD extension of P.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Definition (PSD Extension)

$Q = \{(z, Y) : C_i z + \mathrm{Tr}\,[D_i Y] = d_i, \forall\ i \in [l], Y \in \mathbb{S}^r_+, z \in \mathbb{R}^l\}$, is a PSD extension of P of size r if $\exists\ \pi : \mathbb{R}^l \times \mathbb{S}^r_+ \to \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (PSD Extension Complexity)

$\mathrm{xc}_{psd}(P) :=$ minimum size of any PSD extension of P.

Polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with m facets.

Definition (PSD Extension)

$Q = \{(z, Y) : C_i z + \mathrm{Tr}\,[D_i Y] = d_i, \forall\ i \in [l], Y \in \mathbb{S}^r_+, z \in \mathbb{R}^l\}$, is a PSD extension of P of size r if $\exists\ \pi : \mathbb{R}^l \times \mathbb{S}^r_+ \rightarrow \mathbb{R}^n$ such that

$$P = \pi(Q).$$

Definition (PSD Extension Complexity)

$\mathrm{xc}_{psd}(P) :=$ minimum size of any PSD extension of P.

PSD Factorizations and Extensions

Definition (Slack Matrix)

Slack matrix $S \in \mathbb{R}_+^{m \times N}$ of P (w.r.t. $Ax \leqslant b$ and vertices $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\ i \in [m], j \in [N]$$

Definition

A rank-r PSD factorization of $S \in \mathbb{R}_+^{m \times N}$ is

$S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ where U_i, V_j are $r \times r$ PSD $\forall\ i \in [m], j \in [N]$.

Proposition (Extensions from Factorizations)

$Q = \{(x, Y) : A_i x + \mathrm{Tr}\left[U_i Y\right] = b_i, \forall i \in [m], Y \in \mathbb{S}_+^r\}$
is a PSD extension of P of size r.

PSD Factorizations and Extensions

Definition (Slack Matrix)

Slack matrix $S \in \mathbb{R}_+^{m \times N}$ of P (w.r.t. $Ax \leqslant b$ and vertices $\mathcal{V}$):

$$S_{ij} := b_i - A_i v_j, \forall\, i \in [m], j \in [N]$$

Definition

A rank-r PSD factorization of $S \in \mathbb{R}_+^{m \times N}$ is

$S_{ij} = \mathrm{Tr}\,[U_i V_j]$ where U_i, V_j are $r \times r$ PSD $\forall\, i \in [m], j \in [N]$.

Proposition (Extensions from Factorizations)

$Q = \{(x, Y) : A_i x + \mathrm{Tr}\,[U_i Y] = b_i, \forall i \in [m], Y \in \mathbb{S}_+^r\}$
is a PSD extension of P of size r.

PSD Factorizations and Extensions.

Definition (PSD Rank)

$$\mathrm{rk}_{\mathrm{psd}}(S) := \min\{r \mid \exists \text{ rank-}r \text{ PSD factorization of } S\}$$

Theorem (Factorization Theorem [Gouviea, Thomas, Parillo '11, Fiorini, Massar, Pokutta, Tiwary, de Wolf '12])

For every slack matrix S of P:

$$\mathrm{xc}_{\mathrm{psd}}(P) = \mathrm{rk}_{\mathrm{psd}}(S)$$

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a linear EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized linear EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a linear EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized linear EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a linear EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized linear EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a linear EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized linear EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a linear EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized linear EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Rothvoss's Counting Argument

Goal:
Show existence of $X \subseteq \{0,1\}^n$ with $\mathrm{xc}_{\mathrm{psd}}(\mathrm{conv}(X)) = 2^{\Omega(n)}$.

Let $R = \max_{X \subseteq \{0,1\}^n} \mathrm{xc}_{\mathrm{psd}}(\mathrm{conv}(X))$.

High level:

1. "Discretize" an optimal a PSD EF for $\mathrm{conv}(X)$ to *compress the description* of X.
2. Show that the number of discretized PSD EFs of size R is bounded by $2^{\mathrm{poly}(R,n)}$.

Conclusion: 2^{2^n} subsets of $\{0,1\}^n$ means that $R \geq 2^{\Omega(n)}$.

Discretizing Linear EFs

$X = \{v_1, \ldots, v_N\} \subseteq \{0,1\}^n$. $\mathrm{conv}(X) = \{x \in \mathbb{R}^n : Ax \leq b\}$.

By Hadamard can assume $A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m$
and $\max |A_{ij}|, \max |b_i| \leq \Delta \approx n^{n/2}$.

Let $S_{ij} = b_i - A_i v_j, \forall i \in [m], j \in [N]$, be the slack matrix.

Discretizing Linear EFs

$X = \{v_1, \ldots, v_N\} \subseteq \{0,1\}^n$. $\mathrm{conv}(X) = \{x \in \mathbb{R}^n : Ax \leq b\}$.

By Hadamard can assume $A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m$ and $\max|A_{ij}|, \max|b_i| \leq \Delta \approx n^{n/2}$.

Let $S_{ij} = b_i - A_i v_j, \forall i \in [m], j \in [N]$, be the slack matrix.

Discretizing Linear EFs

$X = \{v_1, \ldots, v_N\} \subseteq \{0,1\}^n$. $\mathrm{conv}(X) = \{x \in \mathbb{R}^n : Ax \leq b\}$.

By Hadamard can assume $A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m$
and $\max |A_{ij}|, \max |b_i| \leq \Delta \approx n^{n/2}$.

Let $S_{ij} = b_i - A_i v_j, \forall i \in [m], j \in [N]$, be the slack matrix.

Discretizing Linear EFs

Let $S = UV$ be rank r nonnegative factorization.

Examine linear EF $Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$.

Let (A_S, U_S), $S \subseteq [m]$, be any maximal set of linearly independent set of rows of (A, U).

Note that $Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}$.

Can represent $\mathrm{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Discretizing Linear EFs

Let $S = UV$ be rank r nonnegative factorization.

Examine linear EF $Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$.

Let (A_S, U_S), $S \subseteq [m]$, be any maximal set of linearly independent set of rows of (A, U).

Note that $Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}$.

Can represent $\mathrm{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Discretizing Linear EFs

Let $S = UV$ be rank r nonnegative factorization.

Examine linear EF $Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$.

Let (A_S, U_S), $S \subseteq [m]$, be any maximal set of linearly independent set of rows of (A, U).

Note that $Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}$.

Can represent $\mathrm{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Discretizing Linear EFs

Let $S = UV$ be rank r nonnegative factorization.

Examine linear EF $Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$.

Let (A_S, U_S), $S \subseteq [m]$, be any maximal set of linearly independent set of rows of (A, U).

Note that $Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}$.

Can represent $\text{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Discretizing Linear EFs

Let $S = UV$ be rank r nonnegative factorization.

Examine linear EF $Q = \{(x, y) : Ax + Uy = b, y \geq 0\}$.

Let (A_S, U_S), $S \subseteq [m]$, be any maximal set of linearly independent set of rows of (A, U).

Note that $Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}$.

Can represent $\mathrm{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Discretizing Linear EFs

$$Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}.$$

Can represent $\text{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Issues with counting:

1. Numbers in U_S may be unbounded.
2. Numbers in U_S don't fall in a discrete set.

Fix idea:

1. Round numbers in U to a grid.
2. Build "rounded" EF $\bar{Q}$ that contains SAME $0/1$ points as Q.

Discretizing Linear EFs

$$Q = \{(x, y) : A_S x + U_S y = b_S, y \geq 0\}.$$

Can represent $\text{conv}(X)$ using only at most $(n + r) \times (n + r + 1)$ numbers!

Issues with counting:

1. Numbers in U_S may be unbounded.
2. Numbers in U_S don't fall in a discrete set.

Fix idea:

1. Round numbers in U to a grid.
2. Build "rounded" EF $\bar{Q}$ that contains SAME $0/1$ points as Q.

Discretizing Linear EFs

Rounded Linear EF:

$$\bar{Q} = \{(x,y) : \|A_S x + \bar{U}_S y - b_S\|_\infty \leq 1/\mathrm{poly}(n,r), \quad y \geq 0, \|y\|_\infty \leq \sqrt{\Delta}\}.$$

High Level Fix:

1. Rescale factorization $S = UV$ so that max entry $U, V \leq \sqrt{\Delta}$.
2. Round $U \to \bar{U}$ to nearest multiple of $1/\mathrm{poly}(n,r,\Delta)$.
3. Choose rows S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n+r) \times (n+r+1)$ numbers in a bounded set!

Discretizing Linear EFs

Rounded Linear EF:

$$\bar{Q} = \{(x, y) : \|A_S x + \bar{U}_S y - b_S\|_\infty \leq 1/\mathrm{poly}(n, r), \\ y \geq 0, \|y\|_\infty \leq \sqrt{\Delta}\}.$$

High Level Fix:

1. Rescale factorization $S = UV$ so that max entry $U, V \leq \sqrt{\Delta}$.
2. Round $U \to \bar{U}$ to nearest multiple of $1/\mathrm{poly}(n, r, \Delta)$.
3. Choose rows S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0, 1\}^n = X$.

Can represent X using only at most $(n + r) \times (n + r + 1)$ numbers in a bounded set!

Discretizing Linear EFs

Rounded Linear EF:
$$\bar{Q} = \{(x,y) : \|A_S x + \bar{U}_S y - b_S\|_\infty \leq 1/\text{poly}(n,r), \\ y \geq 0, \|y\|_\infty \leq \sqrt{\Delta}\}.$$

High Level Fix:

1. Rescale factorization $S = UV$ so that max entry $U, V \leq \sqrt{\Delta}$.
2. Round $U \to \bar{U}$ to nearest multiple of $1/\text{poly}(n, r, \Delta)$.
3. Choose rows S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n+r) \times (n+r+1)$ numbers in a bounded set!

Discretizing Linear EFs

Rounded Linear EF:

$$\bar{Q} = \{(x,y) : \|A_S x + \bar{U}_S y - b_S\|_\infty \leq 1/\mathrm{poly}(n,r), \quad y \geq 0, \|y\|_\infty \leq \sqrt{\Delta}\}.$$

High Level Fix:

1. Rescale factorization $S = UV$ so that max entry $U, V \leq \sqrt{\Delta}$.
2. Round $U \to \bar{U}$ to nearest multiple of $1/\mathrm{poly}(n,r,\Delta)$.
3. Choose rows S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n+r) \times (n+r+1)$ numbers in a bounded set!

Discretizing Linear EFs

Rounded Linear EF:

$$\bar{Q} = \{(x,y) : \|A_S x + \bar{U}_S y - b_S\|_\infty \leq 1/\mathrm{poly}(n,r),\\ y \geq 0, \|y\|_\infty \leq \sqrt{\Delta}\}.$$

High Level Fix:

1. Rescale factorization $S = UV$ so that max entry $U, V \leq \sqrt{\Delta}$.
2. Round $U \to \bar{U}$ to nearest multiple of $1/\mathrm{poly}(n,r,\Delta)$.
3. Choose rows S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n+r) \times (n+r+1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x, y) : \|A_i x + \mathrm{Tr}\,[\bar{U}_i Y] - b_i\|_\infty \le 1/\mathrm{poly}(n, r), \forall i \in S, \\ \|Y\|_{\mathrm{op}} \le ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\,[U_i V_j]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \le ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n, r, ???)$.
3. Choose rows of S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0, 1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x, y) : \|A_i x + \mathrm{Tr}\left[\bar{U}_i Y\right] - b_i\|_\infty \leq 1/\mathrm{poly}(n, r), \forall i \in S, \|Y\|_{\mathrm{op}} \leq ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \leq ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n, r, ???)$.
3. Choose rows of S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0, 1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x,y) : \|A_i x + \mathrm{Tr}\left[\bar{U}_i Y\right] - b_i\|_\infty \leq 1/\mathrm{poly}(n,r), \forall i \in S,$$
$$\|Y\|_{\mathrm{op}} \leq ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \leq ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n, r, ???)$.
3. Choose rows of S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x, y) : \|A_i x + \mathrm{Tr}\left[\bar{U}_i Y\right] - b_i\|_\infty \leq 1/\mathrm{poly}(n, r), \forall i \in S, \\ \|Y\|_{\mathrm{op}} \leq ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \leq ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n, r, ???)$.
3. Choose rows of S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0, 1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x,y) : \|A_i x + \mathrm{Tr}\left[\bar{U}_i Y\right] - b_i\|_\infty \le 1/\mathrm{poly}(n,r), \forall i \in S, \\ \|Y\|_{\mathrm{op}} \le ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \le ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n, r, ???)$.
3. Choose rows of S "carefully".
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Discretizing PSD EFs

Question: What changes for PSD factorizations?

Rounded PSD EF:

$$\bar{Q} = \{(x,y) : \|A_i x + \mathrm{Tr}\left[\bar{U}_i Y\right] - b_i\|_\infty \leq 1/\mathrm{poly}(n,r), \forall i \in S, \\ \|Y\|_{\mathrm{op}} \leq ???, Y \in \mathbb{S}^r_+\}.$$

Solution steps:

1. Rescale factorization $S_{ij} = \mathrm{Tr}\left[U_i V_j\right]$ so that $\max_i \|U_i\|_{\mathrm{op}}, \max_j \|V_j\|_{\mathrm{op}} \leq ???$.
2. Round each $U_i \Rightarrow \bar{U}_i$ to nearest multiple of $1/\mathrm{poly}(n,r,???)$.
3. Choose rows of S “carefully”.
4. Show that $\pi_x(\bar{Q}) \cap \{0,1\}^n = X$.

Can represent X using only at most $(n + r^2) \times (n + r^2 + 1)$ numbers in a bounded set!

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\ i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Rescaling Nonnegative Factorizations

Lemma

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_+(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *nonnegative factorization* $S = UV$ *such that*

$$\|U\|_\infty, \|V\|_\infty \leq \sqrt{\Delta}$$

Rescaling Nonnegative Factorizations

Proof: $U = \begin{pmatrix} u_1 & \dots & u_r \end{pmatrix}$, $V = \begin{pmatrix} v_1^\mathsf{T} \\ \vdots \\ v_r^\mathsf{T} \end{pmatrix}$

$S = \sum_{i=1}^r u_i v_i^\mathsf{T}$.

Rescaling Nonnegative Factorizations

Proof: $U = \begin{pmatrix} u_1 & \dots & u_r \end{pmatrix}$, $V = \begin{pmatrix} v_1^\mathsf{T} \\ \vdots \\ v_r^\mathsf{T} \end{pmatrix}$

$S = \sum_{i=1}^{r} u_i v_i^\mathsf{T}$.

Rescaling Nonnegative Factorizations

Choose $\lambda_1, \ldots, \lambda_r > 0$ such that $\|\lambda_i u_i\|_\infty = \|1/\lambda_i v_i\|_\infty$.

Let $U' = \begin{pmatrix} \lambda_1 u_1 & \ldots & \lambda_r u_r \end{pmatrix}$, $V' = \begin{pmatrix} 1/\lambda_1 v_1^\top \\ \ldots \\ 1/\lambda_r v_r^\top \end{pmatrix}$.

Note that $U'V' = \sum_{i=1}^r (\lambda_i u_i)(1/\lambda_i v_i^\top) = \sum_{i=1}^r u_i v_i^\top = S$.

Rescaling Nonnegative Factorizations

Choose $\lambda_1, \ldots, \lambda_r > 0$ such that $\|\lambda_i u_i\|_\infty = \|1/\lambda_i v_i\|_\infty$.

Let $U' = \begin{pmatrix} \lambda_1 u_1 & \ldots & \lambda_r u_r \end{pmatrix}$, $V' = \begin{pmatrix} 1/\lambda_1 v_1^\mathsf{T} \\ \ldots \\ 1/\lambda_r v_r^\mathsf{T} \end{pmatrix}$.

Note that $U'V' = \sum_{i=1}^r (\lambda_i u_i)(1/\lambda_i v_i^\mathsf{T}) = \sum_{i=1}^r u_i v_i^\mathsf{T} = S$.

Rescaling Nonnegative Factorizations

Choose $\lambda_1, \ldots, \lambda_r > 0$ such that $\|\lambda_i u_i\|_\infty = \|1/\lambda_i v_i\|_\infty$.

Let $U' = \begin{pmatrix} \lambda_1 u_1 & \ldots & \lambda_r u_r \end{pmatrix}$, $V' = \begin{pmatrix} 1/\lambda_1 v_1^\mathsf{T} \\ \ldots \\ 1/\lambda_r v_r^\mathsf{T} \end{pmatrix}$.

Note that $U'V' = \sum_{i=1}^r (\lambda_i u_i)(1/\lambda_i v_i^\mathsf{T}) = \sum_{i=1}^r u_i v_i^\mathsf{T} = S$.

Rescaling Nonnegative Factorizations

$$M \stackrel{\text{def}}{=} \|U'\|_\infty = \max_i \|\lambda_i u_i\|_\infty = \max_i \|1/\lambda v_i\|_\infty = \|V'\|_\infty.$$

Pick $j \in [r]$ such that $M = \|\lambda_j u_j\|_\infty = \|1/\lambda_j v_j\|_\infty$.

$$\begin{aligned}\Delta = \|S\|_\infty = \|U'V'\|_\infty &\geq \|(\lambda_j u_j)(1/\lambda_j v_j^\mathsf{T})\|_\infty \\ &= \|\lambda_j u_j\|_\infty \|1/\lambda_j v_j\|_\infty = M^2.\end{aligned}$$

Therefore $\|U'\|_\infty = \|V'\|_\infty = M \leq \sqrt{\Delta}$ as needed. $\square$

Rescaling Nonnegative Factorizations

$M \stackrel{\text{def}}{=} \|U'\|_\infty = \max_i \|\lambda_i u_i\|_\infty = \max_i \|1/\lambda v_i\|_\infty = \|V'\|_\infty$.

Pick $j \in [r]$ such that $M = \|\lambda_j u_j\|_\infty = \|1/\lambda_j v_j\|_\infty$.

$$\Delta = \|S\|_\infty = \|U'V'\|_\infty \geq \|(\lambda_j u_j)(1/\lambda_j v_j^{\mathsf{T}})\|_\infty$$
$$= \|\lambda_j u_j\|_\infty \|1/\lambda_j v_j\|_\infty = M^2.$$

Therefore $\|U'\|_\infty = \|V'\|_\infty = M \leq \sqrt{\Delta}$ as needed. $\square$

Rescaling Nonnegative Factorizations

$$M \stackrel{\text{def}}{=} \|U'\|_\infty = \max_i \|\lambda_i u_i\|_\infty = \max_i \|1/\lambda v_i\|_\infty = \|V'\|_\infty.$$

Pick $j \in [r]$ such that $M = \|\lambda_j u_j\|_\infty = \|1/\lambda_j v_j\|_\infty$.

$$\begin{aligned} \Delta = \|S\|_\infty = \|U'V'\|_\infty &\geq \|(\lambda_j u_j)(1/\lambda_j v_j^\mathsf{T})\|_\infty \\ &= \|\lambda_j u_j\|_\infty \|1/\lambda_j v_j\|_\infty = M^2. \end{aligned}$$

Therefore $\|U'\|_\infty = \|V'\|_\infty = M \leq \sqrt{\Delta}$ as needed. □

Rescaling Nonnegative Factorizations

$M \stackrel{\text{def}}{=} \|U'\|_\infty = \max_i \|\lambda_i u_i\|_\infty = \max_i \|1/\lambda v_i\|_\infty = \|V'\|_\infty$.

Pick $j \in [r]$ such that $M = \|\lambda_j u_j\|_\infty = \|1/\lambda_j v_j\|_\infty$.

$$\begin{aligned}\Delta = \|S\|_\infty = \|U'V'\|_\infty &\geq \|(\lambda_j u_j)(1/\lambda_j v_j^\mathsf{T})\|_\infty \\ &= \|\lambda_j u_j\|_\infty \|1/\lambda_j v_j\|_\infty = M^2.\end{aligned}$$

Therefore $\|U'\|_\infty = \|V'\|_\infty = M \leq \sqrt{\Delta}$ as needed. $\square$

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\, i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Comparison with Nonnegative Factorizations:

- Nonnegative setting: can rescale entries of non-negative vector independently and maintain non-negativity.
- PSD setting: entries can be NEGATIVE.
  CANNOT rescale entries of PSD matrix independently while maintaining PSD property.

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\, i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Comparison with Nonnegative Factorizations:

- Nonnegative setting: can rescale entries of non-negative vector independently and maintain non-negativity.
- PSD setting: entries can be NEGATIVE.
  CANNOT rescale entries of PSD matrix independently while maintaining PSD property.

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\, i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Comparison with Nonnegative Factorizations:

- Nonnegative setting: can rescale entries of non-negative vector independently and maintain non-negativity.
- PSD setting: entries can be NEGATIVE.
  CANNOT rescale entries of PSD matrix independently while maintaining PSD property.

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_{+}^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_{\infty} = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\, i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Admissible PSD Rescalings: For $A \in \mathbb{R}^{r \times r}$ invertible, send $U_i \to A^{\mathsf{T}} U_i A$ and $V_j \to A^{-1} V_i A^{-\mathsf{T}}$.

Preserves inner product:
$\mathrm{Tr}\left[A^{\mathsf{T}} U_i A A^{-1} V_j A^{-T}\right] = \mathrm{Tr}\left[U_i V_j A^{-\mathsf{T}} A^{\mathsf{T}}\right] = \mathrm{Tr}\left[U_i V_j\right] = S_{ij}$.

Map $U \to A^{\mathsf{T}} U A$ is a symmetry of PSD cone.

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\, i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Admissible PSD Rescalings: For $A \in \mathbb{R}^{r \times r}$ invertible, send $U_i \to A^\mathsf{T} U_i A$ and $V_j \to A^{-1} V_i A^{-\mathsf{T}}$.

Preserves inner product:
$\mathrm{Tr}\left[A^\mathsf{T} U_i A A^{-1} V_j A^{-T}\right] = \mathrm{Tr}\left[U_i V_j A^{-\mathsf{T}} A^\mathsf{T}\right] = \mathrm{Tr}\,[U_i V_j] = S_{ij}$.

Map $U \to A^\mathsf{T} U A$ is a symmetry of PSD cone.

Rescaling PSD Factorizations

Theorem ([Briët, D., Pokutta])

Let $S \in \mathbb{R}_+^{m \times N}$ *with* $\mathrm{rk}_{\mathrm{psd}}(S) = r$ *and* $\|S\|_\infty = \Delta$.
Then $\exists$ *rank* r *PSD factorization* $S_{ij} = \mathrm{Tr}\,[U_i V_j]$, $\forall\ i, j$ *such that*

$$\max_i \|U_i\|_{op}, \max_j \|V_j\|_{op} \leq \sqrt{r\Delta}$$

Admissible PSD Rescalings: For $A \in \mathbb{R}^{r \times r}$ invertible, send $U_i \to A^{\mathsf{T}} U_i A$ and $V_j \to A^{-1} V_i A^{-\mathsf{T}}$.

Preserves inner product:
$\mathrm{Tr}\left[A^{\mathsf{T}} U_i A A^{-1} V_j A^{-T}\right] = \mathrm{Tr}\left[U_i V_j A^{-\mathsf{T}} A^{\mathsf{T}}\right] = \mathrm{Tr}\left[U_i V_j\right] = S_{ij}$.

Map $U \to A^{\mathsf{T}} U A$ is a symmetry of PSD cone.

Rescaling PSD Factorizations

Proof Idea: Variational Argument

1. Choose rescaling $A \in \mathbb{S}^r_+$ such that potential $\max_i \|AU_iA\|_{op} \times \max_j \|A^{-1}V_jA^{-1}\|_{op}$ is minimized.

2. Show that if potential $> r\Delta$, can find infinitessimal pertubation $A \to (I + \epsilon P)A(I + \epsilon P)$ which decreases potential.

Require convex geometric tools
(John's decomposition of the identity) to build perturbation P.

Rescaling PSD Factorizations

Proof Idea: Variational Argument

1. Choose rescaling $A \in \mathbb{S}^r_+$ such that potential $\max_i \|AU_iA\|_{op} \times \max_j \|A^{-1}V_jA^{-1}\|_{op}$ is minimized.
2. Show that if potential $> r\Delta$, can find infinitessimal pertubation $A \to (I + \epsilon P)A(I + \epsilon P)$ which decreases potential.

Require convex geometric tools
(John's decomposition of the identity) to build perturbation P.

Rescaling PSD Factorizations

Proof Idea: Variational Argument

1. Choose rescaling $A \in \mathbb{S}^r_+$ such that potential $\max_i \|AU_iA\|_{op} \times \max_j \|A^{-1}V_jA^{-1}\|_{op}$ is minimized.

2. Show that if potential $> r\Delta$, can find infinitessimal pertubation $A \to (I + \epsilon P)A(I + \epsilon P)$ which decreases potential.

Require convex geometric tools
(John's decomposition of the identity) to build perturbation P.

Open problems and future work.

1. Is the rescaling theorem tight? Do we need dependance on r?
2. Show that any n^r PSD extended formulation for Correlation polytope has integrality gap as large as $O(r)$ levels of Lasserre.
3. Extension complexity of the matching polytope.

Open problems and future work.

1. Is the rescaling theorem tight? Do we need dependance on r?
2. Show that any n^r PSD extended formulation for Correlation polytope has integrality gap as large as $O(r)$ levels of Lasserre.
3. Extension complexity of the matching polytope.

Open problems and future work.

1. Is the rescaling theorem tight? Do we need dependance on r?
2. Show that any n^r PSD extended formulation for Correlation polytope has integrality gap as large as $O(r)$ levels of Lasserre.
3. Extension complexity of the matching polytope.

Thank you!