Positive spectrahedra:

Invariance principles and PRGs

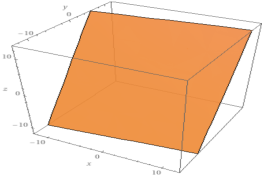
Srinivasan Arunachalam (IBM Research)

joint with Penghui Yao (Nanjing University)

- Optimizing a linear function over a polytope
- A general LP has the form: $w^1, \ldots, w^k, c \in \mathbb{R}^n$ and ${}^i \in \mathbb{R}$

$$\mathrm{OPT} = \max_{x \in \mathbb{R}^n} \{c^T x : w^1 \;\; x \;\;\; {}^1, \ldots, w^k \;\; x \;\;\; {}^k\}$$

- Efficiently solvable!



- Halfspace in $\mathbb{R}^n$ is a constraint that divides the space, i.e., $h_1 : \mathbb{R}^n \to \{0, 1\}$
- Let $w \in \mathbb{R}^n$ and $x \in \{\;1, 1\}^n$, then a halfspace $h_1(x) = 1$ i $\; w \;\; x \;\;$, or $h_1(x) = \;\; w \;\; x$

- **Polytope** is an intersection of halfspaces
- Let $w^i \in \mathbb{R}^n$, ${}^i \in \mathbb{R}$, a k-facet polytope is

$$P = \;\; x : h_1(x) \wedge h_2(x) \wedge \quad \wedge h_k(x)\}$$

where $h_i = [w^i \;\; x \;\;\; {}^i]$

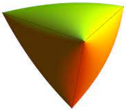
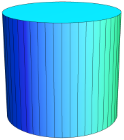
Applications: Optimization, combinatorics, geometry, computational complexity, ...

What is an SDP?

- Optimizing a linear function over a spectrahedron
- A general SDP has the form: $W^1, \dots, W^k, B \in \mathrm{Sym}_n$

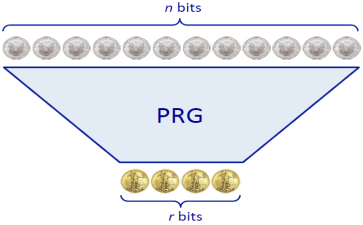
$$\mathrm{OPT} = \max_{x \in \mathbb{R}^n} \{ c^T x : x_1 W^1 + \cdots + x_n W^n \preceq B \},$$

where $C \preceq D$ means $D - C$ is PSD (i.e., all eigenvalues are ≥ 0)

How does it look?  

- Generalizes linear programs and still efficiently solvable!
- Unfortunately, spectrahedra are not very well understood!
- But SDPs have found applications in approximation theory, SoS heirarchy, quantum computing

A PRG is a function that "expands" randomness



PRGs for a class of functions

An ε-PRG for $\mathcal{C} \subseteq \{F : \{0,1\}^n \to \{0,1\}\}$ is a function $G : \{0,1\}^r \to \{0,1\}^n$ such that

$$\text{for every } F \in \mathcal{C}, \quad \left| \mathop{\mathbb{E}}_{x \sim U_r}[F(G(x))] - \mathop{\mathbb{E}}_{u \sim U_n}[F(u)] \right| \leq \varepsilon.$$

The seed length of G is r. **Goal** is to have r = polylog( ) in all relevant parameters

Holy grail. Can we design a PRG against the class of polynomial sized circuits unconditionally? If so, would imply BPP = P

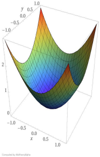
PRGs for geometric objects. Constructing PRGs using geometric properties has been a *rich* area of study this work.

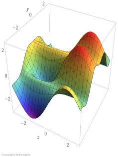
Halfspaces

- Diakonikolas et al.'09
- Meka, Zuckerman'09
- Karnin, Rabani, Shpilka'11
- Kothari, Meka'15
- Gopalan, Kane, Meka'15

Polytopes

- Harsha, Klivans, Meka'13
- Gopalan et al.'13
- Servedio, Tan'17
- O'Donnell, Servedio, Tan'19





Polynomial Threshold function

$T(x) = \mathrm{sign}(p(x_1, \ldots, x_n))$ where p is a polynomial

- Meka, Zuckerman'09
- Diakonikolas'10
- Kane'11, Kane'12, Kane'13
- Kane, Meka'14
- O'Donnell, Servedio, Tan'20

Spectrahedra: generalization of halfspaces, polytopes and PTFs in one framework

In this work: Can we construct PRGs for spectrahedra?

Recall. Spectrahedron is the set $S = \left\{ x \in \{-1,1\}^n : \sum_i x_i A^i \preceq B \right\}$.

1. Positive: $A^1, \ldots, A^n, B$ are $k \times k$ PSD matrices
2. Bounded width: $I \preceq \sum_i (A^i)^2 \preceq M \cdot I$
3. Regular: $(A^i) \preceq \tau \cdot I$ for every i

Main Theorem

There exists a PRG $G : \{0,1\}^r \to \{-1,1\}^n$ with seed length

$$r = (\log n) \cdot \mathrm{poly}(\log k, M, 1/\varepsilon);$$

that ε-fools the class of positive bounded width regular spectrahedron S, i.e.,

$$\left| \mathop{\mathbb{E}}_{x \sim U_r}[G(x) \in S] - \mathop{\mathbb{E}}_{u \sim U_n}[u \in S] \right| \leq \varepsilon.$$

Main technical contributions: Rest of this talk

$\boxed{\text{An invariance principle for } \textit{positive} \text{ regular spectrahedra}}$

Punchline: Invariance principles give pseudorandom generators.

What is an invariance principle? Generalization of Berry-Esseen theorem

Standard central limit theorem states: suppose $x_1, \ldots, x_n$ are random variables satisfying $\mathrm{E}[x] = 0$ and $\mathrm{Var}[x^2] = 1$, then

$$\frac{x_1 + \cdots + x_n}{\sqrt{n}} \to g(0,1),$$

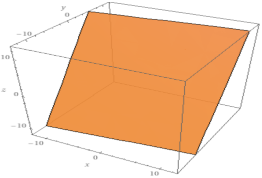
where $g(0,1)$ is a Gaussian

But what about convergence? Berry-Esseen states that for every $u \in \mathbb{R}$

$$\left| \Pr\left[\frac{x_1 + \cdots + x_n}{\sqrt{n}} \le u\right] - \Pr\left[g(0,1) \le u\right] \right| \le \frac{C}{\sqrt{n}},$$

for "C-nice" $x_1, \ldots, x_n$. Proved using the Lindeberg method'22 (aka hybrid method)

Invariance principles: understanding in the Gaussian space is similar to Boolean space



Halfspace (Meka-Zuckerman'09)

- Halfspace is
  $x \in \{\ 1, 1\}^n : \sum_i w_i x_i$
- For smooth $w \in \mathbb{R}^n$
  $$\sum_i w_i x_i \to g(0; 1)$$

Polytope (Harsha-Klivans-Meka'13)

- Polytope is
  $$x \in \{\ 1, 1\}^n : w^1 x \quad _1; \ldots; w^k x \quad _k$$
- Let $w^1; \ldots; w^k \in \mathbb{R}^n$ all be smooth, then
  $$\begin{bmatrix} — & w^1 & — \\ — & w^2 & — \\ & \vdots & \\ — & w^k & — \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \to \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_k \end{bmatrix}$$
- Recently OST'19 removed regularity

$\boxed{\text{Recall: Polytope } F(x) = \quad w^1 \quad x \qquad {}_1 \wedge \qquad \wedge\, w^k \quad x \qquad {}_k \text{ or } W \quad x \qquad \tilde{}}$

Main result of HKM'13 Invariance principle for -regular polytopes (i.e., $kw^i k$)

$$\underset{x}{} \underset{U_n}{\mathrm{E}}[Wx \quad \tilde{}\,] \qquad \underset{g}{} \underset{G^n}{\mathrm{E}}[Wg \quad \tilde{}\,] \qquad \text{polylog}\, k \tag{1}$$

How to prove this?

1. Smooth invariance. Establish (1) for smooth functions $O : \mathrm{R}^k \; ! \; \mathrm{R}$, i.e.,

$$\underset{x}{} \underset{U_n}{\mathrm{E}}[O(Wx)] \qquad \underset{g}{} \underset{G^n}{\mathrm{E}}[O(Wg)] \qquad \log k \quad kO^{(3)}k_1$$

- Lindeberg method: Write out Taylor series for $O : \mathrm{R}^k \; ! \; \mathrm{R}$, since U_n and G^n have matching rst and second moments, we get 3rd derivatives, hence $kO^{(3)}k_1$
- Since O is smooth, all derivatives are \small" so $kO^{(3)}k_1$ is also \small"

Main result of HKM'13 Invariance principle for -regular polytopes

$$\mathop{\mathrm{E}}_{x \quad U_n}[Wx \quad \tilde{}\,] \quad \mathop{\mathrm{E}}_{g \quad G^n}[Wg \quad \tilde{}\,] \qquad \mathrm{polylog}\, k \tag{2}$$

How to prove this?

1. Smooth invariance. Establish (2) for smooth molli ers $O : \mathrm{R}^n \;\; ! \;\; \mathrm{R}$, i.e.,

$$\mathop{\mathrm{E}}_{x \quad U_n}[O(Wx)] \quad \mathop{\mathrm{E}}_{g \quad G^n}[O(Wg)] \qquad \log k \;\; kO^{(3)}k_1 \tag{3}$$

2. Bentkus mollifier. Care about $[Wx \qquad]$ **not** $O(Wx)$. Bentkus'90 established a molli er $B : \mathrm{R}^k \;\; ! \;\; \mathrm{R}$ that approximates the orthant function, i.e.,

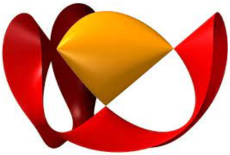
$$B(z_1; : : : ; z_k) \qquad \mathrm{h} \; \max_i z_i \qquad \mathrm{i} \qquad \text{and } kB^{`}k_1 \quad \log^{`=2} k$$

3. Anti-concentration. From above B \approximately agrees" with O.

- Around the \boundary" of the polytope is where B and O disagree
- If probability of $x \;\, 2 \;\, G^n$ lying in boundary is \small", maybe it is ok? YES
- Gaussian surface area! Nazarov'03 showed GSA of k-facet polytopes is $^{p}\overline{\log k}$

Putting everything together. All dependence are logarithmic factors, so (3) $=)$ (2).

A halfspace $F(x) = \sum_i w_i x_i$. Spectrahedron is $F(x) = x_1 A^1 + \quad + x_n A^n \quad B$



1. Hybrid method?

1. Spectrahedron naturally deals with eigenvalues of matrices
2. Unknown if Lindeberg-type argument works for spectral molli ers (i.e., smooth functions acting on the eigenspectrum of matrices)

An invariance principle for the Bentkus molli er of *arbitrary* regular spectrahedra

2. Anti-concentration? Even if GSA of spectrahedra are small, they are \funky" geometric objects, not clear how to go from molli er-closeness to CDF closeness

Prove a Littlewood-O ord theorem for positive regular spectrahedra

1. Spectral function $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ acts on the eigenvalues of matrices:

$$f(M) = g \circ \lambda(M) = g\big(\lambda_1(M), \ldots, \lambda_n(M)\big)$$

for some $g : \mathbb{R}^n \to \mathbb{R}$. Examples include determinants, trace, matrix norms

2. Derivatives of matrix-valued functions $f : \mathbb{R}^{n \times n} \to \mathbb{R}^{m \times m}$.

Taylor series. Let $h : \mathbb{R} \to \mathbb{R}$, then Taylor series of h is

$$h(x) = h(a) + \frac{h'(a)}{1!}(x - a) + \frac{h''(a)}{2!}(x - a)^2 + \frac{h'''(a)}{3!}(x - a)^3 + \cdots ;$$

where

$$h'(a) = \lim_{s \to 0} \frac{1}{s}\big(h(a + s) - h(a)\big)$$

Fréchet derivatives. Derivatives in Banach spaces. "Similar" to standard calculus. For $A, B \in \mathbb{R}^{n \times n}$, we have

$$Df(A)[B] = \lim_{s \to 0} \frac{1}{s}\big(f(A + sB) - f(A)\big);$$

$$D^t f(A)[B] = \lim_{s \to 0} \frac{1}{s}\big(D^{t-1} f(A + sB)[B] - D^{t-1} f(A)[B]\big)$$

Frechet derivatives are hard to compute. Poorly understood: basic properties as continuity, Lipschitz continuity, differentiability proven in last 2 decades.

Goal: Invariance principle for Bentkus molli er B

$$\Pr_{x \quad U_n}[(B \quad)(\sum_i x_i A^i \quad B)] \quad \Pr_{g \quad G^n}[(B \quad)(\sum_i g_i A^i \quad B)] \qquad \mathrm{polylog}\, k$$

1. **Hybrid method.** Hash the sum over $[n]$ into t blocks: let $Q_x = \sum_{i=1}^{n=t} x_i A^i$,

$$\mathop{\mathrm{E}}_{x;g}[(B \quad)\ Q_x + P_{x;g}\ ] \quad \mathop{\mathrm{E}}_{x;g}[(B \quad)(Q_g + P_{x;g})] \qquad (4)$$

2. **Taylor expansion.** Write out Frechet series for both these terms.

$$\mathcal{B}_\lambda\ \ Q_x + P_{x,g} \ = \mathcal{B}_\lambda(P_{x,g}) + D\mathcal{B}_\lambda(P_{x,g})[Q_x] + \frac{1}{2}D^2\mathcal{B}_\lambda(P_{x,g})[Q_x, Q_x] + \frac{1}{6}D^3\mathcal{B}_\lambda(P'_{x,g})[Q_x, Q_x, Q_x]$$

$$\mathcal{B}_\lambda\ \ Q_g + P_{x,g} \ = \mathcal{B}_\lambda(P_{x,g}) + D\mathcal{B}_\lambda(P_{x,g})[Q_g] + \frac{1}{2}D^2\mathcal{B}_\lambda(P_{x,g})[Q_g, Q_g] + \frac{1}{6}D^3\mathcal{B}_\lambda(R'_{x,g})[Q_g, Q_g, Q_g]$$

Same colour terms are equal in expectation

So bounding Eq (4) amounts to proving. **Goal**: upper bound

$$\mathop{\mathrm{E}}_{x;g}\Big[\ D^3B\ \ (P^0_{x;g})[Q_x; Q_x; Q_x] \quad D^3B\ \ (R^0_{x;g})[Q_g; Q_g; Q_g]\Big] \qquad \mathrm{polylog}\, k$$

3. **Sendov to the rescue.** For us, Sendov provided a tensorial representation of Frechet series for spectral functions

Recall: Goal is to upper bound

$$\mathop{\mathbb{E}}_{x;g}\left[D^3 B \; (P^0_{x;g})[Q_x; Q_x; Q_x] \quad D^3 B \; (R^0_{x;g})[Q_g; Q_g; Q_g] \right]$$

Hope: use Sendov's tensor-result. BUT, if you write it out, we get:

$H = VQV^T$. Then $\boxed{D^3F\,(P)\,[Q,Q,Q]}$ is the summation of the following terms.

1. $\sum_{i_1} \nabla^3_{i_1,i_1,i_1} f\,(x)\, H^3_{i_1,i_1}$
2. $\sum_{i_1 \neq i_2} \nabla^3_{i_1,i_2,i_1} f\,(x)\, H^2_{i_1,i_1} H_{i_2,i_2}$
3. $\sum_{i_1 \neq i_2 \neq i_3} (\nabla^3_{i_1,i_2,i_2} f\,(x)) \cdot H_{i_1,i_1} H_{i_2,i_2} H_{i_3,i_3}$

DIAGONAL ELEMENTS

4. $\sum_{i_1 \neq i_2} \left(\frac{\nabla^2_{i_2,i_2} - \nabla^2_{i_1,i_2}}{x_{i_2} - x_{i_1}} - \frac{\nabla_{i_2} - \nabla_{i_1}}{(x_{i_2} - x_{i_1})^2} \right) f\,(x)\, H_{i_2,i_2} H^2_{i_2,i_1}$
5. $\sum_{i_1 \neq i_2 \neq i_3} \frac{\nabla^2_{i_2,i_3} - \nabla^2_{i_1,i_3}}{x_{i_2} - x_{i_1}} f\,(x)\, H^2_{i_1,i_2} H_{i_3,i_3}$
6. $\sum_{i_1 \neq i_2 \neq i_3} \left(\frac{\nabla_{i_3} - \nabla_{i_1}}{(x_{i_3} - x_{i_2})(x_{i_3} - x_{i_1})} - \frac{\nabla_{i_2} - \nabla_{i_1}}{(x_{i_3} - x_{i_2})(x_{i_2} - x_{i_1})} \right) f\,(x)\, H_{i_1,i_2} H_{i_2,i_3} H_{i_3,i_1}$
7. $\sum_{i_1 \neq i_2 \neq i_3} \left(\frac{\nabla_{i_2} - \nabla_{i_3}}{(x_{i_3} - x_{i_1})(x_{i_2} - x_{i_3})} - \frac{\nabla_{i_2} - \nabla_{i_1}}{(x_{i_3} - x_{i_1})(x_{i_2} - x_{i_1})} \right) f\,(x)\, H_{i_1,i_3} H_{i_2,i_1} H_{i_3,i_2},$

Recall: Goal is to upper bound

$$\mathop{\mathbb{E}}_{x,g}\left[D^3 B\,(P'_{x,g})[Q_x, Q_x, Q_x] \quad D^3 B\,(R'_{x,g})[Q_g, Q_g, Q_g] \right]$$

Technical contribution.

- Bound each of the 7 terms by $\operatorname{polylog} k$ times norms of $Q_g, Q_x, P'_{x,g}, R'_{x,g}$.
- Completely open up the Bentkus mollifier (prior works used it as a blackbox)

4. Final step. Understand $\mathbb{E}_x[\|Q_x\|_4^4]$ and similar quantities.

- We use matrix Rosenthal's inequality (proved "recently") gives good concentration for Schatten norms of $\|\sum_i x_i A^i\|_p^p$
- Also matrix Rosenthal is true when $(x_1, \ldots, x_n)$ is p-wise independent

Putting everything together.

$$\Pr_{x \quad U_n}\left[(B \quad)\left(\sum_i x_i A^i \quad B\right)\right] \quad \Pr_{g \quad G^n}\left[(B \quad)\left(\sum_i g_i A^i \quad B\right)\right] \qquad \operatorname{polylog} k$$

$$\boxed{\text{Recall: positive spectrahedron } S = \{x : x_1A^1 + \quad + x_nA^n \quad B\} \text{ where } A^i, B \quad 0}$$

So far.

$$\Pr_{x \quad U_n}\Big(\underbrace{B \quad)(\sum_i x_iA^i \quad B}_{\sum_i x_iA^i \quad B}\Big) \qquad \Pr_{g \quad G^n}\Big(\underbrace{B \quad)(\sum_i g_iA^i \quad B}_{\sum_i g_iA^i \quad B}\Big) \qquad \text{polylog}\,k \tag{5}$$

But we care about CDF distance

$$\Pr_{x \quad U_n}\Big[\sum_i x_iA^i \quad B\Big] \qquad \Pr_{g \quad G^n}\Big[\sum_i g_iA^i \quad B\Big] \qquad \text{polylog}\,k \tag{6}$$

Intuition for this approximation: What means?

If $\quad_1, \ldots, \quad_k \in [\quad 1/100, 1/100]$, then $(B \quad)(\sum_i x_iA^i \quad B) \neq \sum_i x_iA^i \quad B$

Else $\quad_1, \ldots, \quad_k \notin [\quad 1/100, 1/100]$, then $(B \quad)(\sum_i x_iA^i \quad B) \quad \sum_i x_iA^i \quad B$

$$\boxed{\text{For uniform } x, \ \ _{\max}\Big(\sum_i x_iA^i \quad B\Big) \in \Big[\tfrac{1}{100}, \tfrac{1}{100}\Big] \text{ with tiny probability}}$$

Recall: positive spectrahedron $S = \{x : x_1A^1 + \cdots + x_nA^n \ B\}$ where $A^i; B \ 0$

So far.

$$\Pr_{x \ U_n}\underbrace{(B \ )(\sum_i x_iA^i \ B)}_{\sum_i x_iA^i \ B} \quad \Pr_{g \ G^n}\underbrace{(B \ )(\sum_i g_iA^i \ B)}_{\sum_i g_iA^i \ B} \quad \mathrm{polylog}\, k \qquad (7)$$

But we care about CDF distance

$$\Pr_{x \ U_n}\Big[\sum_i x_iA^i \ B\Big] \quad \Pr_{g \ G^n}\Big[\sum_i g_iA^i \ B\Big] \quad \mathrm{polylog}\, k \qquad (8)$$

Our result: Littlewood-Offord for spectrahedra

Let $A^1; \ldots; A^n$ be positive matrices s.t. $\sum_i \|A^i\|^2 \ 1$. For every

$$\Pr_{x \ U_n}\Big[\max \ \sum_i x_iA^i \ B \ 2 [\ ; \ ]\Big] \ O(\ ):$$

Prior: Littlewood-Offord'43, Erdős'45 proved it for halfspaces, OST'19 for polytopes

Hence (7) implies (8) except tiny probability. Done!

Recall: positive spectrahedron $F(x) = x_1 A^1 + \quad + x_n A^n \quad B$ where $A^i; B \quad 0$

A PRG that -fools the class of positive width-M spectrahedra with seed length $\mathrm{poly}(\log n; \log k; M; 1= \ )$

Open questions:

1. Remove regularity?
2. Remove positivity?
3. What is the Gaussian surface area of spectrahedron?
4. Improve the $1=$ dependence?
5. A general invariance principle for spectral functions?

THANK YOU