CS 59300 – Algorithms for Data Science

Classical and Quantum approaches

Lecture 8 (09/30)

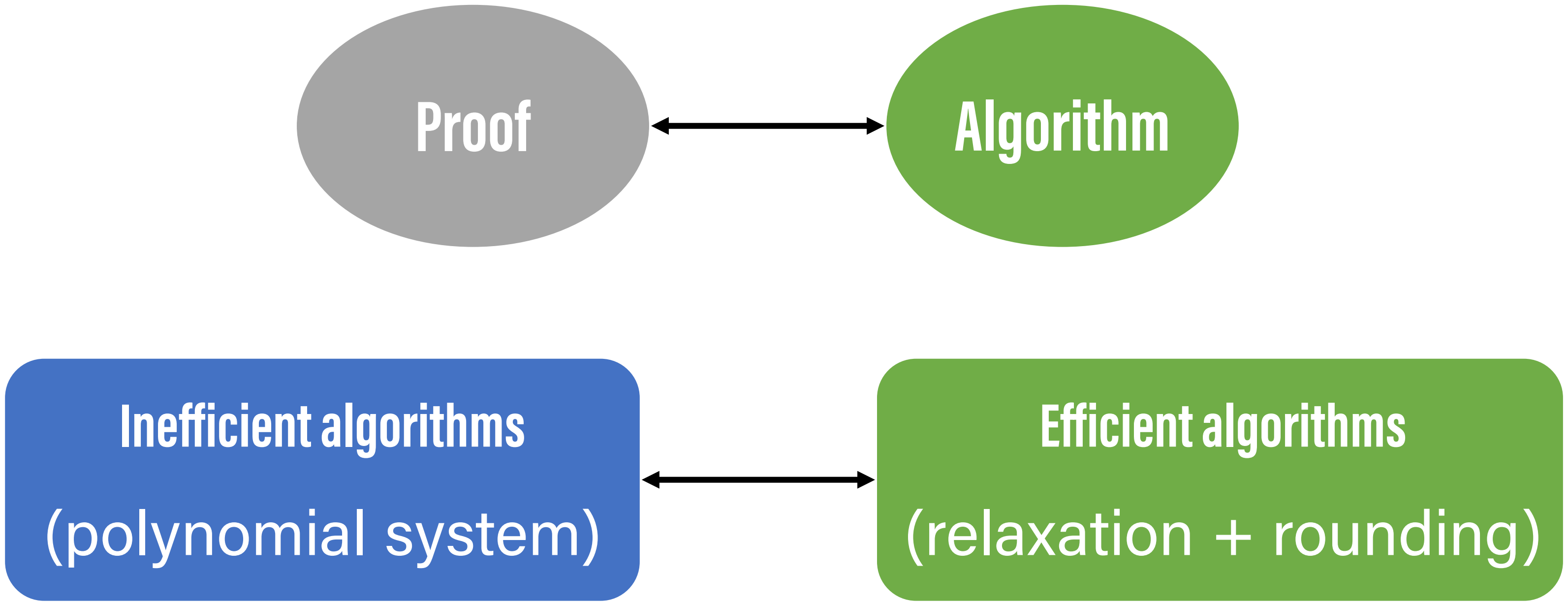
Sum-of-Squares (I)

https://ruizhezhang.com/course_fall_2025.html

Part of the slides is based on Pravesh Kothari's talk

Sum-of-Squares (SoS)

Powerful generic framework for algorithm design/nonconvex optimization



Yields the most powerful approximation algorithms for many statistical/ML problems

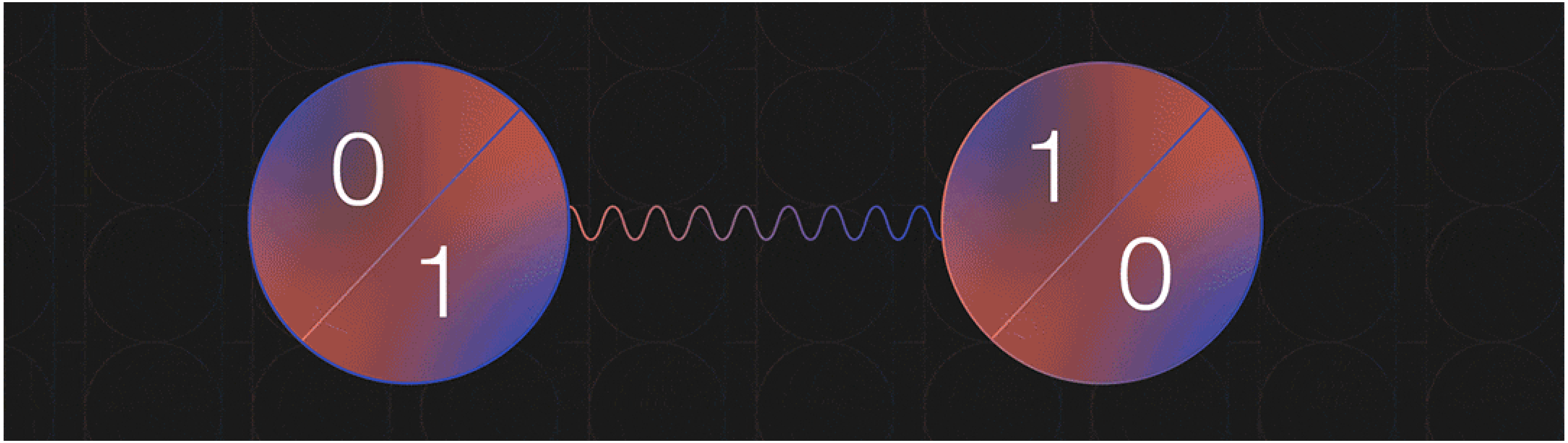
- Max-cut, tensor decomposition, dictionary learning, matrix/tensor completion, sparse PCA, Gaussian mixture models, planted clique, robust statistics, quantum separability, ...

Quantum separability

What is quantum entanglement?

Quantum separability

What is quantum entanglement?



Hilbert space: $\mathcal{H}_A \quad \otimes \quad \mathcal{H}_B$

Quantum separability

$$\mathcal{H}_A \otimes \mathcal{H}_B \neq \{|\phi\rangle_A \otimes |\psi\rangle_B \;:\; |\phi\rangle \in \mathcal{H}_A, |\psi\rangle \in \mathcal{H}_B\}$$

Quantum separability

$$\mathcal{H}_A \otimes \mathcal{H}_B = \left\{ \sum_i c_i |\phi_i\rangle_A \otimes |\psi_i\rangle_B \; : |\phi_i\rangle \in \mathcal{H}_A, |\psi_i\rangle \in \mathcal{H}_B, c_i \in \mathbb{C} \right\}$$

- (Pure) separable state:

$$|\Psi\rangle = |\phi\rangle_A \otimes |\psi\rangle_B$$

E.g. $|01\rangle = |0\rangle \otimes |1\rangle$

- (Pure) entangled state:

$$|\Psi\rangle \neq |\phi\rangle_A \otimes |\psi\rangle_B \quad \forall\, |\phi\rangle \in \mathcal{H}_A, |\psi\rangle \in \mathcal{H}_B$$

E.g. $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ (Bell state)

Quantum separability

Sometimes we don't have a complete knowledge of the quantum system

- Mixed state: probabilistic mixtures of pure states, e.g. $\Pr[\rho = |0\rangle] = \frac{1}{2}, \Pr[\rho = |1\rangle] = \frac{1}{2}$
- Represented by the density **matrix** $\rho \in B(\mathcal{H})$:

$$\rho \succcurlyeq 0 \quad \text{and} \quad \mathrm{tr}[\rho] = 1$$

e.g.

$$\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

- Mixed separable state:

$$\rho_{AB} = \sum_i p_i \sigma_i^A \otimes \tau_i^B = \sum_i q_i |\phi_i\rangle_A\langle\phi_i| \otimes |\psi_i\rangle_B\langle\psi_i|$$

- Mixed entangled state: any ρ_{AB} not separable

Quantum separability

- The Werner state (2-qubit):

$$\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)I_4$$

  where $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$
  - $p = 1$: pure entangled state
  - $1/3 < p < 1$: mixed entangled state
  - $0 < p \leq 1/3$: mixed separable state
  - $p = 0$: maximally mixed state (noise)

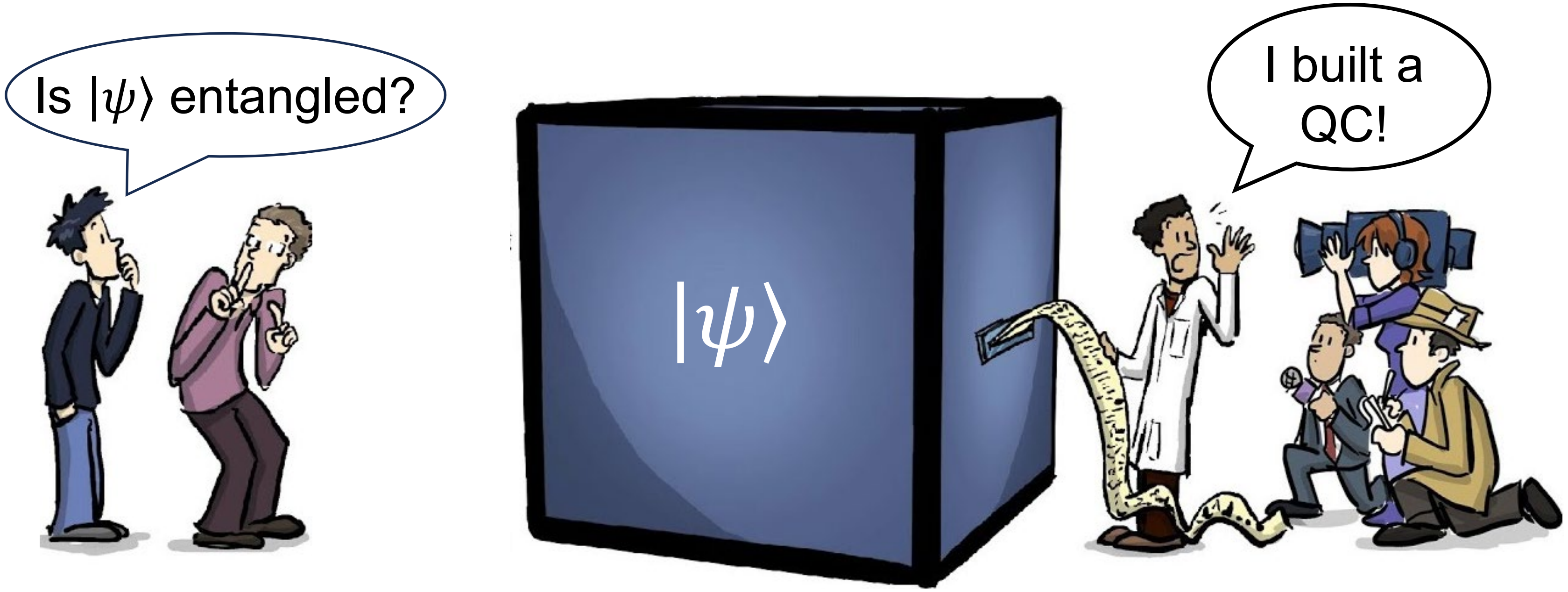
Quantum separability: given the full description of a density matrix $\rho_{AB} \in \mathbb{C}^{n^2 \times n^2}$, decide whether ρ_{AB} is entangled or separable

Quantum separability

Quantum separability: given a density matrix $\rho_{AB} \in \mathbb{C}^{n^2 \times n^2}$, decide:

- **YES:** ρ_{AB} is separable; or
- **NO:** ρ_{AB} is ϵ-far from the set of separable states

i.e. the weak membership problem for the set of separable states



Entanglement witness and the best separable state (BSS)

Entanglement witness for an entangled state ρ is a measurement $\mathcal{M}$ such that

1. $\mathcal{M}$ accepts ρ with probability 1;
2. For any separable state ρ', $\mathcal{M}$ accepts ρ' with probability at most $1 - \epsilon$

Quantum measurement

- $\mathcal{M} \in \mathbb{C}^{n^2 \times n^2}$ is Hermitian and $0 \preccurlyeq \mathcal{M} \preccurlyeq I$
- The probability that $\mathcal{M}$ accepts ρ is $\mathrm{tr}[\mathcal{M}\rho]$

Best separable state problem ($BSS_{1,s}$): given an n^2-by-n^2 measurement $\mathcal{M}$, distinguish:

- **YES:** there is a separable state ρ such that $\mathrm{tr}[\mathcal{M}\rho] = 1$
- **NO:** for any separable state ρ, $\mathrm{tr}[\mathcal{M}\rho] \leq s$

History and motivations of BSS

Best separable state problem ($BSS_{c,s}$): given an n^2-by-n^2 measurement $\mathcal{M}$ and $0 \leq s < c \leq 1$, distinguish the following two cases:

- **YES:** there is a separable state ρ such that $\mathrm{tr}[\mathcal{M}\rho] \geq c$
- **NO:** for any separable state ρ, $\mathrm{tr}[\mathcal{M}\rho] \leq s$

- Brute-force: $2^{\mathcal{O}(n)}$-time
- **Blier-Tapp '09, Gurvits '03:** $BSS_{c,s}$ is NP-hard for any $c - s \geq 1/\mathrm{poly}(n)$
- **Harrow-Montanaro '13:** no $n^{o(\log n)}$-time algorithm for $BSS_{1,0.5}$ assuming **ETH**
- Quantum complexity implication: if there is a quasi-polynomial time algorithm for $BSS_{0.99,0.5}$, then $\mathbf{QMA(2)} \subset \mathbf{EXP}$
- **Brandão et al. '11, Brandão-Harrow '15:** quasi-poly algorithm for special family of measurements

Main result

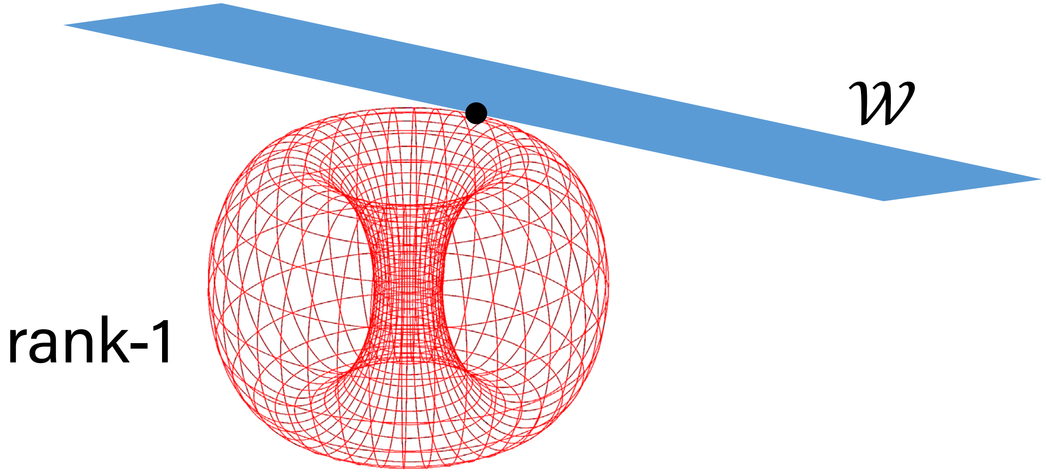
Theorem (Barak-Kothari-Steurer '17).

For every $s < 1$, there is a $2^{\tilde{O}(\sqrt{n})}$-time algorithm for $BSS_{1,s}$, based on rounding an SoS relaxation

Classical version of BSS

Find a rank-1 matrix in a subspace:

- Input: the basis of a subspace $\mathcal{W} \subset \mathbb{R}^{n\times n}$ of $n \times n$ matrices
- Promise: $uv^\top \in \mathcal{W}$ for $u, v \in \mathbb{S}^{n-1}$
- Goal: find $x, y \in \mathbb{S}^{n-1}$ such that $\mathrm{dist}(xy^\top, \mathcal{W}) < 0.01$

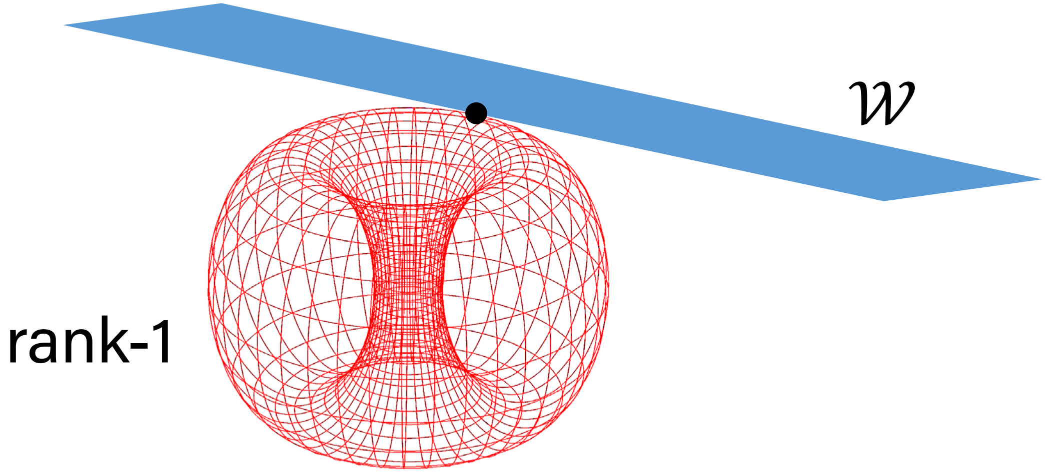


Classical version of BSS

Find a rank-1 matrix in a subspace:

- Input: the basis of a subspace $\mathcal{W} \subset \mathbb{R}^{n \times n}$ of $n \times n$ matrices
- Promise: $uu^\top \in \mathcal{W}$ for $u \in \mathbb{S}^{n-1}$
- Goal: find $x \in \mathbb{S}^{n-1}$ such that $\mathrm{dist}(xx^\top, \mathcal{W}) < 0.01$

Only "symmetric" version for simplicity



Polynomial optimization formulation

Find a rank-1 matrix in a subspace:

- **Input:** $\Pi \in \mathbb{R}^{n^2 \times n^2}$ the orthogonal projector for the subspace $\mathcal{W}$
- **Promise:** $\Pi(u \otimes u) = u \otimes u$ for $u \in \mathbb{S}^{n-1}$
- **Goal:** find $x \in \mathbb{S}^{n-1}$ such that $\|\Pi(x \otimes x)\|_F^2 > 0.99$

Degree-4 polynomial optimization:

$$\max_{x_1,\ldots x_n \in \mathbb{R}} P(x) := (x \otimes x)^\top \Pi(x \otimes x)$$

$$\text{s.t.} \quad x_1^2 + \cdots x_n^2 = 1$$

Approximation algorithm = Relaxation + Rounding

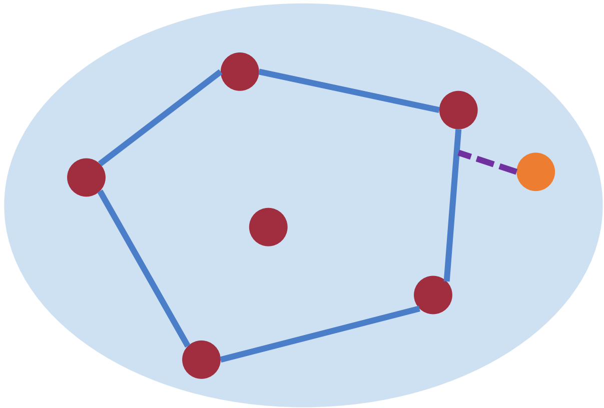
Is there a rank-1 matrix in $\mathcal{W}$?

Is the polynomial system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

General paradigm in approximation algorithm design

- Choose a convex relaxation
- Solve a convex program
- Round to a feasible solution



Blueprint

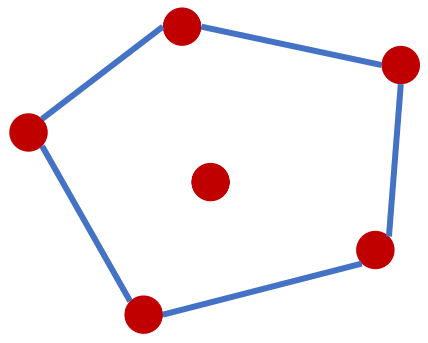
Is there a rank-1 matrix in $\mathcal{W}$?

Is the polynomial system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

1. Choose the "tightest" convex relaxation

   → Feasible points = distributions over solutions

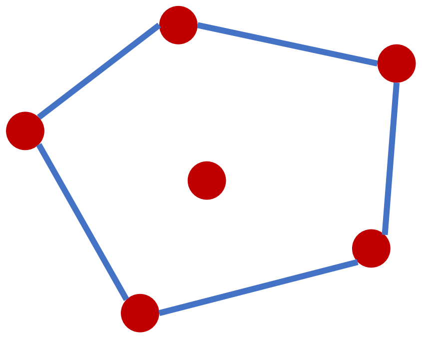


2. Show how to round any solution

3. Dive in to find the constraints actually used

4. Relax to a convex program by including only useful constraints

Solutions as distributions



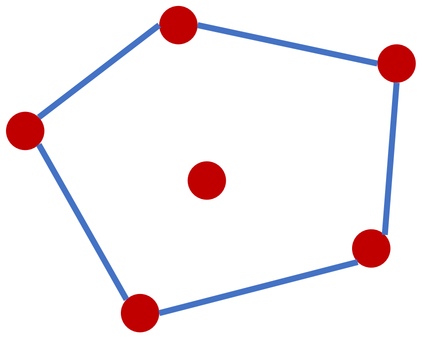
Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Let μ be a probability distribution over x satisfying (1) and (2)

Consider $\mathbb{E}_\mu$ and any polynomial $q(x)$:

- **Positivity:** $\mathbb{E}_\mu[q] \geq 0$ whenever $q(x) \geq 0 \; \forall \, x$
- **Normalization:** $\mathbb{E}_\mu[1] = 1$
- **Constraints satisfiability:** $\mathbb{E}_\mu[\Pi(x \otimes x)] = \mathbb{E}_\mu[x \otimes x]$ and $\mathbb{E}_\mu[\|x\|^2] = 1$

Solutions as distributions



Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Let μ be a probability distribution over x satisfying (1) and (2)

Consider $\mathbb{E}_\mu$ and any polynomial $q(x)$:

- **Positivity:** $\mathbb{E}_\mu[q] \geq 0$ whenever $q(x) \geq 0 \; \forall \, x$
- **Normalization:** $\mathbb{E}_\mu[1] = 1$
- **Constraints satisfiability:** $\mathbb{E}_\mu[q \cdot \Pi(x \otimes x)] = \mathbb{E}_\mu[q \cdot (x \otimes x)]$ and $\mathbb{E}_\mu[q \cdot \|x\|^2] = \mathbb{E}_\mu[q]$

μ requires exponentially many parameters to describe

- Only use degree-d moments of μ:

$$\{\mathbb{E}_\mu[x^S] \mid |S| \subset [n], |S| \leq d\}$$

Rounding low-degree moments

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Rounding: Given degree-d moments of μ satisfying (1) and (2) $\{\mathbb{E}_\mu[x^S] \mid |S| \subset [n], |S| \leq d\}$, find an approximation solution $x^\star \in \mathbb{S}^{n-1}$ such that $\|\Pi(x^\star \otimes x^\star)\| \geq 0.99$

Thought experiment:

- Say, μ is uniform over $u_1, \ldots, u_N \in \mathbb{S}^{n-1}$ such that $\Pi(u_i \otimes u_i) = u_i \otimes u_i$
- Say, the set of rank-1 matrices were convex
- By linearity,

$$\mathbb{E}_\mu[xx^\top] = \frac{1}{N} \sum_i u_i u_i^\top \in \mathcal{W}$$

Done!

Approximate "convexity"

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Rounding: Given degree-d moments of μ satisfying (1) and (2) $\{\mathbb{E}_\mu[x^S] \mid |S| \subset [n], |S| \le d\}$, find an approximation solution $x^\star \in \mathbb{S}^{n-1}$ such that $\|\Pi(x^\star \otimes x^\star)\| \ge 0.99$

Lemma (Approximately rank-1 suffices). Suppose that

$$\lambda_{\max}\left(\mathbb{E}_\mu[xx^\top]\right) \ge 0.99\left\|\mathbb{E}_\mu[xx^\top]\right\|_F$$

Then the top eigenvector y of $\mathbb{E}_\mu[xx^\top]$ satisfies $\|\Pi(y \otimes y)\| \ge 0.9$

Proof.

- $\|\lambda yy^\top - \mathbb{E}[xx^\top]\|_F^2 = \|\mathbb{E}[xx^\top]\|_F^2 - \lambda^2 \le 0.01 \cdot \lambda^2$
- $\|\Pi(y \otimes y)\| \ge \|\Pi\mathbb{E}[x \otimes x]\|/\lambda - 0.1 = \|\mathbb{E}[\Pi(x \otimes x)]\|/\lambda - 0.1 = \|\mathbb{E}[x \otimes x]\|/\lambda - 0.1 \ge 0.9$

$\blacksquare$

Blueprint

Is there a rank-1 matrix in $\mathcal{W}$? → Is the polygonal system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

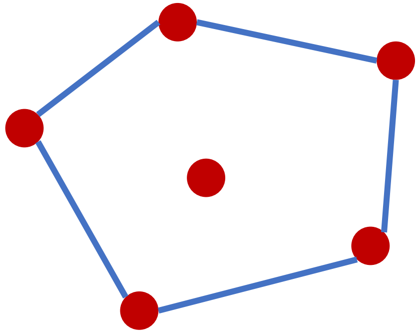
1. Choose the "tightest" convex relaxation

   → Feasible points = distributions over solutions

2. Show how to round any solution

   → "Bulk" of the proof

3. Dive in to find the constraints actually used



4. Relax to a convex program by including only useful constraints

Rounding via conditioning

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Rounding:

$$\lambda_{\max}\left(\mathbb{E}_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}_\mu[xx^\top]\right\|_F$$

Thought experiment 2:

- Say, μ is uniform over $\mathbb{S}^{n-1}$
- $\mathbb{E}_\mu[xx^\top] = \frac{1}{n}I$, i.e. not rank-1

Rounding via conditioning

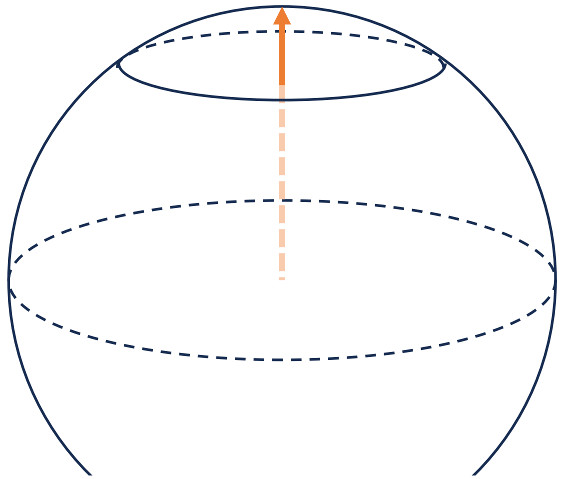
Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Rounding:

$$\lambda_{\max}\left(\mathbb{E}_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}_\mu[xx^\top]\right\|_F$$

Thought experiment 2:

- Say, μ is uniform over $\mathbb{S}^{n-1}$
- $\mathbb{E}_\mu[xx^\top \mid |x_1| \geq t] \approx \mathrm{diag}\left(t^2, \frac{1-t^2}{n-1}, \dots, \frac{1-t^2}{n-1}\right)$
- Define μ' to be the μ conditioned on $|x_1| \geq t$
- Then, $\lambda_{\max}\left(\mathbb{E}_{\mu'}[xx^\top]\right) = t^2$ and $\left\|\mathbb{E}_{\mu'}[xx^\top]\right\|_F \approx \sqrt{t^4 + 1/n}$
- We just need $t \gg n^{-1/4}$



How to conditioning on some event if only access to moments $\mathbb{E}[x^S]$?

Rounding via ~~conditioning~~ polynomial reweighing

Key idea: polynomial reweighing!

- Define a linear operator $\mathbb{E}'_\mu$ such that for any polynomial $p(x)$,

$$\mathbb{E}'_\mu[p] \coloneqq \frac{\mathbb{E}_\mu[w(x)p(x)]}{\mathbb{E}_\mu[w(x)]}$$

- If w is the indicator function for $\mathcal{E}$, then $\mathbb{E}'_\mu$ is just the conditional expectation
- Instead, we'll choose w to be some low-degree polynomial

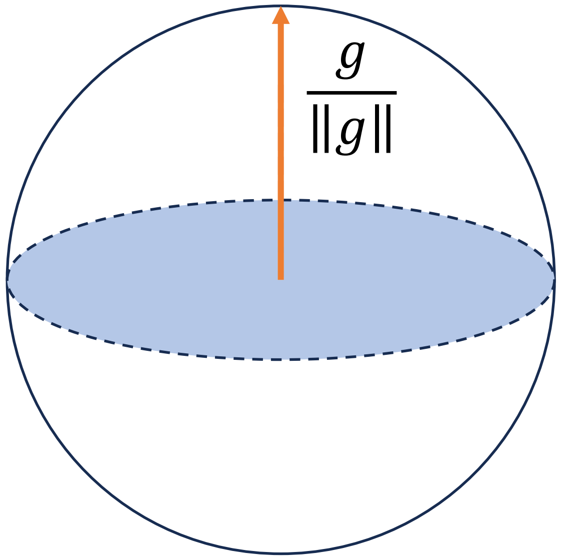
Claim. There is a degree-$\mathcal{O}(\sqrt{n})$ reweighing of the uniform distribution on $\mathbb{S}^{n-1}$ which has the approximate rank-1 property: $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F$

Rounding via polynomial reweighing

Claim. There is a degree-$\mathcal{O}(\sqrt{n})$ reweighing of the uniform distribution μ on $\mathbb{S}^{n-1}$ which has the approximate rank-1 property: $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F$

Proof.

- Let $g \sim \mathcal{N}(0, I)$ be a Gaussian vector and $w(x) \coloneqq \langle x, g\rangle^k$
- $\mathbb{E}'[xx^\top] = \frac{\mathbb{E}[\langle x,g\rangle^k xx^\top]}{\mathbb{E}[\langle x,g\rangle^k]}$
- $\lambda_{\max} = \frac{g^\top}{\|g\|}\mathbb{E}'[xx^\top]\frac{g}{\|g\|} = \frac{\mathbb{E}[\langle x,g\rangle^{k+2}]}{\mathbb{E}[\langle x,g\rangle^k]\|g\|^2}$
- $\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F^2 = \lambda_{\max}^2 + (1-\lambda_{\max})^2/(n-1)$
- If $\lambda_{\max} \geq \frac{k+2}{2n}$, then we are done by taking $k = \mathcal{O}(\sqrt{n})$!

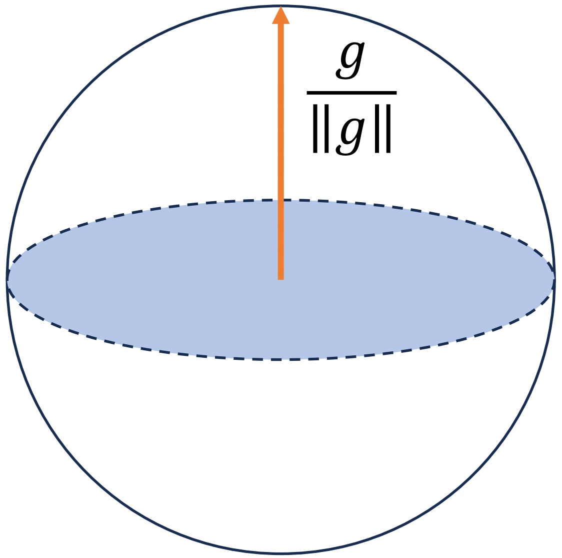


Rounding via polynomial reweighing

Claim. There is a degree-$\mathcal{O}(\sqrt{n})$ reweighing of the uniform distribution μ on $\mathbb{S}^{n-1}$ which has the approximate rank-1 property: $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F$

Proof.

- Let $g \sim \mathcal{N}(0, I)$ be a Gaussian vector and $w(x) \coloneqq \langle x, g\rangle^k$
- Goal: $\mathbb{E}_x[\langle x, g\rangle^{k+2}] \geq \frac{k+2}{2n} \mathbb{E}_x[\langle x, g\rangle^k] \|g\|^2$



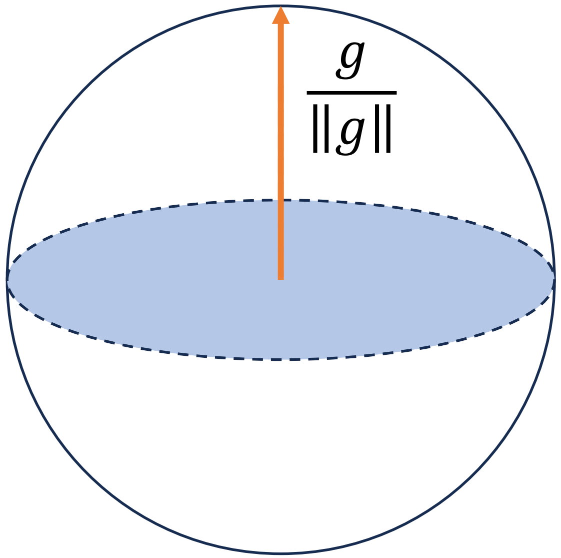
Rounding via polynomial reweighing

Claim. There is a degree-$\mathcal{O}(\sqrt{n})$ reweighing of the uniform distribution μ on $\mathbb{S}^{n-1}$ which has the approximate rank-1 property: $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F$

Proof.

- Let $g \sim \mathcal{N}(0, I)$ be a Gaussian vector and $w(x) \coloneqq \langle x, g\rangle^k$
- Goal: $\mathbb{E}_g\left[\mathbb{E}_x[\langle x, g\rangle^{k+2}]\right] \geq \frac{k+2}{2n}\mathbb{E}_g[\mathbb{E}_x[\langle x, g\rangle^k]\|g\|^2]$
- LHS $= \mathbb{E}_{z\sim\mathcal{N}(0,1)}[z^{k+2}] = (k+2)!!$
- RHS $= \frac{k+2}{2n}\mathbb{E}_{z\sim\mathcal{N}(0,1)}\left[z^k\left(z^2 + (n-1)\right)\right] = \frac{k+2}{2n}\left((k+2)!! + (n-1)k!!\right)$

$\blacksquare$



Rounding via polynomial reweighing

Structure theorem.

For any distribution μ over $\mathbb{S}^{n-1}$, there exists a degree-$\mathcal{O}(\sqrt{n})$ reweighing $\mathbb{E}'_\mu$ such that $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \geq 0.99\left\|\mathbb{E}_\mu[xx^\top]\right\|_F$

Easy to make $\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F \approx \left\|\mathbb{E}_\mu[xx^\top]\right\|_F$

Proof.

- Let $k = \mathcal{O}(\sqrt{n})$ and $g \sim \mathcal{N}\left(0, \mathbb{E}_\mu[xx^\top]\right)$
- Let $\mathbb{E}'_a$ be the reweighing with $\langle x, g\rangle^a$ for even $0 \leq a \leq k-2$
- Then, we have $\lambda_{\max}(\mathbb{E}'_a[xx^\top]) = \frac{\mathbb{E}\left[\langle x,g\rangle^{a+2}\right]}{\mathbb{E}[\langle x,g\rangle^a]\|g\|^2}$
- If $\prod_{a=0}^{k-2} \lambda_{\max}(\mathbb{E}'_a[xx^\top]) = \frac{\mathbb{E}\left[\langle x,g\rangle^k\right]}{\|g\|^k} \geq 0.99^{k/2}\|\mathbb{E}[xx^\top]\|_F^{k/2}$ with positive probability, then we are done! (why?)

Rounding via polynomial reweighing

$$\mathbb{E}_g\left[\mathbb{E}_\mu[\langle x, g\rangle^k]\right] \geq 0.99^{k/2}\left\|\mathbb{E}_\mu[xx^\top]\right\|_F^{k/2}\mathbb{E}_g[\|g\|^k] \text{ where } g \sim \mathcal{N}\left(0, \mathbb{E}_\mu[xx^\top]\right)$$

- For the LHS, we have

(Jensen)

$$\mathbb{E}_\mu\left[\mathbb{E}_g[\langle x, g\rangle^k]\right] = (k-1)!!\,\mathbb{E}_\mu\left[\left(x^\top \mathbb{E}_\mu[xx^\top]x\right)^{k/2}\right] \geq (k-1)!!\,\mathbb{E}_\mu\left[x^\top \mathbb{E}_\mu[x^\top x]x\right]^{k/2}$$
$$= (k-1)!!\,\left\|\mathbb{E}_\mu[xx^\top]\right\|_F^k$$

- For the RHS, it is more complicated:

$$\mathbb{E}_g[\|g\|^k] \leq \sum_{p \leq k/2} \binom{k/2}{p}(2p-1)!!\,\left\|\mathbb{E}_\mu[xx^\top]\right\|_F^p$$
$$\leq 1.01^{k/2}(k-1)!!\,\left\|\mathbb{E}_\mu[xx^\top]\right\|_F^{k/2}$$

if $k \geq \frac{C}{\left\|\mathbb{E}_\mu[xx^\top]\right\|_F} = \Theta(\sqrt{n})$

■

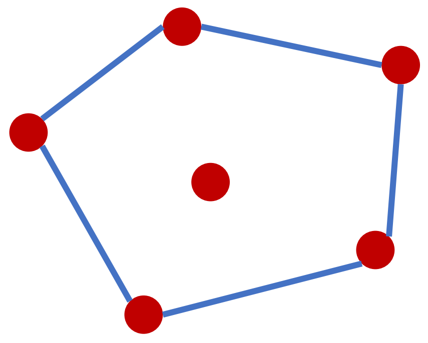
Blueprint

Is there a rank-1 matrix in $\mathcal{W}$?

Is the polynomial system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

1. Choose the "tightest" convex relaxation
   → Feasible points = distributions over solutions
2. Show how to round any solution
   → "Bulk" of the proof
3. Dive in to find the constraints actually used
   → Obtain a degree-d Sum-of-Squares proof
4. Relax to a convex program by including only useful constraints



Algorithm from rounding

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Rounding: Given degree-d moments of μ satisfying (1) and (2) $\left\{\mathbb{E}_\mu[x^S] \;\middle|\; |S| \subset [n], |S| \le d\right\}$, find an approximation solution $x^\star \in \mathbb{S}^{n-1}$ such that $\|\Pi(x^\star \otimes x^\star)\| \ge 0.99$

Structure theorem: For any μ over $\mathbb{S}^{n-1}$, there exists a degree-$\tilde{O}(\sqrt{n})$ reweighing $\mathbb{E}'_\mu$ such that $\lambda_{\max}\left(\mathbb{E}'_\mu[xx^\top]\right) \ge 0.99\left\|\mathbb{E}'_\mu[xx^\top]\right\|_F$

Issue: can't hope to compute low-degree moments of μ efficiently!

Key idea: "simple proof" cannot distinguish between distributions and "pseudo-distributions"

- We'll show that the rounding algorithm works for a less-constrained version of probability distributions, that one can efficiently optimized over

Solutions as distributions

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Let μ be a probability distribution over x satisfying (1) and (2)

Consider $\mathbb{E}_\mu$ and any polynomial $q(x)$ of degree $\tilde{O}(\sqrt{n})$:

- ? **Positivity:** $\mathbb{E}_\mu[q] \geq 0$ whenever $q(x) \geq 0 \ \forall \ x$
- ✓ **Normalization:** $\mathbb{E}_{\boldsymbol{\mu}}[1] = 1$
- ✓ **Constraints satisfiability:** $\mathbb{E}_\mu[q \cdot \Pi(x \otimes x)] = \mathbb{E}_\mu[q \cdot (x \otimes x)]$ and $\mathbb{E}_\mu[q \cdot \|x\|^2] = \mathbb{E}_\mu[q]$
- Only use degree-$\tilde{O}(\sqrt{n})$ moments of μ:

$$\left\{\mathbb{E}_\mu[x^S] \,\middle|\, |S| \subset [n], |S| \leq d\right\}$$

Positivity is tricky

- No hope for an efficient algorithm even restrict to low degree polynomials
- Conditioning relies on positivity to ensure polynomial reweighing gives another valid distribution

Pseudo-distribution

Polynomial formulation: (1) $\Pi(x \otimes x) = x \otimes x$ (2) $\|x\|^2 = 1$

Let $\tilde{\mu}$ be a degree-d pseudo-distribution over x satisfying (1) and (2)

Consider $\widetilde{\mathbb{E}} := \widetilde{\mathbb{E}}_{\tilde{\mu}}$ (pseudo-expectation):

- **Linearity:** $\widetilde{\mathbb{E}}$ is a linear operator, described by pseudo-moments $\widetilde{\mathbb{E}}[x^S]$
- **Normalization:** $\widetilde{\mathbb{E}}[1] = 1$
- **Constraints satisfiability:** for any degree$\leq (d-2)$ polynomial q,

$$\widetilde{\mathbb{E}}[q \cdot (I - \Pi)(x \otimes x)] = 0 \quad \text{and} \quad \widetilde{\mathbb{E}}[q \cdot (\|x\|^2 - 1)] = 0$$

- **Positive semi-definiteness:** for any degree$\leq d/2$ polynomial q, $\widetilde{\mathbb{E}}[q^2] \geq 0$

Computed using an SDP with $n^{\mathcal{O}(d)}$ variables and constraints

Degree-d SoS algorithm

Computed using an SDP with $n^{\mathcal{O}(d)}$ variables and constraints

Variables: $\widetilde{\mathbb{E}}[x^S]$ $\forall\, S \subset [n] : |S| \le d$

Linear constraints:

- $\widetilde{\mathbb{E}}[q \cdot (x_1^2 + \cdots + x_n^2 - 1)] = 0$ $\forall\, q$ of degree at most $d-2$
- $\widetilde{\mathbb{E}}[q \cdot (\Pi(x \otimes x) - x \otimes x)] = 0$ $\forall\, q$ of degree at most $d-2$
- $\widetilde{\mathbb{E}}[1] = 1$

Linearity $\Rightarrow$ only checking $q(x) = x^T$ for $|T| \le d-2$

PSD constraint: $\mathcal{M}_d \in \mathbb{R}^{n^{\mathcal{O}(d)} \times n^{\mathcal{O}(d)}}$ defined as $\mathcal{M}_d[S,T] \coloneqq \widetilde{\mathbb{E}}[x^{S \cup T}]$ $\forall\, S, T \subset [n] : |S|, |T| \le d/2$

$$\mathcal{M}_d \succcurlyeq 0$$

Solvable in $n^{\mathcal{O}(d)}$ time using e.g., ellipsoid method

Rounding Pseudo-distribution

Q: Are degree-$\tilde{O}(\sqrt{n})$ pseudo-distributions enough for our rounding?

Is the restricted degree-$\tilde{O}(\sqrt{n})$ SoS positivity constraints enough?

There is a growing toolkit to show such statement

"*…if the inequality $f \geq 0$ is 'classical' and 'famous' enough, then f usually turns out to be representable as a sum of squares, although such a representation is not always easy to find.*"

(Frenkel-Horváth '14)

SoS toolkit

Cauchy-Schwarz inequality: If $\widetilde{\mathbb{E}}$ is a degree-d pseudoexpectation, then

$$\widetilde{\mathbb{E}}[p \cdot q] \leq \widetilde{\mathbb{E}}[p^2]^{1/2} \widetilde{\mathbb{E}}[q^2]^{1/2}$$

Proof.

- We may assume that $\widetilde{\mathbb{E}}[p^2] = \widetilde{\mathbb{E}}[q^2] = 1$
- SoS positivity $\Longrightarrow \widetilde{\mathbb{E}}[(p-q)^2] \geq 0 \Longrightarrow \widetilde{\mathbb{E}}[p \cdot q] \leq 1$

$\blacksquare$

Jensen inequality: If $\widetilde{\mathbb{E}}$ is a degree-d pseudoexpectation and p is of degree $\leq d/2$, then

$$\widetilde{\mathbb{E}}[p^2] \geq \widetilde{\mathbb{E}}[p]^2$$

Proof.

- Apply SoS Cauchy-Schwarz with $q = 1$

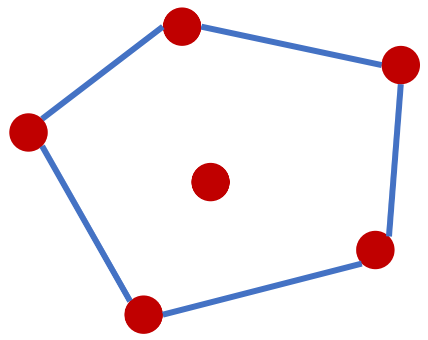
$\blacksquare$

Blueprint

Is there a rank-1 matrix in $\mathcal{W}$? → Is the polynomial system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

1. Choose the tightest convex "relaxation"
   → Feasible points = distributions over solutions
2. Show how to round any solution
   → "Bulk" of the proof
3. Dive in to find the constraints actually used
   → Obtain a degree-d Sum-of-Squares proof
4. Relax to a convex program by including only useful constraints
   → Optimize over degree-d pseudo-distributions



Full algorithm

1. Run Sum-of-Squares relaxation to obtain a degree-$\tilde{O}(\sqrt{n})$ pseudo-distribution $\tilde{\mu}$ over $\mathbb{S}^{n-1}$
2. Apply structure theorem to obtain a degree-$\tilde{O}(\sqrt{n})$ reweighing $\tilde{\mu}'$
3. Return rank-1 matrix on top eigenvector of $\tilde{\mathbb{E}}_{\tilde{\mu}'}[xx^\top]$

Find a rank-1 matrix in a subspace:

- **Input:** $\Pi \in \mathbb{R}^{n^2 \times n^2}$ the orthogonal projector for the subspace $\mathcal{W}$
- **Promise:** $\Pi(u \otimes u) = u \otimes u$ for $u \in \mathbb{S}^{n-1}$
- **Goal:** find $x \in \mathbb{S}^{n-1}$ such that $\|\Pi(x \otimes u)\|_F^2 > 0.99$

Theorem (Barak-Kothari-Steurer '17). For every $s < 1$, there is a $2^{\tilde{O}(\sqrt{n})}$-time algorithm for $BSS_{1,s}$, based on rounding an SoS relaxation

Blueprint

Is there a rank-1 matrix in $\mathcal{W}$?

→

Is the polynomial system feasible?

$$\begin{cases} \Pi(x \otimes x) = x \otimes x \\ \|x\|^2 = 1 \end{cases}$$

1. Choose the tightest convex "relaxation"
   → Feasible points = distributions over solutions
2. Show how to round any solution
   → "Bulk" of the proof (polynomial reweighing + structure theorem)
3. Dive in to find the constraints actually used
   → Obtain a degree-d Sum-of-Squares proof
4. Relax to a convex program by including only useful constraints
   → Optimize over degree-d pseudo-distributions