Sum-of-Squares Lower Bounds for Non-Gaussian Component Analysis

Challenges and New Techniques

Shuo Pang

University of Copenhagen

TTIC, June 14, 2024

With Ilias Diakonikolas, Sushrut Karmalkar, and Aaron Potechin

Outline

1 Problem Formulation

2 Pseudo-Calibration

3 PSDness via Representation

4 Error Analysis

Non-Gaussian Component Analysis

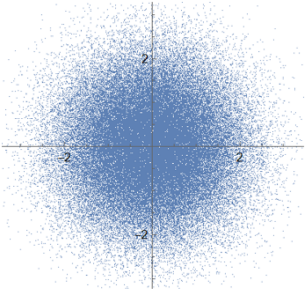
How many samples are needed to distinguish $N(0, \mathrm{Id}_n)$ from a planted distribution D?

- $D = A \times N(0, \mathrm{Id}_{n-1})_{v^\perp}$, v unknown
- A matches first $k - 1$ moments with $N(0, 1)$.

Non-Gaussian Component Analysis

How many samples are needed to distinguish $N(0, \mathrm{Id}_n)$ from a planted distribution D?

- $D = A \times N(0, \mathrm{Id}_{n-1})_{v^\perp}$, v unknown
- A matches first $k-1$ moments with $N(0,1)$.



vs

Non-Gaussian Component Analysis

How many samples are needed to distinguish $N(0, \mathrm{Id}_n)$ from a planted distribution D?

- $D = A \times N(0, \mathrm{Id}_{n-1})_{v^\perp}$, v unknown
- A matches first $k-1$ moments with $N(0,1)$.

Under mild conditions, information-theoretically $O(n)$.

- **Statistical Query:** $\geq n^{\Omega(k)}$ [Diakonikolas-Kane-Stewart 17]
- **Spectral (k-tensor):** $\leq n^{k/2}$ [Dudeja-Hsu 20]

Non-Gaussian Component Analysis

How many samples are needed to distinguish $N(0, \mathrm{Id}_n)$ from a planted distribution D?

- $D = A \times N(0, \mathrm{Id}_{n-1})_{v^\perp}$, v unknown
- A matches first $k-1$ moments with $N(0,1)$.

Under mild conditions, information-theoretically $O(n)$.

- **Statistical Query:** $\geq n^{\Omega(k)}$ [Diakonikolas-Kane-Stewart 17]
- **Spectral (k-tensor):** $\leq n^{k/2}$ [Dudeja-Hsu 20]
- **Sum-of-Squares?**

Problem Formulation

Given m i.i.d. samples $\sim N(0, \mathrm{Id}_n)$, can SoS efficiently rule out the existence of v?

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

- **Variables** $\{\mathrm{v}_1^{i_1} \cdots \mathrm{v}_n^{i_n} : \ i_1 + \cdots + i_n \leq d\}$ (formal names, no relation)

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

- **Variables** $\{\mathrm{v}_1^{i_1} \cdots \mathrm{v}_n^{i_n} : \ i_1 + \cdots + i_n \leq d\}$
  (formal names, no relation)
- **Constraints**
  1. Match statistics of A
  2. Positivity
  3. Booleaness (optional)

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

- **Variables** $\{\mathrm{v}_1^{i_1} \cdots \mathrm{v}_n^{i_n} : \; i_1 + \cdots + i_n \leq d\}$
  (formal names, no relation)
- **Constraints**
  1. Match statistics of A
     $$\left| \frac{1}{m} \sum_{u=1}^{m} He_i(\langle x_u, \mathrm{v} \rangle) - \mathop{\mathbb{E}}_{A}[He_i] \right| \leq O_A\left(\frac{1}{\sqrt{m}}\right), \quad \forall i \leq d$$
  2. Positivity
  3. Booleaness (optional)

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

- **Variables** $\{\mathrm{v}_1^{i_1} \cdots \mathrm{v}_n^{i_n} : \; i_1 + \cdots + i_n \leq d\}$
  (formal names, no relation)
- **Constraints**
  1. Match statistics of A
     $$\left| \frac{1}{m} \sum_{u=1}^{m} He_i(\langle x_u, \mathrm{v} \rangle) - \mathop{\mathbb{E}}_{A}[He_i] \right| \leq O_A(\frac{1}{\sqrt{m}}), \quad \forall i \leq d$$
  2. Positivity $p^2(\mathrm{v}) \geq 0$ for all low degree polynomial p.
  3. Booleaness (optional)

Sum-of-Squares Relaxation

Degree-d SoS

Given input $x_1, \ldots, x_m \sim N(0, \mathrm{Id}_n)$, run an SDP.

- **Variables** $\{\mathrm{v}_1^{i_1} \cdots \mathrm{v}_n^{i_n} : \; i_1 + \cdots + i_n \leq d\}$ (formal names, no relation)
- **Constraints**
  1. Match statistics of A
  $$\left| \frac{1}{m} \sum_{u=1}^{m} He_i(\langle x_u, \mathrm{v} \rangle) - \mathop{\mathbb{E}}_{A}[He_i] \right| \leq O_A(\frac{1}{\sqrt{m}}), \quad \forall i \leq d$$
  2. Positivity $p^2(\mathrm{v}) \geq 0$ for all low degree polynomial p.
  3. Booleaness (optional) $\mathrm{v}^I \mathrm{v}_i^2 = \frac{1}{n} \mathrm{v}^I$

Sum-of-Squares Relaxation

- To show lower bounds, given $x_1, \ldots, x_m$, we find a feasible solution
$$\widetilde{E} : \{\mathrm{v}^I\} \to \mathbb{R}.$$
- We need to consider arbitrary $x_1, \ldots, x_m$.

Sum-of-Squares Relaxation

- To show lower bounds, given $x_1, \ldots, x_m$, we find a feasible solution
$$\widetilde{E} : \{\mathrm{v}^I\} \to \mathbb{R}.$$
- We need to consider arbitrary $x_1, \ldots, x_m$. Pseudo-calibration

Outline

1 Problem Formulation

2 Pseudo-Calibration

3 PSDness via Representation

4 Error Analysis

Designing $\widetilde{E}$

Idea:

1. Let $\widetilde{E}(\mathrm{v}^I)$ be **low-degree polynomial** in $\vec{x}$.

Designing $\widetilde{E}$

Idea:

1. Let $\widetilde{E}(\mathrm{v}^I)$ be **low-degree polynomial** in $\vec{x}$.

$$\widetilde{E}(\mathrm{v}^I) = \sum_{a \in \mathbb{N}^{mn}:\ \text{low}} c_{I,a} \cdot He_a(\vec{x})$$

$$He_a(\vec{x}) = \prod_{u,i} He_{a_{u,i}}(x_{u,i}) \text{ Hermite polynomials.}$$

Designing $\widetilde{E}$

Idea:

1. Let $\widetilde{E}(\mathrm{v}^I)$ be **low-degree polynomial** in $\vec{x}$.

$$\widetilde{E}(\mathrm{v}^I) = \sum_{a \in \mathbb{N}^{mn}:\ \text{low}} c_{I,a} \cdot He_a(\vec{x})$$

$$He_a(\vec{x}) = \prod_{u,i} He_{a_{u,i}}(x_{u,i}) \text{ Hermite polynomials.}$$

$$\begin{aligned}
He_0(x) &= 1, \\
He_1(x) &= x, \\
He_2(x) &= x^2 - 1, \\
He_3(x) &= x^3 - 3x, \\
He_4(x) &= x^4 - 6x^2 + 3, \\
He_5(x) &= x^5 - 10x^3 + 15x, \\
He_6(x) &= x^6 - 15x^4 + 45x^2 - 15, \\
He_7(x) &= x^7 - 21x^5 + 105x^3 - 105x, \\
He_8(x) &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105, \\
He_9(x) &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x, \\
He_{10}(x) &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945.
\end{aligned}$$

Designing $\widetilde{E}$

Idea:

1. Let $\widetilde{E}(\mathrm{v}^I)$ be **low-degree polynomial** in $\vec{x}$.

$$\widetilde{E}(\mathrm{v}^I) = \sum_{a \in \mathbb{N}^{mn}:\ \text{low}} c_{I,a} \cdot He_a(\vec{x})$$

2. Choose coefficients by **pseudo-calibration**. [Barak-Hopkins-Kelner-Kothari-Moitra-Potechin 16]

Designing $\widetilde{E}$

Idea:

1. Let $\widetilde{E}(\mathrm{v}^I)$ be **low-degree polynomial** in $\vec{x}$.

$$\widetilde{E}(\mathrm{v}^I) = \sum_{a \in \mathbb{N}^{mn}:\ \text{low}} c_{I,a} \cdot He_a(\vec{x})$$

2. Choose coefficients by **pseudo-calibration**. [Barak-Hopkins-Kelner-Kothari-Moitra-Potechin 16]

$$\begin{aligned} c_{I,a} &:= \text{average correlation over planted cases} \\ &= \mathop{\mathbb{E}}_{\substack{v \sim \{\frac{\pm 1}{\sqrt{n}}\}^n \\ x \sim \mathcal{D}_{v,A}}} \langle v^I, He_a(x) \rangle \end{aligned}$$

Our result

Under mild conditions on A,

Theorem (SoS Lower Bounds for NGCA)

W.p. $1 - o_n(1)$ *over* $m = n^{(1-\epsilon)k/2}$ *many samples from* $N(0, \mathrm{Id}_n)$*, degree* $\sqrt{\log n}$ *pseudo-calibration is a feasible solution.*

Our result

Under mild conditions on A,

Theorem (SoS Lower Bounds for NGCA)

W.p. $1 - o_n(1)$ *over* $m = n^{(1-\epsilon)k/2}$ *many samples from* $N(0, \mathrm{Id}_n)$*, degree* $\sqrt{\log n}$ *pseudo-calibration is a feasible solution.*

Conditions on A

A matches $k-1$ moments and $\exists C > 0$ s.t.

1. **Moment Bounds**
2. **Non-singular**

Our result

Under mild conditions on A,

Theorem (SoS Lower Bounds for NGCA)

W.p. $1 - o_n(1)$ over $m = n^{(1-\epsilon)k/2}$ many samples from $N(0, \mathrm{Id}_n)$, degree $\sqrt{\log n}$ pseudo-calibration is a feasible solution.

Conditions on A

A matches $k-1$ moments and $\exists C > 0$ s.t.

1. **Moment Bounds** $|\mathbb{E}_A[He_i]| \leq (\log n)^{C \cdot i}$ for all $i \leq \sqrt{\log n}$.
2. **Non-singular**

Our result

Under mild conditions on A,

Theorem (SoS Lower Bounds for NGCA)

W.p. $1 - o_n(1)$ *over* $m = n^{(1-\epsilon)k/2}$ *many samples from* $N(0, \mathrm{Id}_n)$, *degree* $\sqrt{\log n}$ *pseudo-calibration is a feasible solution.*

Conditions on A

A matches $k-1$ moments and $\exists C > 0$ s.t.

1. **Moment Bounds** $\left|\mathbb{E}_A[He_i]\right| \leq (\log n)^{C \cdot i}$ for all $i \leq \sqrt{\log n}$.
2. **Non-singular** $\mathbb{E}_A[q^2] \geq (\log n)^{-C\sqrt{\log n}}$, $\forall q$: deg $\leq \sqrt{\log n}$, ℓ_2-unit in $N(0,1)$.

Our Result

In other words, degree $\sqrt{\log n}$ SoS algorithms require $n^{(1-\epsilon)k/2}$ samples to solve NGCA.

Our Result

In other words, degree $\sqrt{\log n}$ SoS algorithms require $n^{(1-\epsilon)k/2}$ samples to solve NGCA.

- Almost tight, matching $n^{k/2}$ [Dudeja-Hsu 22]
- Super-constant degree
- Applications:

  Robust mean estimation
  List-decodable mean estimation
  Robust covariance estimation (additive, multiplicative)
  Learning k-mixed Gaussians ($k \geq 2$)
  Noisy planted planes [GJJPR 21]
  ...

Rest of talk:

- $m = n^{(1-\epsilon)k}$, and $\mathop{\mathbb{E}}_{A}[He_i] = 0, \forall i \in [1, k-1]$.
- positivity constraints

Our Goal

- Positivity $\Leftrightarrow$ moment matrix M is PSD

$$M(I, J) := \widetilde{E}(\mathrm{v}^{I+J}), \quad I, J \in \binom{[n]}{\leq d_{SoS}}$$

Our Goal

- Positivity $\Leftrightarrow$ moment matrix M is PSD

$$M(I,J) := \widetilde{E}(\mathrm{v}^{I+J}), \quad I,J \in \binom{[n]}{\leq d_{SoS}}$$

where

$$\widetilde{E}(\mathrm{v}^I) = \sum_{\substack{a\in(\mathbb{N}^n)^m:\ \text{low},\\ \text{some more conditions}}} n^{-\frac{\|I\|_1+\|a\|_1}{2}} \frac{1}{a!} \left(\prod_{u=1}^{m} \mathop{\mathbb{E}}_{A} \left[He_{\|a_u\|_1} \right] \right) He_a,$$

$$\text{where } a! := \prod_{u,i} a_{u,i}!.$$

Our Goal

- Positivity $\Leftrightarrow$ moment matrix M is PSD

$$M(I,J) := \widetilde{E}(\mathrm{v}^{I+J}),\quad I,J \in \binom{[n]}{\leq d_{SoS}}$$

  - Entries of M are low-deg in $x_{u,i}$ (u for sample, i for coordinate)
  - Invariant under $\mathcal{S}_m \times \mathcal{S}_n$
  - Tool: Graph matrices

Graph Matrices [Medarametla-Potechin 16, Ahn-M-P 20]

Graph matrices $\{M_\alpha\}$

- A basis of such matrix functions. (low-deg, "graph-theoretic")

Graph Matrices [Medarametla-Potechin 16, Ahn-M-P 20]

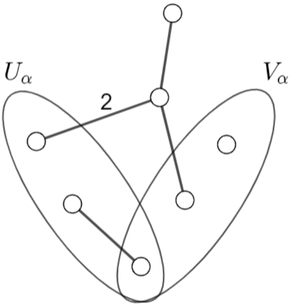
Graph matrices $\{M_\alpha\}$

- A basis of such matrix functions.

Definition (Shape)

A **shape** $\alpha = (V(\alpha), E(\alpha))$ *is a edge-weighted graph, plus two "sides"* U_α, V_α.

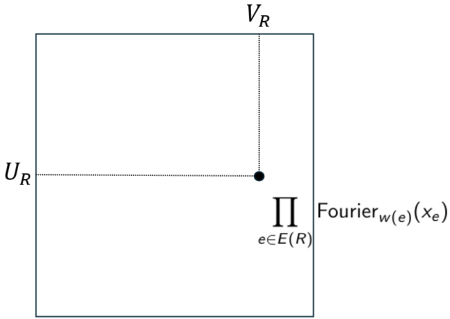
A shape α



Graph Matrices [Medarametla-Potechin 16, Ahn-M-P 20]

Graph matrices $\{M_\alpha\}$

- A basis of such matrix functions.
- Shapes can be realized on $[N]$. Realization R gives matrix M_R.
- $M_\alpha = \sum_{\text{realization of } \alpha} M_R$.

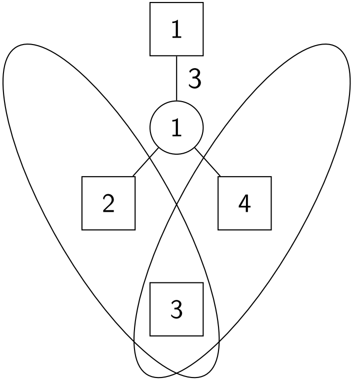


For NGCA [GJJPR 21]

- $He_t(x_{u,i})$: bipartite shapes, edge $\{ⓤ, \boxed{i}\}$ of weight t.
  $(u \in [m], i \in [n])$

For NGCA [GJJPR 21]

- $He_t(x_{u,i})$: bipartite shapes, edge $\{ⓤ, \boxed{i}\}$ of weight t. $(u \in [m], i \in [n])$



$$M_R\left(\{\boxed{2}, \boxed{3}\}, \{\boxed{3}, \boxed{4}\}\right) = He_3(x_{1,1}) \cdot He_1(x_{1,2}) \cdot He_1(x_{1,4})$$

Norm Bounds

Theorem [Ahn-Medarametla-Potechin 20]

W.h.p. over $\vec{x}$, simultaneously for all small shapes α:

$$\|M_\alpha\| \lesssim n^{\frac{w(V)-w(S_{min})}{2}+o(1)}.$$

- $w(\square) = 1$, $w(\bigcirc) = \log_n m$.
- S_{min}: minimum weight vertex separator.

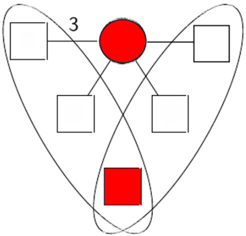
Norm Bounds

Theorem [Ahn-Medarametla-Potechin 20]

W.h.p. over $\vec{x}$, simultaneously for all small shapes α:

$$\|M_\alpha\| \lesssim n^{\frac{w(V)-w(S_{min})}{2}+o(1)}.$$

- $w(\square) = 1$, $w(\bigcirc) = \log_n m$.
- S_{min}: minimum weight vertex separator.



Norm Bounds

Theorem [Ahn-Medarametla-Potechin 20]

W.h.p. over $\vec{x}$, simultaneously for all small shapes α:

$$\|M_\alpha\| \lesssim n^{\frac{w(V)-w(S_{min})}{2}+o(1)}.$$

- $w(\square) = 1$, $w(\bigcirc) = \log_n m$.
- S_{min}: minimum weight vertex separator.

Takeaway: $\|M_\alpha\|$ is determined by $w(V) - w(S_{min})$.

Analyzing Moment Matrix

$$M = \sum_{\substack{\alpha:\ \text{small,} \\ \text{other conditions}}} \underbrace{(\text{Hermite coefficient})}_{\frac{\prod_{v:\ \text{circle}} \mathbb{E}_A \left[He_{\deg_\alpha(v)}\right]}{\prod_{e \in E(\alpha)} w(e)!}} \cdot \underbrace{(\text{Scaling coefficient})}_{n^{-w(E(\alpha))/2}} \cdot M_\alpha$$

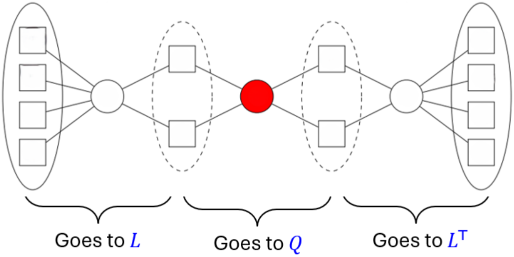
Analyzing Moment Matrix

$$M = \sum_{\substack{\alpha:\ \text{small},\\ \text{other conditions}}} \underbrace{(\text{Hermite coefficient})}_{\frac{\prod_{v:\ \text{circle}} \mathbb{E}_A\left[He_{\deg_\alpha(v)}\right]}{\prod_{e\in E(\alpha)} w(e)!}} \cdot \underbrace{(\text{Scaling coefficient})}_{n^{-w(E(\alpha))/2}} \cdot M_\alpha$$

First step. Factorize $M \approx LQL^\top$.
[BHKKMP16, PR20, JPRTX21, P21, JPRX23]

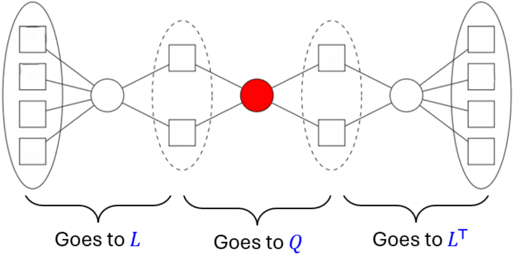
Factorize $M \approx LQL^\top$

- Intuition: shape composition $\leftrightsquigarrow$ matrix product



Factorize $M \approx LQL^\top$

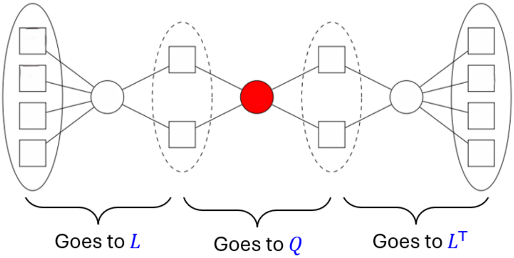
- Intuition: shape composition $\leftrightsquigarrow$ matrix product



- Key: use vertex separator to decompose shapes

Factorize $M \approx LQL^\top$

- Intuition: shape composition $\leftrightsquigarrow$ matrix product



- Key: use vertex separator to decompose shapes
- We use minimum square-vertex separators.

Factorization with Minimum Square Separators

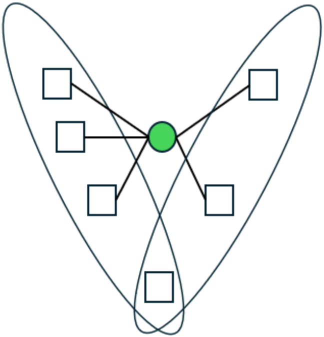
Factorization Lemma

We have $M \approx LQL^\top$, where L is okayish-conditioned, and

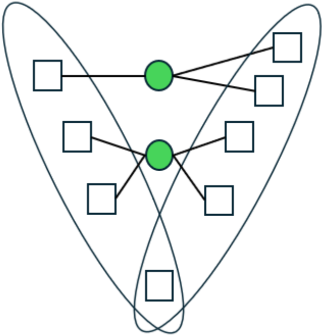
$$Q = Q_{\mathrm{main}} + n^{-\epsilon}, \; Q_{\mathrm{main}} \text{ is sum of special shapes.}$$

Special shapes

Simple spider disjoint unions



Simple spider $S(3,2;1)$



A simple spider disjoint union

Outline

1 Problem Formulation

2 Pseudo-Calibration

3 PSDness via Representation

4 Error Analysis

Many Dominant Terms

Goal: prove Q_{main} is positive-definite
In all previous works, it's a constant matrix.

Many Dominant Terms

Goal: prove Q_{main} is positive-definite
In all previous works, it's a constant matrix.

Challenge

Q_{main} contains a whole family of non-constant, equally dominant shapes.

- Simple spider disjoint unions
- With recursive coefficients (involving products of $\mathbb{E}_A[He_i]$)

Many Dominant Terms

Goal: prove Q_{main} is positive-definite
In all previous works, it's a constant matrix.

Challenge

Q_{main} contains a whole family of non-constant, equally dominant shapes.

- Simple spider disjoint unions
- With recursive coefficients (involving products of $\mathbb{E}_A[He_i]$)

Idea: study **multiplicative structure** of them

Algebra of Simple Spiders

$$S_\alpha := (\text{scaled } M_\alpha) = n^{\frac{-w(E(\alpha))}{2}} M_\alpha.$$

Simple Spider Algebra (SA)

Basis: simple spiders with side size $\leq d$.

Multiplication $\star$: includes only simple spiders in $S_\alpha \cdot S_\beta$, with idealized coefficients.

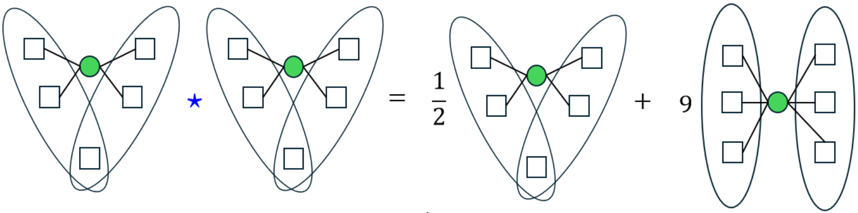
Algebra of Simple Spiders

$$S_\alpha := (\text{scaled } M_\alpha) = n^{\frac{-w(E(\alpha))}{2}} M_\alpha.$$

Simple Spider Algebra (SA)

Basis: simple spiders with side size $\leq d$.

Multiplication $\star$: includes only simple spiders in $S_\alpha \cdot S_\beta$, with idealized coefficients.



$$S(2,2;1) \star S(2,2;1) = \frac{1}{2} \cdot S(2,2;1) + 9 \cdot S(3,3;0)$$

Algebra of Simple Spiders

$$S_\alpha := (\text{scaled } M_\alpha) = n^{\frac{-w(E(\alpha))}{2}} M_\alpha.$$

Simple Spider Algebra (SA)

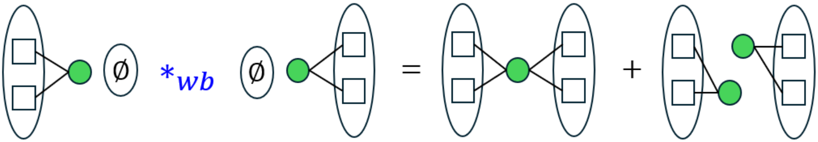
Basis: simple spiders with side size $\leq d$.

Multiplication $\star$: includes only simple spiders in $S_\alpha \cdot S_\beta$, with idealized coefficients.

Disjoint Union Algebra $(\mathrm{SA}_{\mathrm{disj}})$

On simple spider disjoint unions. $*_{\mathrm{wb}}$: well-behaved product.

Algebra of Simple Spiders



Algebra of Simple Spiders

Basic Properties

- (Associativity) Both are associative $\mathbb{R}$-algebras.
- (Compatibility) If restrict $*_{\mathrm{wb}}$ to simple spiders, we get $\star$.
- (Approximation)

$$\|S_\alpha \cdot S_\beta - S_\alpha *_{\mathrm{wb}} S_\beta\| \leq n^{-\epsilon}$$

assuming all circles have $\geq k$ legs to each side.

Algebra of Simple Spiders

Basic Properties

- (Associativity) Both are associative $\mathbb{R}$-algebras.
- (Compatibility) If restrict $*_{\mathrm{wb}}$ to simple spiders, we get $\star$.
- (Approximation)

$$\|S_\alpha \cdot S_\beta - S_\alpha *_{\mathrm{wb}} S_\beta\| \leq n^{-\epsilon}$$

assuming all circles have $\geq k$ legs to each side.

Using these algebras, we can nail down Q_{main}.

Determining Q_{main}

Lemma

Q_{main} is uniquely determined by

$$L *_{\text{wb}} Q_{\text{main}} *_{\text{wb}} L^{\top} = P. \quad (L, P \text{ explicit}) \tag{1}$$

Moreover,

$$L \star Q_{\text{main}} \star L^{\top} = P_{\mathsf{SS}}. \tag{2}$$

L and P look like this:

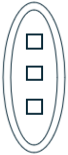
$$\sum_{\substack{\alpha: \\ \text{simple spider disjoint union} \\ \text{with certain conditions}}} \left(\prod_{v:\ \text{circle vertex}} \mathbb{E}_A[He_{\deg(v)}] \right) \cdot S_\alpha$$

Determining Q_{main}

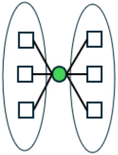
L and P look like this:

$$\sum_{\substack{\alpha:\\ \text{simple spider disjoint union}\\ \text{with certain conditions}}} \left(\prod_{v:\ \text{circle vertex}} \mathop{\mathbb{E}}_{A}[He_{\deg(v)}] \right) \cdot S_\alpha$$

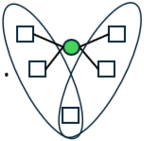
Some terms:



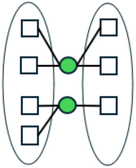
$+ E_A[He_6] \cdot$



$+ E_A[He_4] \cdot$



$+ E_A[He_4] \cdot E_A[He_3] \cdot$



$+ \cdots$

Determining Q_{main}

Lemma

Q_{main} is uniquely determined by

$$L *_{\text{wb}} Q_{\text{main}} *_{\text{wb}} L^\top = P. \quad (L, P \text{ explicit}) \tag{1}$$

Moreover,

$$L \star Q_{\text{main}} \star L^\top = P_{SS}. \tag{2}$$

The proof relies on intricate **error analysis**.

Proving PSDness

$$L *_{\mathrm{wb}} Q_{\mathrm{main}} *_{\mathrm{wb}} L^{\top} = P \quad \Rightarrow \quad Q_{\mathrm{main}} \succ 0$$

Proving PSDness

$$L *_{\mathrm{wb}} Q_{\mathrm{main}} *_{\mathrm{wb}} L^{\top} = P \quad \Rightarrow \quad Q_{\mathrm{main}} \succ 0$$

Proof Overview

PSDness in simple spider world $\rightarrow$ disjoint unions $\rightarrow$ real world

Proving PSDness

$$L *_{\mathrm{wb}} Q_{\mathrm{main}} *_{\mathrm{wb}} L^\top = P \quad \Rightarrow \quad Q_{\mathrm{main}} \succ 0$$

Proof Overview

PSDness in simple spider world → disjoint unions → real world

1. Show that $P_{SS} = a \star a^\top$.
2. Show that $Q_{\mathrm{main}} = b *_{\mathrm{wb}} b^\top$.
3. $Q_{\mathrm{main}} \approx b \cdot b^\top \succ 0$.

Proving PSDness

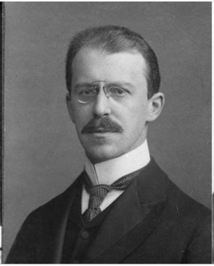
$$L *_{\mathrm{wb}} Q_{\mathrm{main}} *_{\mathrm{wb}} L^{\top} = P \quad \Rightarrow \quad Q_{\mathrm{main}} \succ 0$$

Proof Overview

PSDness in simple spider world → disjoint unions → real world

1. Show that $P_{SS} = a \star a^{\top}$.
   - Using representation
2. Show that $Q_{\mathrm{main}} = b *_{\mathrm{wb}} b^{\top}$.
3. $Q_{\mathrm{main}} \approx b \cdot b^{\top} \succ 0$.






Ferdinand Frobenius
1849 – 1917

William Brunside
1852 – 1937

Issai Schur
1875 – 1941

Richard Brauer
1901 – 1977

Simple Spider Algebra

Basis $\{S(k_1, k_2; u) \mid k_1 + u, k_2 + u \leq d\}$.

Structural Constants

$$S(k_1 - u, k_2 - u; u) \star S(k_2 - v, k_3 - v; v) =$$

$$\sum_{i=\max\{0,u+v-k_2\}}^{\min\{u,v\}} \binom{k_1 - i}{k_1 - u} \binom{k_3 - i}{k_3 - v} / (k_2 + i - u - v)! \cdot S(k_1 - i, k_3 - i; i)$$

Simple Spider Algebra

Basis $\{S(k_1, k_2; u) \mid k_1 + u, k_2 + u \leq d\}$.

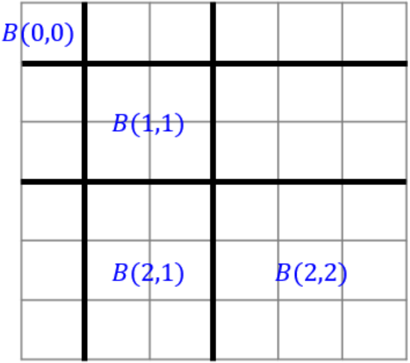
Structural Constants

$$S(k_1 - u, k_2 - u; u) \star S(k_2 - v, k_3 - v; v) =$$
$$\sum_{i=\max\{0,u+v-k_2\}}^{\min\{u,v\}} \binom{k_1 - i}{k_1 - u}\binom{k_3 - i}{k_3 - v}/(k_2 + i - u - v)! \cdot S(k_1 - i, k_3 - i; i)$$

- **Representation**: homomorphism to a matrix algebra.
- We will construct $\rho : \mathrm{SA} \to M_{1+\cdots+(d+1)}(\mathbb{R})$.

Simple Spider Algebra

- **Representation**: homomorphism to a matrix algebra.
- We will construct $\rho : \mathrm{SA} \to M_{1+\cdots+(d+1)}(\mathbb{R})$.



Block structure for $M_6(\mathbb{R})$

Representation

Definition (Representation ρ)

$\rho : SA \to M_{1+\ldots+(d+1)}(\mathbb{R})$ maps each $S(k_1 - t, k_2 - t; t)$ to a matrix supported on block (k_1, k_2), where nonzero entries appear "diagonally bottom-up":

$$\frac{\sqrt{i!j!}}{(k_1 - t)!(k_2 - t)!\big(t - (k_1 - i)\big)!} \text{ if } j - i = k_2 - k_1.$$

Representation

Definition (Representation ρ)

$\rho : SA \to M_{1+\ldots+(d+1)}(\mathbb{R})$ maps each $S(k_1 - t, k_2 - t; t)$ to a matrix supported on block (k_1, k_2), where nonzero entries appear "diagonally bottom-up":

$$\frac{\sqrt{i!j!}}{(k_1 - t)!(k_2 - t)!\big(t - (k_1 - i)\big)!} \text{ if } j - i = k_2 - k_1.$$

$$\mathrm{Im}(\rho): \left(\begin{array}{c|cc|ccc} * & 0 & * & 0 & 0 & * \\ \hline 0 & * & 0 & 0 & * & 0 \\ * & 0 & * & 0 & 0 & * \\ \hline 0 & 0 & 0 & * & 0 & 0 \\ 0 & * & 0 & 0 & * & 0 \\ * & 0 & * & 0 & 0 & * \end{array}\right)$$

Representation

Definition (Representation ρ)

$\rho : SA \to M_{1+\ldots+(d+1)}(\mathbb{R})$ *maps each* $S(k_1 - t, k_2 - t; t)$ *to a matrix supported on block* (k_1, k_2), *where nonzero entries appear "diagonally bottom-up":*

$$\frac{\sqrt{i!j!}}{(k_1 - t)!(k_2 - t)!\big(t - (k_1 - i)\big)!} \text{ if } j - i = k_2 - k_1.$$

Proposition

The linear extension of ρ is an algebra isomorphism $\mathrm{SA} \xrightarrow{\cong} \mathrm{Im}(\rho)$.

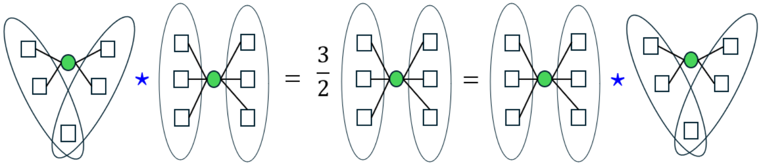
Example

Simple spiders with equal #legs to sides: commutative subalgebra

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.



$$S_2 \star S_3 = \frac{3}{2} S_3 = S_3 \star S_2$$

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.
- Elements are linear combinations $a = \sum_{i \leq s} a_i \cdot S_i$,

$$\frac{(a \star b)_i}{i!} = \sum_{(j,j',):\ 0 \leq j,j' \leq i \leq j+j'} \binom{i}{j} \binom{j}{i-j'} \frac{a_j}{j!} \cdot \frac{b_{j'}}{(j')!}$$

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.
- Elements are linear combinations $a = \sum_{i \leq s} a_i \cdot S_i$,

$$\frac{(a \star b)_i}{i!} = \sum_{(j,j',):\ 0 \leq j,j' \leq i \leq j+j'} \binom{i}{j} \binom{j}{i-j'} \frac{a_j}{j!} \cdot \frac{b_{j'}}{(j')!}$$

What does this formula look like?

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.
- Elements are linear combinations $a = \sum_{i \leq s} a_i \cdot S_i$,

$$\frac{(a \star b)_i}{i!} = \sum_{(j,j',):\ 0 \leq j,j' \leq i \leq j+j'} \binom{i}{j} \binom{j}{i-j'} \frac{a_j}{j!} \cdot \frac{b_{j'}}{(j')!}$$

- Take **symmetric function on** $\{0,1\}^s$:

$$f_a := \sum_{I \subseteq \{0,1\}^s} \frac{a_{|I|}}{|I|!} x^I, \quad \text{then } f_a \cdot f_b = f_{a \star b}$$

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.
- Elements are linear combinations $a = \sum_{i \leq s} a_i \cdot S_i$,

$$\frac{(a \star b)_i}{i!} = \sum_{(j,j',):\ 0 \leq j,j' \leq i \leq j+j'} \binom{i}{j} \binom{j}{i-j'} \frac{a_j}{j!} \cdot \frac{b_{j'}}{(j')!}$$

- Take **symmetric function on** $\{0,1\}^s$:

$$f_a := \sum_{I \subseteq \{0,1\}^s} \frac{a_{|I|}}{|I|!} x^I, \quad \text{then } f_a \cdot f_b = f_{a \star b}$$

- a is sum-of-squares in $\mathrm{SA} \Leftrightarrow f_a$ is point-wise nonnegative

Example

Simple spiders with equal #legs to sides: commutative subalgebra

- Fixing side-size s.
- Elements are linear combinations $a = \sum_{i \leq s} a_i \cdot S_i$,

$$\frac{(a \star b)_i}{i!} = \sum_{(j,j',):\ 0 \leq j,j' \leq i \leq j+j'} \binom{i}{j} \binom{j}{i-j'} \frac{a_j}{j!} \cdot \frac{b_{j'}}{(j')!}$$

- Take **symmetric function on** $\{0,1\}^s$:

$$f_a := \sum_{I \subseteq \{0,1\}^s} \frac{a_{|I|}}{|I|!} x^I, \quad \text{then } f_a \cdot f_b = f_{a \star b}$$

- a is sum-of-squares in $\mathrm{SA} \Leftrightarrow f_a$ is point-wise nonnegative
- f values are diagonals of ρ.

PSDness via Representation

Lemma (Structure of ρ)

1. $\mathrm{Im}(\rho) \cong \bigoplus_{i=0}^{d} M_{i+1}(\mathbb{R})$.

$$
\left(\begin{array}{c|cc|ccc}
\bullet & 0 & \bullet & 0 & 0 & \bullet \\
\hline
0 & \triangle & 0 & 0 & \triangle & 0 \\
\bullet & 0 & \bullet & 0 & 0 & \bullet \\
\hline
0 & 0 & 0 & \times & 0 & 0 \\
0 & \triangle & 0 & 0 & \triangle & 0 \\
\bullet & 0 & \bullet & 0 & 0 & \bullet
\end{array}\right)
$$

PSDness via Representation

Lemma (Structure of ρ)

1. $\mathrm{Im}(\rho) \cong \bigoplus_{i=0}^{d} M_{i+1}(\mathbb{R})$.
2. $\rho(X^\top) = \rho(X)^\top$ for all $X \in \mathrm{SA}$.

PSDness via Representation

Lemma (Structure of ρ)

1. $\mathrm{Im}(\rho) \cong \bigoplus_{i=0}^{d} M_{i+1}(\mathbb{R})$.
2. $\rho(X^\top) = \rho(X)^\top$ for all $X \in \mathrm{SA}$.

Item 1: ρ "contains all irreducible representations" of SA.

(Wedderburn-Artin reified)

Item 2: $X \in \mathrm{SA}$ is a sum-of-squares iff $\rho(X)$ is PSD.

PSDness via Representation

Lemma (Structure of ρ)

1. $\mathrm{Im}(\rho) \cong \bigoplus_{i=0}^{d} M_{i+1}(\mathbb{R})$.
2. $\rho(X^\top) = \rho(X)^\top$ for all $X \in \mathrm{SA}$.

Target P under $\rho(\cdot)$:

Lemma (PSDness)

$\rho(P) = \bigoplus_{i=0}^{d} P_i$, each P_i a leading principal minor of P_d, and

$$P_d = \mathop{\mathbb{E}}_{z\sim A}\left[v(z)\cdot v(z)^\top\right] \text{ where } v(z) = \left(He_0(z), \ldots, He_d(z)\right).$$

PSDness via Representation

Lemma (Structure of ρ)

1. $\mathrm{Im}(\rho) \cong \bigoplus_{i=0}^{d} M_{i+1}(\mathbb{R})$.
2. $\rho(X^\top) = \rho(X)^\top$ for all $X \in \mathrm{SA}$.

Target P under $\rho(\cdot)$:

Lemma (PSDness)

$\rho(P) = \bigoplus_{i=0}^{d} P_i$, each P_i a leading principal minor of P_d, and

$$P_d = \mathop{\mathbb{E}}_{z \sim A}[v(z) \cdot v(z)^\top] \text{ where } v(z) = (He_0(z), \ldots, He_d(z)) .$$

Proof boils down to Hermite multiplication formula.

Lift to True PSDness

1. Using ρ back-and-forth, it's not too hard to show

$$(Q_{\text{main}})_{SS} = a \star a^\top, \; a \in \mathrm{SA}.$$

Lift to True PSDness

1. Using ρ back-and-forth, it's not too hard to show

$$(Q_{\mathrm{main}})_{SS} = a \star a^\top, \;\; a \in \mathrm{SA}.$$

2. Lift to disjoint unions: *d*-combination operator $[\cdot]^d$

$$[x \star y]^d = [x]^d *_{\mathrm{wb}} ([y]^d)^\top \text{ for special } x, y.$$

From here,

$$Q_{\mathrm{main}} = [a]^d *_{\mathrm{wb}} ([a]^d)^\top \approx [a]^d \cdot ([a]^d)^\top.$$

Lift to True PSDness

1. Using ρ back-and-forth, it's not too hard to show
$$(Q_{\text{main}})_{SS} = a \star a^\top, \; a \in \mathrm{SA}.$$

2. Lift to disjoint unions: *d*-combination operator $[\cdot]^d$
$$[x \star y]^d = [x]^d *_{\text{wb}} ([y]^d)^\top \text{ for special } x, y.$$
From here,
$$Q_{\text{main}} = [a]^d *_{\text{wb}} ([a]^d)^\top \approx [a]^d \cdot ([a]^d)^\top.$$

3. $[a]^d$ is sufficiently non-singular.

Remark

$[\cdot]^d$ operator: combinatorial construction

Outline

1 Problem Formulation

2 Pseudo-Calibration

3 PSDness via Representation

4 Error Analysis

Error Analysis: Highlights

- Error analysis is needed throughout the proof.

Error Analysis: Highlights

- Error analysis is needed throughout the proof.
- Advanced charging argument in a systematic language.
  For norm estimates in sequential matrix multiplication.
- Interplay between min-square and min-weight separators.

Summary

- We prove SoS lower bounds for Non-Gaussian Component Analysis, an important problem.
- This closes the gap between statistical query/low-degree polynomial lower bounds and SoS lower bounds for NGCA, giving further evidence for the low-degree conjecture.
- The SoS lower bound problem presents intrinsic challenges. We introduce algebro-combinatorial techniques to address them.

Thank you