Polynomial optimization on the sphere and quantum entanglement testing

Kun Fang

Joint work with Hamza Fawzi

Presented at ICCOPT 2019, Berlin

UNIVERSITY OF CAMBRIDGE

Talk Outline

- Polynomial Optimization and SOS Hierarchy
- An improved Convergence Rate
  Main Result and Proof Strategy
- Relation to Entanglement Testing
  SOS Hierarchy (polynomial) v.s. DPS Hierarchy (quantum)
- Summary and Discussions

Polynomial Optimization on the Sphere

Polynomial Optimization on the Sphere

Given a multivariate polynomial $p(x)$ with $x = (x_1, \cdots, x_d)$

Computing the maximal value $p_{\max} = \max_{x \in S^{d-1}} p(x)$

Over the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d : x_1^2 + \cdots + x_d^2 = 1\}$

Applications:

- the largest stable/independent set of a graph
  - *Degree 3* polynomial opt. on the sphere (e.g. [Nesterov'03, De Klerk'08])
- 2 → 4 norm of a matrix A, $p(x) = \|Ax\|_4^4$
  - *Degree 4* polynomial opt. on the sphere (e.g. [Barak et al.'12])
- Best Separable State problem in quantum information theory
  - *Degree 4* polynomial opt. on the product of spheres (e.g. [Barak-Kothari-Steurer'17])
- ...

Polynomial Optimization on the Sphere

Q

Given a multivariate polynomial $p(x)$ with $x = (x_1, \cdots, x_d)$

Computing the maximal value $p_{\max} = \max_{x \in S^{d-1}} p(x)$

Over the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d : x_1^2 + \cdots + x_d^2 = 1\}$

Difficulty:

- Degree = 2, efficiently solved as an eigenvalue problem;
- Degree > 2, **NP-hard** in general!

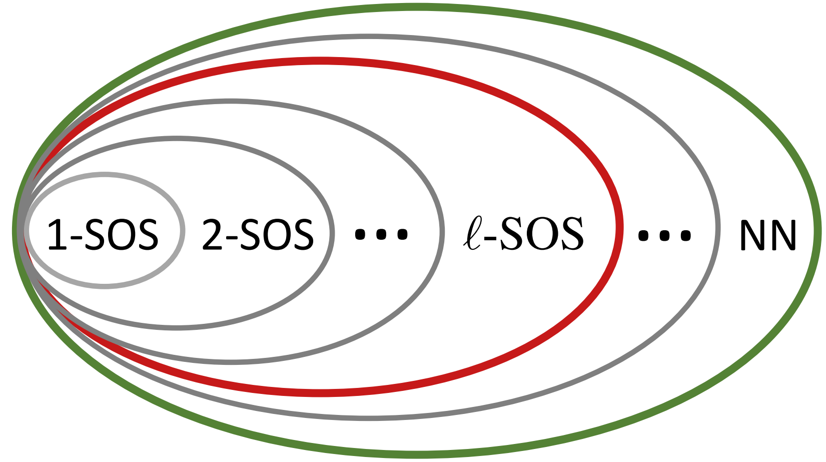
Solution:

- Sum-of-square (SOS) hierarchy [Parrilo'00; Lasserre'01]
  where each level is efficiently computable by semidefinite program

Sum-of-Square (SOS) Hierarchy

$$\ell\text{-SOS} = \Big\{ p(x) = \sum_i q_i(x)^2 \text{ on } S^{d-1} \quad \text{s.t. } \deg(q_i) \leq \ell \Big\}$$

polynomial polynomial



$$\mathrm{NN} = \big\{ p(x) : p(x) \geq 0, \forall x \in S^{d-1} \big\}$$

Relation with polynomial optimization:

$$p_{\max} = \max_{x \in S^{d-1}} p(x) = \min\{\gamma \in \mathbb{R} : \boxed{\gamma - p \in \mathrm{NN}} \text{ on } S^{d-1}\}$$

$$\leq \min \{\gamma \in \mathbb{R} : \boxed{\gamma - p \in \ell\text{-SOS}} \text{ on } S^{d-1}\}$$

restriction

$$= p_\ell$$ [SDP of size $d^{O(\ell)}$]

Approaching $p_{\max}$ from above:

Main Result: improved Convergence Rate

Q: How fast does p_ℓ converge to $p_{\max}$?

A: [Reznick'95; Doherty-Wehner'12], convergence rate at least $O(d/\ell)$

Q: Can we further sharpen the convergence rate?
(see recent works by de Klerk & Laurent 1811.05439 & 1904.08828)

A: *Positive answer in this work*, convergence rate at least $O((d/\ell)^2)$

Quadratic improvement

Main Result (technical statement)

Suppose $p(x_1, \cdots, x_d)$ is a homo. poly. of degree $2n$ in d variables with $n \le d$,

$$1 \le \frac{p_\ell - \boxed{p_{\min}}}{p_{\max} - \boxed{p_{\min}}} \le 1 + \left(\mathsf{C}_n \cdot \frac{d}{\ell}\right)^2 \quad \text{for all} \quad \ell \ge \boxed{\mathsf{C}_n} d$$

reference point

$$p_{\min} = \min_{x \in S^{d-1}} p(x)$$

A constant depends only on n.

A Stronger Result [take home message]

Matrix-valued polynomials: $x = (x_1, \cdots, x_d)$

$F(x) \in \mathbf{S}^k[x]$: k by k matrix with polynomial entries, symmetric for any x .

$F(x) \geq 0$: positive semidefinite matrix for any x .

$F(x) \in \ell\text{-SOS}$ if $F(x) = \sum_j U_j(x) U_j(x)^\mathsf{T}, \quad \deg(U_j) \leq \ell$

Remark:
- These definitions reduce to (scalar-valued) polynomial if k = 1;
- But the results cannot be trivially extended. (e.g. Nonnegative quadratic polynomial is necessarily a SOS. But not true for matrix-valued case.)

For any homo. matrix-valued poly. $F(x) \in \mathbf{S}^k[x]$ of <u>degree $2n$</u> in <u>d variables</u> with $n \leq d$ and $0 \leq F(x) \leq I$ for all $x \in S^{d-1}$,

$$F + \boxed{\mathsf{C}'_n \left(\frac{d}{\ell}\right)^2} \cdot I \in \ell\text{-SOS on } S^{d-1} \quad \text{for all } \ell \geq \mathsf{C}_n d$$

A small perturbation δ Identity

A remarkable fact: the result is totally independent on the size of the matrix F(x).

A Stronger Result

Main Result (technical statement)

Suppose $p(x_1, \cdots, x_d)$ is a homo. poly. of degree $2n$ in d variables with $n \leq d$,

$$1 \leq \frac{p_\ell - p_{\min}}{p_{\max} - p_{\min}} \leq 1 + \left(\mathsf{C}_n \cdot \frac{d}{\ell}\right)^2 \quad \text{for all } \ \ell \geq \mathsf{C}_n d$$

$$F = \frac{p_{\max} - p}{p_{\max} - p_{\min}}$$

For any homo. matrix-valued poly. $F(x) \in \mathbf{S}^k[x]$ of degree $2n$ in d variables with $n \leq d$ and $0 \preceq F(x) \preceq I$ for all $x \in S^{d-1}$,

$$F + \boxed{\mathsf{C}'_n \left(\frac{d}{\ell}\right)^2} \cdot I \in \ell\text{-SOS on } S^{d-1} \quad \text{for all } \ \ell \geq \mathsf{C}_n d$$

A small perturbation δ

Identity

A remarkable fact: the result is totally independent on the size of the matrix F(x).

Convergence Rate

Proof Outline

Proof Outline

Goal: Given poly. $0 \leq F \leq 1$. Find $\delta > 0$ such that $\widetilde{F} = F + \delta$ is ℓ-SOS

How to construct SOS:

For any polynomial: $q(t) : [-1,1] \to \mathbb{R}$ of $\deg(q) = \ell$, consider $K(x,y) = q(\langle x, y\rangle)^2$

$$\underline{(Kh)}(x) = \int_{y \in S^{d-1}} K(x,y)\underline{h(y)}d\sigma(y) \qquad \forall x \in S^{d-1}$$

$$= \underbrace{\int_{y \in S^{d-1}}}_{\text{sum}} \underbrace{q(\langle x, y\rangle)^2}_{\text{Poly^2}} \underbrace{h(y)\, d\sigma(y)}_{\text{nonnegative weight}} \qquad \forall x \in S^{d-1}$$

Key observation: $h(y) \geq 0 \implies Kh \in \ell\text{-SOS}$ [Reznick'95; Doherty-Wehner'12 ;Parrilo'13]

$\widetilde{F} \in \ell\text{-SOS}$ $\Longleftarrow$ $K^{-1}\widetilde{F} \geq 0$ $\Longleftarrow$ $\boxed{\|K^{-1}\widetilde{F} - \widetilde{F}\|_\infty \leq \delta}$

$\widetilde{F} = K(\underline{K^{-1}\widetilde{F}})$

$\widetilde{F} = F + \delta \geq \delta$

$\widetilde{F} = F + \delta$ is ℓ-SOS $\Leftarrow$ $\|K^{-1}\widetilde{F} - \widetilde{F}\|_\infty \le \delta$

Estimate $\|K^{-1}\widetilde{F} - \widetilde{F}\|_\infty \le \delta$ Rotation invariant kernel $K(x,y) = q(\langle x, y\rangle)^2$

Some well-studied results

Harmonic Decomposition $F = F_0 + F_2 + \cdots F_{2n}$ $\widetilde{F} = (F_0 + \delta) + F_2 + \cdots F_{2n}$

Kernel Decomposition $\phi = q^2 = \lambda_0 C_0 + \lambda_1 C_1 + \cdots \lambda_{2\ell} C_{2\ell}$ [C_k Gegenbauer]

Funk-Hecke formula $K\widetilde{F} = \lambda_0(F_0 + \delta) + \lambda_2 F_2 + \cdots \lambda_{2n} F_{2n}$

$$K^{-1}\widetilde{F} = \lambda_0^{-1}(F_0 + \delta) + \lambda_2^{-1} F_2 + \cdots \lambda_{2n}^{-1} F_{2n}$$

$$\|K^{-1}(\widetilde{F}) - \widetilde{F}\|_\infty = \left\| \sum_{k=1}^{n} \left(\frac{1}{\lambda_{2k}} - 1 \right) F_{2k} \right\|_\infty \le \underline{\sum_{k=1}^{n} \left| \frac{1}{\lambda_{2k}} - 1 \right|} \, \underline{\|F_{2k}\|_\infty}$$

- Since $0 \le F \le 1$, its harmonic components won't be too large $\|F_{2k}\|_\infty \le B_{2n}\|F\|_\infty$
- Estimate $\sum_{k=1}^{n} \left| \frac{1}{\lambda_{2k}} - 1 \right|$ $\Leftarrow$ $\sum_{k=1}^{n} \left| \frac{1}{\lambda_{2k}} - 1 \right| \le 2 \boxed{\sum_{k=1}^{n} (1 - \lambda_{2k}),}$

$\ell \ge 2nd$

$\widetilde{F} = F + \delta$ is ℓ-SOS $\Leftarrow$ $\|K^{-1}\widetilde{F} - \widetilde{F}\|_\infty \le \delta$ $\Leftarrow$ $\sum_{k=1}^{n}(1-\lambda_{2k}) \le \delta$

Estimate $\sum_{k=1}^{n}(1-\lambda_{2k}) \le \delta$

$$\lambda_i = \frac{\omega_{d-1}}{\omega_d}\int_{-1}^{1}\phi(t)\frac{C_i(t)}{C_i(1)}(1-t^2)^{\frac{d-3}{2}}dt \qquad \phi = q^2 = \lambda_0 C_0 + \lambda_1 C_1 + \cdots \lambda_{2\ell}C_{2\ell}$$

- If we choose polynomial $q(t) \propto t^\ell$, each coefficient can be computed explicitly. Observe that λ_i scales as $O(d/\ell)$. Recover results by [Reznick'95; Doherty-Wehner'12].
- **To obtain a better result, we do not choose specific $q(t)$ at this moment.**

$$\phi(t) = [q(t)]^2 = \left[\sum_{i=0}^{\ell} e_i \frac{C_i(t)}{\sqrt{C_i(1)}}\right]^2 \qquad e = [e_0\ e_1\ \cdots\ e_\ell]^\mathsf{T}$$

$$\lambda_{2k} = e^\mathsf{T}\underline{\mathcal{T}[C_{2k}/C_{2k}(1)]}e \qquad \mathcal{T}[g]_{i,j} = \frac{\omega_{d-1}}{\omega_d}\int_{-1}^{1}\frac{C_i(t)}{\sqrt{C_i(1)}}\frac{C_j(t)}{\sqrt{C_j(1)}}g(t)(1-t^2)^{\frac{d-3}{2}}dt$$

Generalized Toeplitz matrix

$$\sum_{k=1}^{n}(1-\lambda_{2k}) = n\left(1 - e^\mathsf{T}\mathcal{T}[h]\,e\right) \Leftarrow \boxed{1-\lambda_{\max}(\mathcal{T}[h]) \le \delta}$$

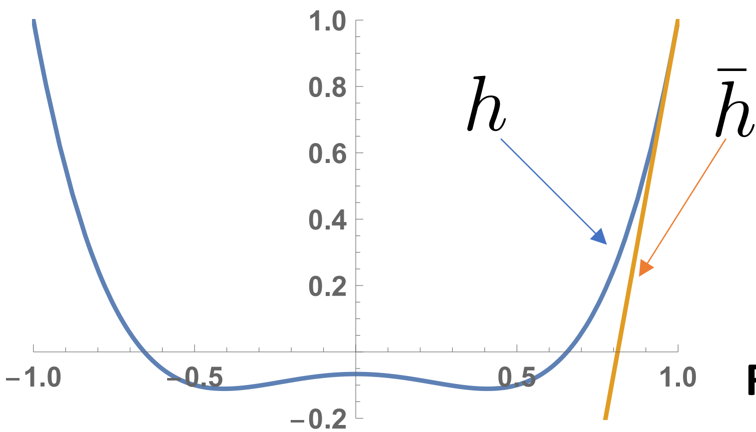
$$h = \frac{1}{n}\sum_{k=1}^{n}\frac{C_{2k}}{C_{2k}(1)}$$

$\boxed{\widetilde{F} = F + \delta \text{ is } \ell\text{-SOS}} \Leftarrow \boxed{\|K^{-1}\widetilde{F} - \widetilde{F}\|_\infty \le \delta} \Leftarrow \boxed{\sum_{k=1}^{n}(1-\lambda_{2k}) \le \delta} \Leftarrow \boxed{1 - \lambda_{\max}(\mathcal{T}[h]) \le \delta}$

Estimate $1 - \lambda_{\max}(\mathcal{T}[h]) \le \delta \qquad h = \frac{1}{n}\sum_{k=1}^{n} \frac{C_{2k}}{C_{2k}(1)}$

- For linear polynomial f, we have a good understanding of the eigenvalues of $\mathcal{T}[f]$.
- h is non-linear -> Consider the tangent line at t = 1, $\bar{h}(t) = h'(1)(t-1) + h(1)$

$$\lambda_{\max}(\mathcal{T}[h]) \ge \lambda_{\max}(\mathcal{T}[\bar{h}]) = \bar{h}(x_{\ell+1,\ell+1}) \ge 1 - \underbrace{\boxed{\frac{7n}{12}\frac{d^2}{\ell^2}}}_{\delta}$$



d = 4, n = 2

The largest root of $C_{\ell+1}$, $x_{\ell+1,\ell+1} \ge 1 - \frac{1}{4}\frac{d^2}{\ell^2}$

This completes the proof.

Remark: The estimation of $\lambda_{\max}(\mathcal{T}[h])$ is tight.

Some Remarks

Recall the result

For any homo. matrix-valued poly. $F(x) \in \mathbf{S}^k[x]$ of $\underline{\text{degree } 2n}$ in $\underline{d \text{ variables}}$ with $n \leq d$ and $0 \leq F(x) \leq I$ for all $x \in S^{d-1}$

$$F + \mathrm{C}_n' \left(\frac{d}{\ell}\right)^2 \cdot I \in \ell\text{-SOS on } S^{d-1} \quad \text{for all } \ell \geq \mathrm{C}_n d$$

Some Remarks:

- The proof works for *matrix-valued polynomials*.
- The proof works for polynomials on the *complex sphere*.
- We can estimate the convergence for *all values of level, not just* $\ell \geq \mathrm{C}_n d$.

Entanglement Testing

Quantum States

Quantum state $\mathcal{S}(\mathcal{H}_A) = \left\{ \sum_i p_i x_i x_i^\dagger : p_i \geq 0, x_i \in \mathcal{H}_A \right\}$

Separable state $\mathcal{SEP}(\mathcal{H}_A \otimes \mathcal{H}_B) = \left\{ \sum_i p_i (x_i x_i^\dagger) \otimes (y_i y_i^\dagger) : p_i \geq 0, x_i \in \mathcal{H}_A, y_i \in \mathcal{H}_B \right\}$

Entangled state Any quantum state that is *not separable*

Q: Whether a given quantum state ρ_{AB} is entangled or not?

A: Doherty-Parrilo-Spedalieri (DPS) hierarchy

$$\mathcal{DPS}_\ell(\mathcal{H}_A \otimes \mathcal{H}_B) = \{\rho_{AB} : \exists \rho_{AB_1 \cdots B_\ell} \text{ s.t. } (1,2,3) \text{ holds}\}$$

1. Reduction under partial trace:

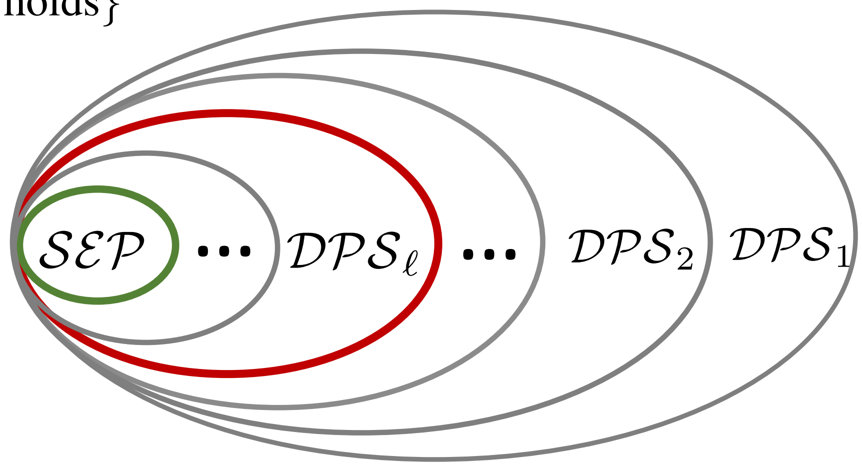
$$\mathrm{Tr}_{B_2 \cdots B_\ell}[\rho_{AB_1 \underline{B_2 \cdots B_\ell}}] = \rho_{AB}$$

2. Symmetry on B systems:

$$(I \otimes \Pi_{B_1 \cdots B_\ell}) \rho_{\underline{AB_1 \cdots B_\ell}} (I \otimes \Pi_{B_1 \cdots B_\ell}) = \rho_{AB_1 \cdots B_\ell}$$

3. Positive partial transpose (PPT):

$$(I_A \otimes \boxed{\mathsf{T}_{B_1} \otimes \cdots \mathsf{T}_{B_s}} \otimes I_{B_{i+1}} \otimes I_{B_\ell})(\rho_{AB_1 \cdots B_\ell}) \geq 0$$



Form a complete hierarchy [Doherty-Parrilo-Spedalieri'02&04]

Quantum States

Quantum state $\mathcal{S}(\mathcal{H}_A) = \left\{ \sum_i p_i x_i x_i^\dagger : p_i \geq 0, x_i \in \mathcal{H}_A \right\}$

Separable state $\mathcal{SEP}(\mathcal{H}_A \otimes \mathcal{H}_B) = \left\{ \sum_i p_i (x_i x_i^\dagger) \otimes (y_i y_i^\dagger) : p_i \geq 0, x_i \in \mathcal{H}_A, y_i \in \mathcal{H}_B \right\}$

Entangled state Any quantum state that is *not separable*

Q: Whether a given quantum state ρ_{AB} is entangled or not?

A: Doherty-Parrilo-Spedalieri (DPS) hierarchy

$\mathcal{DPS}_\ell(\mathcal{H}_A \otimes \mathcal{H}_B) = \{\rho_{AB} : \exists \rho_{AB_1 \cdots B_\ell} \text{ s.t. } (1,2,3) \text{ holds}\}$

1. Reduction under partial trace:

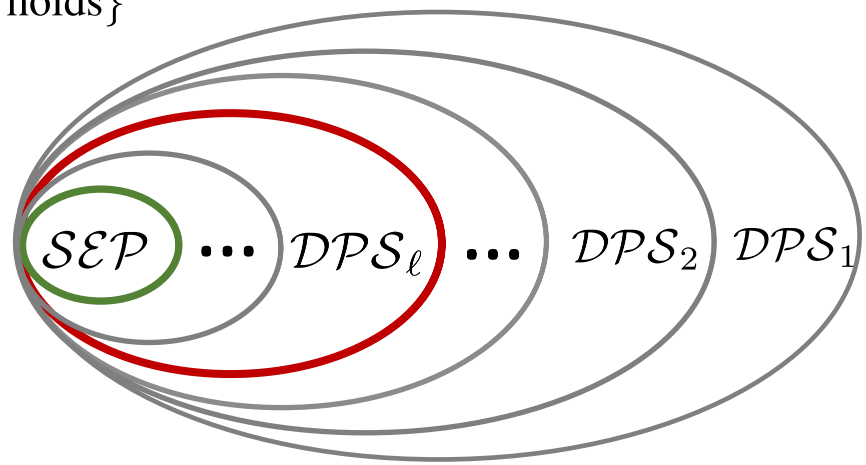
$$\mathrm{Tr}_{B_2 \cdots B_\ell}[\rho_{AB_1 B_2 \cdots B_\ell}] = \rho_{AB}$$

2. Symmetry on B systems:

$$(I \otimes \Pi_{B_1 \cdots B_\ell}) \rho_{AB_1 \cdots B_\ell} (I \otimes \Pi_{B_1 \cdots B_\ell}) = \rho_{AB_1 \cdots B_\ell}$$

3. Positive partial transpose (PPT):

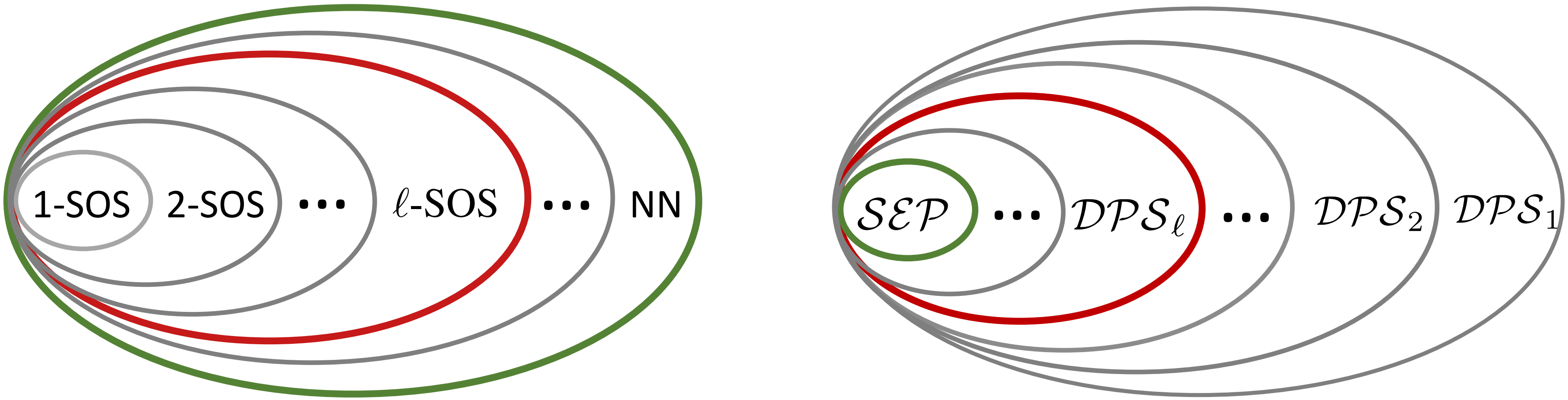
$$(I_A \otimes \boxed{\mathsf{T}_{B_1} \otimes \cdots \mathsf{T}_{B_s}} \otimes I_{B_{i+1}} \otimes I_{B_\ell})(\rho_{AB_1 \cdots B_\ell}) \geq 0$$



Without PPT conditions $\mathcal{EXT}_\ell(\mathcal{H}_A \otimes \mathcal{H}_B) = \{\rho_{AB} : \exists \rho_{AB_1 \cdots B_\ell} \text{ s.t. } (1,2) \text{ holds}\}$

This is a weaker hierarchy but it is still complete.

Duality Relation



SOS hierarchy (polynomials) $\xleftrightarrow{\text{dual ?}}$ DPS hierarchy (quantum states)

For any Hermitian operator M on $\mathcal{H}_A \otimes \mathcal{H}_B$, define its associated Hermitian polynomial

$$p_M(x,\bar{x},y,\bar{y}) = (x\otimes y)^\dagger M(x\otimes y) = \sum_{i,j,k,l} M_{ij,kl}x_i\bar{x}_k y_j \bar{y}_\ell \quad \forall x\in\mathbb{C}^{d_A}, y\in\mathbb{C}^{d_B}$$

Duality relation

$$\mathcal{SEP}^* = \{M\in \mathrm{Herm}(\mathcal{H}_A\otimes\mathcal{H}_B) : p_M \text{ is nonnegative}\}$$

$$\mathcal{DPS}_\ell^* = \left\{M\in \mathrm{Herm}(\mathcal{H}_A\otimes\mathcal{H}_B) : \|y\|^{2(\ell-1)}p_M \text{ is rSOS}\right\} \dashrightarrow \sum_i q_i(x,\bar{x},y,\bar{y})^2$$

$$\mathcal{EXT}_\ell^* = \left\{M\in \mathrm{Herm}(\mathcal{H}_A\otimes\mathcal{H}_B) : \|y\|^{2(\ell-1)}p_M \text{ is cSOS}\right\} \dashrightarrow \sum_i |g_i(x,y)|^2$$

PPT conditions determine the choice of monomials in the SOS decomposition.

Relation with [NOP'09]

[Navascues-Owari-Plenio'09]

For any quantum state $\rho_{AB} \in \mathrm{DPS}_\ell$ with reduced state $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$

$$(1-t)\rho_{AB} + \boxed{t\rho_A \otimes \frac{I_B}{d_B}} \text{ is seperable} \quad \text{with} \quad t = O\left(\frac{d_B^2}{\ell^2}\right)$$

A small perturbation

Utilizing the duality between DPS and SOS, this is equivalent to our result of matrix-valued polynomial with **degree 2**.

Recall our result in this work

For any homo. matrix-valued poly. $F(x) \in \mathbf{S}^k[x]$ of degree $2n$ in d variables with $n \leq d$ and $0 \leq F(x) \leq I$ for all $x \in S^{d-1}$

$$F + \boxed{\mathsf{C}'_n \left(\frac{d}{\ell}\right)^2 \cdot I} \in \ell\text{-SOS on } S^{d-1} \qquad \text{for all } \ell \geq \mathsf{C}_n d$$

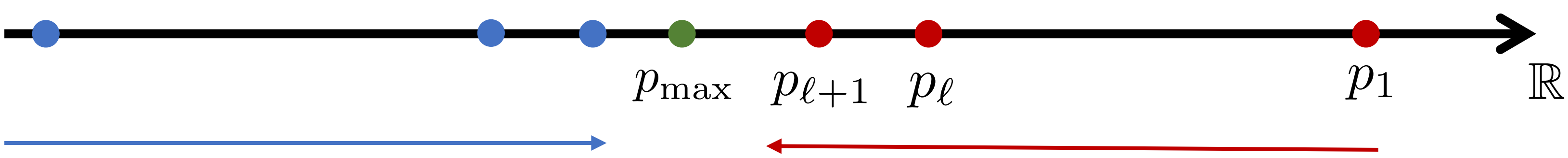
A small perturbation

Summary & Discussions

Summary

- **An quadratic improvement of convergence rate of the SOS hierarchy**
  - Works for matrix-valued polynomials
  - Works for complex variables
  - Works for all values of the level
- **Exact Duality relation between SOS and DPS hierarchies**
- **Connection with [Navascues-Owari-Plenio'09] from quantum community**

Other related works:



- Lasserre hierarchy appr. from below
  [de Klerk-Laurent 1904.08828]

SOS hierarchy appr. from above
[Fang-Fawzi-This work]
Empirically much faster

- Analysis of the SOS hierarchy from the computer science community
  (e.g. [Bhattiprolu et al.'17; Barak-Kothari-Steurer'17])

Question: Further sharpening the convergence rate? New techniques are required.

Thanks for your attention!

Full paper will be online soon.