Polynomial optimization on the sphere and quantum entanglement testing

(Full paper will be online soon)

Kun Fang

Joint work with Hamza Fawzi

Presented at QAS 2019, Shenzhen
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
Given a multivariate polynomial \( p(x) \) with \( x = (x_1, \cdots, x_d) \)

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

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
  - \textit{Degree 3} polynomial opt. on the sphere (e.g. [Nesterov’03, De Klerk’08])

- \( 2 \rightarrow 4 \) norm of a matrix \( A \), \( p(x) = \|Ax\|_4^4 \)
  - \textit{Degree 4} polynomial opt. on the sphere (e.g. [Barak et al.’12])

- Best Separable State problem in quantum information theory
  - \textit{Degree 4} polynomial opt. on the product of spheres (e.g. [Barak-Kothari-Steurer’17])

- ...
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
\[ \ell\text{-SOS} = \left\{ p(x) = \sum_i q_i(x)^2 \text{ on } S^{d-1} \right\} \]

\[ \text{s.t. } \deg(q_i) \leq \ell \}

\[ \text{NN} = \{ p(x) : p(x) \geq 0, \forall x \in S^{d-1} \} \]

**Relation with polynomial optimization:**

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

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

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

**Approaching \( p_{\max} \) from above:**

\[ p_{\max} \quad p_{\ell+1} \quad p_\ell \quad p_1 \]
Main Result: improved Convergence Rate

**Q:** How fast does $p_\ell$ converge to $p_{\text{max}}$?

**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 $O((d/\ell)^2)$

---

**Main Result (technical statement)**

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

$$1 \leq \frac{p_\ell - p_{\text{min}}}{p_{\text{max}} - p_{\text{min}}} \leq 1 + \left( C_n \cdot \frac{d}{\ell} \right)^2$$

for all $\ell \geq C_n d$

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

reference point

A constant depends only on $n$. 

---

Quadratic improvement
A Stronger Result [take home message]

Matrix-valued polynomials:

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

\[ F(x) \geq 0 : \text{positive semidefinite matrix for any } x. \]

\[ F(x) \in \ell\text{-SOS} \quad \text{if } F(x) = \sum_j U_j(x)U_j(x)^\top, \quad \text{deg}(U_j) \leq \ell \]

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. \( F(x) \in S^k[x] \) of degree \( 2n \) in \( d \) variables with \( n \leq d \) and \( 0 \leq F(x) \leq I \) for all \( x \in S^{d-1} \),

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

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 $n \leq d$, $1 \leq \frac{p_{\ell} - p_{\min}}{p_{\max} - p_{\min}} \leq 1 + \left( C_n \cdot \frac{d}{\ell} \right)^2$ for all $\ell \geq C_n d$

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

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

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

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 $\tilde{F} = F + \delta$ is $\ell$-SOS

**How to construct SOS:**

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

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

$$
= \int_{y \in S^{d-1}} q(\langle x, y \rangle)^2 h(y) d\sigma(y) \quad \forall x \in S^{d-1}
$$

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

\[ \tilde{F} \in \ell\text{-SOS} \quad \xrightarrow{\text{Key observation}} \quad K^{-1}\tilde{F} \geq 0 \quad \xrightarrow{\text{Key observation}} \quad \|K^{-1}\tilde{F} - \tilde{F}\|_{\infty} \leq \delta \]
Estimate $\|K^{-1}\tilde{F} - \tilde{F}\|_{\infty} \leq \delta$

Rotation invariant kernel $K(x, y) = q(\langle x, y \rangle)^2$

Some well-studied results

Harmonic Decomposition

$F = F_0 + F_2 + \cdots F_{2n}$

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

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

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

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

Since $0 \leq F \leq 1$, its harmonic components won’t be too large $\|F_{2k}\|_{\infty} \leq B_{2n} \|F\|_{\infty}$

Estimate $\sum_{k=1}^{n} \left| \frac{1}{\lambda_{2k}} - 1 \right| \leq 2 \sum_{k=1}^{n} (1 - \lambda_{2k})$, $\ell \geq 2nd$
Estimate \( \sum_{k=1}^{n} (1 - \lambda_{2k}) \leq \delta \)

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

- If we choose polynomial \( q(t) \propto t^\ell \), each coefficient can be computed explicitly. Observe that \( \lambda_i \) 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.

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

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

**Generalized Toeplitz matrix**

\[
\sum_{k=1}^{n} (1 - \lambda_{2k}) = n \left( 1 - e^T \mathcal{T} [h] e \right) \\
1 - \lambda_{\max}(\mathcal{T}[h]) \leq \delta
\]

\[
h = \frac{1}{n} \sum_{k=1}^{n} \frac{C_{2k}}{C_{2k}(1)}
\]
\( \hat{F} = F + \delta \) is \( \ell \)-SOS \( \| K^{-1} \hat{F} - F \|_\infty \leq \delta \) \( \sum_{k=1}^{n} (1 - \lambda_{2k}) \leq \delta \) \( 1 - \lambda_{\text{max}}(T[h]) \leq \delta \)

Estimate \( 1 - \lambda_{\text{max}}(T[h]) \leq \delta \)

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

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

\[
\lambda_{\text{max}}(T[h]) \geq \lambda_{\text{max}}(T[\bar{h}]) = \bar{h}(x_{\ell+1,\ell+1}) \geq 1 - \frac{7n d^2}{12 \ell^2}
\]

This completes the proof.
Some Remarks

Recall the result

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

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

Some Remarks:

- The proof works for matrix-valued polynomials.
- The proof works for polynomials on the complex sphere.
- We can estimate the perturbation for all values of level, not just $\ell \geq C_n d$. 
Entanglement Testing
Quantum States

**Quantum state**

\[ S(\mathcal{H}_A) = \left\{ \sum_i p_i x_i x_i^\dagger : p_i \geq 0, x_i \in \mathcal{H}_A \right\} \quad \text{unnormalized} \]

**Separable state**

\[ S\mathcal{EP}(\mathcal{H}_A \otimes \mathcal{H}_B) = \left\{ \sum_i p_i (x_i x_i^\dagger) \otimes (y_i y_i^\dagger) : p_i \geq 0, x_i \in \mathcal{H}_A, y_i \in \mathcal{H}_B \right\} \]

**Entangled state**

Any quantum state that is *not separable*

### Q:

Whether a given quantum state \( \rho_{AB} \) is entangled or not?

### A:

Doherty-Parrilo-Spedalieri (DPS) hierarchy

\[ \mathcal{DPS}_\ell(\mathcal{H}_A \otimes \mathcal{H}_B) = \{ \rho_{AB} : \exists \rho_{A1\ldots B_\ell} \text{ s.t. } (1, 2, 3) \text{ holds} \} \]

1. Reduction under partial trace:

\[ \operatorname{Tr}_{B_2\ldots B_\ell}[\rho_{AB_1B_2\ldots B_\ell}] = \rho_{AB} \]

2. Symmetry on B systems:

\[ (I \otimes \Pi_{B_1\ldots B_\ell})\rho_{AB_1\ldots B_\ell}(I \otimes \Pi_{B_1\ldots B_\ell}) = \rho_{AB_1\ldots B_\ell} \]

3. Positive partial transpose (PPT):

\[ (I_A \otimes T_{B_1} \otimes \cdots T_{B_s} \otimes I_{B_{i+1}} \otimes I_{B_\ell})(\rho_{AB_1\ldots B_\ell}) \geq 0 \]

Form a complete hierarchy [Doherty-Parrilo-Spedalieri’02&04]
Quantum States

**Quantum state**

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

**Separable state**

\[ S\mathcal{E}\mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B) = \left\{ \sum_i p_i (x_i x_i^\dagger) \otimes (y_i y_i^\dagger) : p_i \geq 0, x_i \in \mathcal{H}_A, y_i \in \mathcal{H}_B \right\} \]

**Entangled state**

Any quantum state that is not separable

**Q:** Whether a given quantum state \( \rho_{AB} \) is entangled or not?

**A:** Doherty-Parrilo-Spedalieri (DPS) hierarchy

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

1. Reduction under partial trace:

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

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):

\[ (I_A \otimes T_{B_1} \otimes \cdots T_{B_s} \otimes I_{B_{i+1}} \otimes I_{B_\ell})(\rho_{AB_1 \cdots B_\ell}) \geq 0 \]

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

This is a weaker hierarchy but it is still complete.
Duality relation

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

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

**SEP**

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

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

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

PPT conditions determine the choice of monomials in the SOS decomposition.
For any quantum state \( \rho_{AB} \in \text{DPS}_\ell \) with reduced state \( \rho_A = \text{Tr}_B[\rho_{AB}] \)

\[
(1 - t)\rho_{AB} + t\rho_A \otimes \frac{I_B}{d_B} \text{ is separable with } t = O\left(\frac{d_B^2}{\ell^2}\right)
\]

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. \( F(x) \in S^k[x] \) of degree \( 2n \) in \( d \) variables with \( n \leq d \) and \( 0 \leq F(x) \leq I \) for all \( x \in S^{d-1} \)

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

A small perturbation
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.
Open Quantum Problems List (https://oqp.iqoqi.univie.ac.at/)

**Problem 38: The PPT-squared conjecture** (Matthias Christandl, 2012)

**Conjecture:**

- **PPT channel:** completely positive and completely co-positive
- **Entanglement breaking channel:**
  \[ \text{id}_R \otimes M_{A \rightarrow B}(\rho_{RA}) \in \text{SEP} \quad \forall \rho_{RA} \]

**Quantum (key) repeater**

A new perspective, PPT Square Conjecture is equivalent to:

\[ \mathcal{N} \text{ PPT, } M \text{ positive map} \quad \Rightarrow \quad \text{Tr} \left[ \mathcal{N}(|x\rangle\langle x|) \right] \left[ M(|y\rangle\langle y|) \right] \text{ is SOS} \]

\[ \mathcal{N} \text{ positive map, } M \text{ positive map} \quad \Rightarrow \quad \text{Tr} \left[ \mathcal{N}(|x\rangle\langle x|) \right] \left[ M(|y\rangle\langle y|) \right] \text{ is nonnegative} \]
Thanks for your attention!

Full paper will be online soon.