A Positivity-First Approach to Sum-of-Squares (Lower Bounds) Over the Hypercube

Tim Kunisky

Courant Institute of Mathematical Sciences, NYU

September 22, 2020

I. Problem Setting

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}$$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①

$\boldsymbol{W} \sim \mathsf{GOE}(N)$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①

$\boldsymbol{W} \sim \mathsf{GOE}(N)$

ground state of
Sherrington-Kirkpatrick
(SK) spin glass model

$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①

$\boldsymbol{W} \sim \mathsf{GOE}(N)$

ground state of Sherrington-Kirkpatrick (SK) spin glass model

$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$

②

$\boldsymbol{W}$ = Laplacian of $G \sim \mathbb{Q}'$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①

$\boldsymbol{W} \sim \mathsf{GOE}(N)$

ground state of Sherrington-Kirkpatrick (SK) spin glass model

$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$

②

$\boldsymbol{W}$ = Laplacian of $G \sim \mathbb{Q}'$

maximum cut in sparse (deg $\approx \gamma$) random graph (regular or Erdős-Rényi)

$\mathsf{OPT}(\boldsymbol{W})/N \approx \frac{\gamma}{4} + \Theta(\gamma^{\frac{1}{2}})$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x}\in\{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①		②
$\boldsymbol{W} \sim \mathsf{GOE}(N)$	" $\xleftarrow{\gamma\uparrow\infty}$ "	$\boldsymbol{W}$ = Laplacian of $G \sim \mathbb{Q}'$
ground state of Sherrington-Kirkpatrick (SK) spin glass model		**maximum cut** in sparse (deg $\approx \gamma$) random graph (regular or Erdős-Rényi)
$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$		$\mathsf{OPT}(\boldsymbol{W})/N \approx \frac{\gamma}{4} + \Theta(\gamma^{\frac{1}{2}})$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①		②
$\boldsymbol{W} \sim \mathsf{GOE}(N)$	“ $\overset{\gamma \uparrow \infty}{\longleftarrow}$ ”	$\boldsymbol{W}$ = Laplacian of $G \sim \mathbb{Q}'$
ground state of Sherrington-Kirkpatrick (SK) spin glass model		**maximum cut** in sparse (deg $\approx \gamma$) random graph (regular or Erdős-Rényi)
$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$		$\mathsf{OPT}(\boldsymbol{W})/N \approx \frac{\gamma}{4} + \Theta(\gamma^{\frac{1}{2}})$

$$\mathsf{OPT}(\boldsymbol{W}) := \max_{\boldsymbol{x}\in\{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}, \text{ for random } \boldsymbol{W} \sim \mathbb{Q}.$$

①		②
$\boldsymbol{W} \sim \mathsf{GOE}(N)$	"$\xleftarrow{\gamma\uparrow\infty}$"	$\boldsymbol{W}$ = Laplacian of $G \sim \mathbb{Q}'$
ground state of Sherrington-Kirkpatrick (SK) spin glass model		**maximum cut** in sparse (deg $\approx \gamma$) random graph (regular or Erdős-Rényi)
$\mathsf{OPT}(\boldsymbol{W})/N \approx 1.526$		$\mathsf{OPT}(\boldsymbol{W})/N \approx \frac{\gamma}{4} + \Theta(\gamma^{\frac{1}{2}})$

But is it easy or hard to **compute** $\mathsf{OPT}(\boldsymbol{W})$?

Computational Tasks for $\mathsf{OPT}(W)$

Computational Tasks for $\mathsf{OPT}(\boldsymbol{W})$

1. Search: Compute $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{W})$ with $\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \approx \mathsf{OPT}(\boldsymbol{W})$.

- Local search
- Relax & round
- Message passing (Ahmed's talk—can get near-equality!)

Computational Tasks for $\mathsf{OPT}(\boldsymbol{W})$

1. Search: Compute $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{W})$ with $\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \approx \mathsf{OPT}(\boldsymbol{W})$.

- Local search
- Relax & round
- Message passing (Ahmed's talk—can get near-equality!)

2. Certification: Compute $c(\boldsymbol{W}) \in \mathbb{R}$ with $\mathsf{OPT}(\boldsymbol{W}) \leq c(\boldsymbol{W})$ *for all* $\boldsymbol{W}$, and as small as possible.

- LP relaxations (Sherali-Adams)
- SDP relaxations (Goemans-Williamson, sum-of-squares)

Computational Tasks for $\mathsf{OPT}(W)$

1. Search: Compute $x = x(W)$ with $x^\top W x \approx \mathsf{OPT}(W)$.

- Local search
- Relax & round
- Message passing (Ahmed's talk—can get near-equality!)

2. Certification: Compute $c(W) \in \mathbb{R}$ with $\mathsf{OPT}(W) \leq c(W)$ *for all* W, and as small as possible.

- LP relaxations (Sherali-Adams)
- SDP relaxations (Goemans-Williamson, sum-of-squares)

↑

This talk—what happens here?

The Spectral Certificate

The Spectral Certificate

The simplest way to bound $\mathsf{OPT}(\boldsymbol{W})$: ignore all special structure in $\boldsymbol{x}$, to get

$$\begin{aligned} \mathsf{OPT}(\boldsymbol{W}) &= \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \\ &\le \max_{\|\boldsymbol{x}\| = \sqrt{N}} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \\ &= \lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N \\ &\approx 2N \qquad\qquad (\text{for } \boldsymbol{W} \sim \mathsf{GOE}(N)) \end{aligned}$$

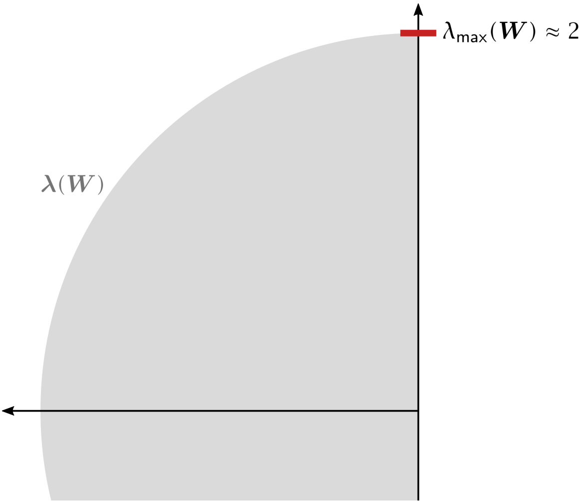
The Spectral Certificate

The simplest way to bound $\mathsf{OPT}(\boldsymbol{W})$: ignore all special structure in $\boldsymbol{x}$, to get

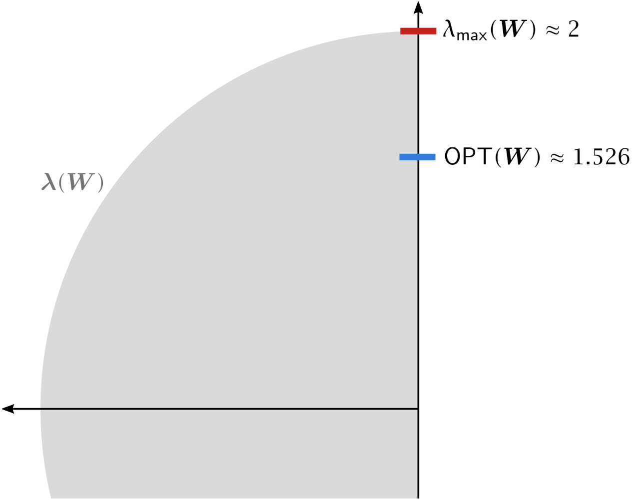
$$\begin{aligned}
\mathsf{OPT}(\boldsymbol{W}) &= \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \\
&\leq \max_{\|\boldsymbol{x}\| = \sqrt{N}} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \\
&= \lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N \\
&\approx 2N \qquad\qquad \text{(for } \boldsymbol{W} \sim \mathsf{GOE}(N)\text{)}
\end{aligned}$$

Afonso's talk (sort of)—indirect evidence suggests among poly(N)-time certificates this is **optimal!**

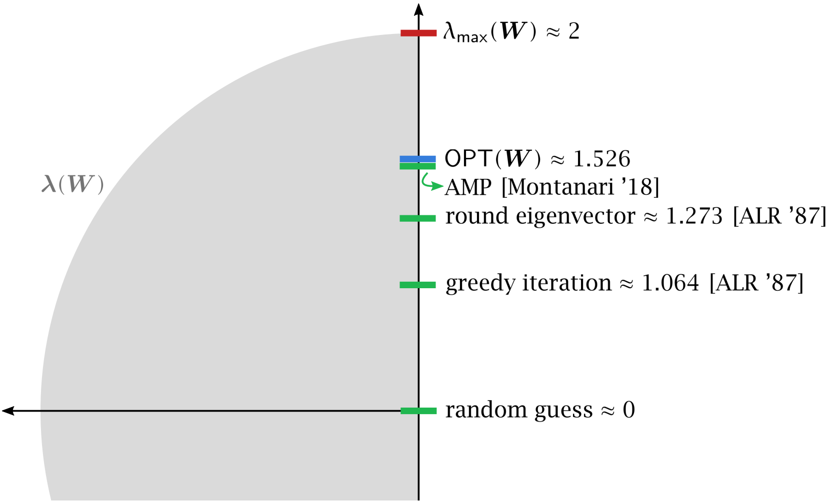
Search-Certification Gap in SK Hamiltonian



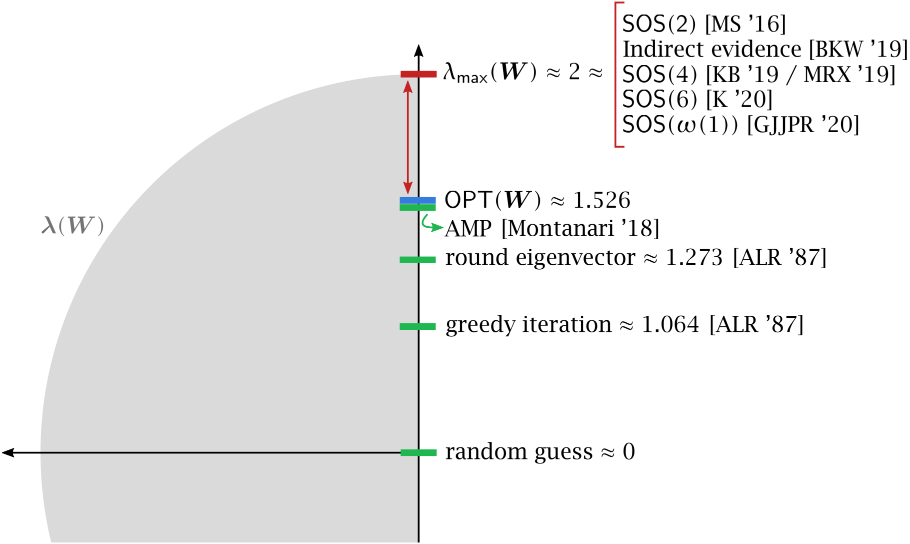
Search-Certification Gap in SK Hamiltonian



Search-Certification Gap in SK Hamiltonian



Search-Certification Gap in SK Hamiltonian



II. Sum-of-Squares Lower Bounds

Sum-of-Squares (SOS) Relaxations

Recall, $\mathsf{OPT}(\boldsymbol{W}) = \max\{\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} : \boldsymbol{x} \in \{\pm 1\}^N\}$.

Sum-of-Squares (SOS) Relaxations

Recall, $\mathsf{OPT}(\boldsymbol{W}) = \max\{\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} : x_i^2 - 1 = 0\}$.

Sum-of-Squares (SOS) Relaxations

Recall, $\mathsf{OPT}(W) = \max\{\boldsymbol{x}^\top W \boldsymbol{x} : x_i^2 - 1 = 0\}$.

$\widetilde{\mathbb{E}} : \mathbb{R}[\boldsymbol{x}]_{\leq 2d} \to \mathbb{R}$ is a degree $2d$ pseudoexpectation if

- Linear
- $\widetilde{\mathbb{E}}[1] = 1$ ("normalization")
- $\widetilde{\mathbb{E}}[(x_i^2 - 1)p(\boldsymbol{x})] = 0$ ("ideal annihilation")
- $\widetilde{\mathbb{E}}[p(\boldsymbol{x})^2] \geq 0$ ("positivity")

Sum-of-Squares (SOS) Relaxations

Recall, $\mathsf{OPT}(\boldsymbol{W}) = \max\{\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} : x_i^2 - 1 = 0\}$.

$\widetilde{\mathbb{E}} : \mathbb{R}[\boldsymbol{x}]_{\leq 2d} \to \mathbb{R}$ is a degree $2d$ pseudoexpectation if

- Linear
- $\widetilde{\mathbb{E}}[1] = 1$ ("normalization")
- $\widetilde{\mathbb{E}}[(x_i^2 - 1)p(\boldsymbol{x})] = 0$ ("ideal annihilation")
- $\widetilde{\mathbb{E}}[p(\boldsymbol{x})^2] \geq 0$ ("positivity")

Optimizing over these gives a certificate:

$$\mathsf{OPT}(\boldsymbol{W}) \leq \mathsf{SOS}_{2d}(\boldsymbol{W}) := \max_{\substack{\widetilde{\mathbb{E}} \text{ degree } 2d \\ \text{pseudoexpectation}}} \widetilde{\mathbb{E}}[\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}]$$

Sum-of-Squares (SOS) Relaxations

Recall, $\mathsf{OPT}(\boldsymbol{W}) = \max\{\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} : x_i^2 - 1 = 0\}$.

$\widetilde{\mathbb{E}} : \mathbb{R}[\boldsymbol{x}]_{\leq 2d} \to \mathbb{R}$ is a degree $2d$ pseudoexpectation if

- Linear
- $\widetilde{\mathbb{E}}[1] = 1$ (“normalization”)
- $\widetilde{\mathbb{E}}[(x_i^2 - 1)p(\boldsymbol{x})] = 0$ (“ideal annihilation”)
- $\widetilde{\mathbb{E}}[p(\boldsymbol{x})^2] \geq 0$ (“positivity”)

Optimizing over these gives a **very powerful** certificate:

$$\mathsf{OPT}(\boldsymbol{W}) \leq \mathsf{SOS}_{2d}(\boldsymbol{W}) := \max_{\substack{\widetilde{\mathbb{E}} \text{ degree } 2d \\ \text{pseudoexpectation}}} \widetilde{\mathbb{E}}[\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}]$$

SOS Lower Bound Constructions

Based on heuristic arguments, we think that, for any constant $2d$,

$$\mathsf{SOS}_{2d}(\boldsymbol{W}) \approx \lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N$$

SOS Lower Bound Constructions

Based on heuristic arguments, we think that, for any constant $2d$,

$$\mathsf{SOS}_{2d}(\boldsymbol{W}) \approx \lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N$$

To *prove* this, we need to build $\widetilde{\mathbb{E}}$ with

$$\lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N \approx \widetilde{\mathbb{E}}[\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}] \approx \langle \widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top], \boldsymbol{W} \rangle$$

SOS Lower Bound Constructions

Based on heuristic arguments, we think that, for any constant $2d$,

$$\mathsf{SOS}_{2d}(\boldsymbol{W}) \approx \lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N$$

To *prove* this, we need to build $\widetilde{\mathbb{E}}$ with

$$\lambda_{\mathsf{max}}(\boldsymbol{W}) \cdot N \approx \widetilde{\mathbb{E}}[\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}] \approx \langle \widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top], \boldsymbol{W} \rangle$$

Since $\mathsf{Tr}(\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top]) = N$, $\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top]$ needs most of its mass near the top few eigenvectors of $\boldsymbol{W}$. Roughly:

$\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top] \underset{\sim}{\propto}$ uniformly random low-dimensional projection

How Everyone Builds $\widetilde{\mathbb{E}}$—“Entries First”

How Everyone Builds $\widetilde{\mathbb{E}}$—"Entries First"

Argue using Bayesian ideas (pseudocalibration ~ Jess' talk), symmetry, or whatever else to predict for each monomial,

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S] := f_S(\boldsymbol{W})$$

How Everyone Builds $\widetilde{\mathbb{E}}$—“Entries First”

Argue using Bayesian ideas (pseudocalibration ∼ Jess’ talk), symmetry, or whatever else to predict for each monomial,

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S] := f_S(\boldsymbol{W})$$

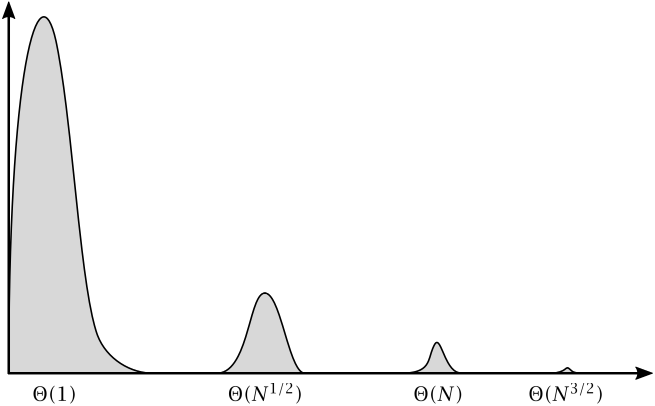
That leaves just one little detail...

$$\text{positivity} \iff \underbrace{\left(\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T]\right)_{S,T \in \binom{[N]}{\leq d}}}_{\text{pseudomoment matrix}} \succeq \boldsymbol{0}$$

These arguments are hard!

Why Entries-First Makes Things Hard

There is not much intuition for *why* we have positivity, and the matrices involved have complicated *multiscale* spectra:



“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

Desiderata:

- $\boldsymbol{G}^{(d)} \in \mathsf{Sym}^d(\mathbb{R}^N)$ (some symmetry)

“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

Desiderata:

- $\boldsymbol{G}^{(d)} \in \mathsf{Sym}^d(\mathbb{R}^N)$ (some symmetry)
- $G_\varnothing^{(0)} = \mathbb{E}[(G_S^{(d)})^2] = 1$ (normalization)

“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

Desiderata:

- $\boldsymbol{G}^{(d)} \in \mathsf{Sym}^d(\mathbb{R}^N)$ (some symmetry)
- $G_\emptyset^{(0)} = \mathbb{E}[(G_S^{(d)})^2] = 1$ (normalization)
- $G_{\{i,i\}+S}^{(d)} = G_S^{(d-2)}$ $(x_i^2 \boldsymbol{x}^S = \boldsymbol{x}^S)$

“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

Desiderata:

- $\boldsymbol{G}^{(d)} \in \mathsf{Sym}^d(\mathbb{R}^N)$ (some symmetry)
- $G_{\emptyset}^{(0)} = \mathbb{E}[(G_S^{(d)})^2] = 1$ (normalization)
- $G_{\{i,i\}+S}^{(d)} = G_S^{(d-2)}$ ($x_i^2 \boldsymbol{x}^S = \boldsymbol{x}^S$)
- $(G_{\{i\}+S}^{(d)})_{i=1}^N$ in top eigenspace of $\boldsymbol{W}$ ($\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}$ large)

“Positivity First” Using Surrogate Tensors

A different approach: build a random *surrogate* tensor $\boldsymbol{G}^{(d)} \in (\mathbb{R}^N)^{\otimes d}$ for $\boldsymbol{x}^{\otimes d}$, and take

$$\widetilde{\mathbb{E}}[\boldsymbol{x}^S \boldsymbol{x}^T] := \mathbb{E}[G_S^{(d)} G_T^{(d)}].$$

Desiderata:

- $\boldsymbol{G}^{(d)} \in \mathsf{Sym}^d(\mathbb{R}^N)$ (some symmetry)
- $G_\emptyset^{(0)} = \mathbb{E}[(G_S^{(d)})^2] = 1$ (normalization)
- $G_{\{i,i\}+S}^{(d)} = G_S^{(d-2)}$ ($x_i^2 \boldsymbol{x}^S = \boldsymbol{x}^S$)
- $(G_{\{i\}+S}^{(d)})_{i=1}^N$ in top eigenspace of $\boldsymbol{W}$ ($\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}$ large)

Otherwise, take $\boldsymbol{G}^{(d)}$ *as random as possible* $\rightsquigarrow$ condition canonical gaussian symmetric tensor.

What Comes Out

The result, upon squinting the right way, has the joint symmetry in $S + T$ we wanted! This has a diagrammatic interpretation:

$$\mathbb{E}[G_S^{(d)} G_T^{(d)}] \approx \sum_{F \text{ forest on } 2d \text{ leaves}} \mu(F) p_F(\boldsymbol{W}, S + T)$$

What Comes Out

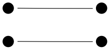
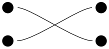
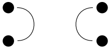
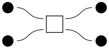
The result, upon squinting the right way, has the joint symmetry in $S + T$ we wanted! This has a diagrammatic interpretation:

$$\mathbb{E}[G_S^{(d)} G_T^{(d)}] \approx \sum_{F \text{ forest on } 2d \text{ leaves}} \mu(F) p_F(\boldsymbol{W}, S + T)$$

Ex. If $2d = 4$, $S = \{i, j\}$, $T = \{k, \ell\}$, $\boldsymbol{M} \approx \widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top]$ is the rescaled projector to top eigenspace of $\boldsymbol{W}$, then

$M_{ik}M_{j\ell}$ $\qquad$ $M_{i\ell}M_{jk}$ $\qquad$ $M_{ij}M_{k\ell}$ $\qquad$ $\sum_{a=1}^{N} M_{ia}M_{ja}M_{ka}M_{\ell a}$

Cancellations for Ideal Constraint: $i = j$

 +  +  − 2 

$M_{ik}M_{j\ell}$ $\quad$ $M_{i\ell}M_{jk}$ $\quad$ $M_{ij}M_{k\ell}$ $\quad$ $\sum_{a=1}^{N} M_{ia}M_{ja}M_{ka}M_{\ell a}$

Cancellations for Ideal Constraint: $i = j$

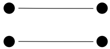
$$M_{ik}M_{i\ell} \qquad M_{i\ell}M_{ik} \qquad M_{k\ell} \qquad \sum_{a=1}^{N} M_{ia}^2 M_{ka} M_{\ell a}$$

Cancellations for Ideal Constraint: $i = j$

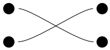
$$M_{ik}M_{i\ell} \qquad M_{i\ell}M_{ik} \qquad M_{k\ell} \qquad \sum_{a=1}^{N} M_{ia}^2 M_{ka} M_{\ell a}$$

Therefore, $\widetilde{\mathbb{E}}[x_i^2 x_k x_\ell] \approx \widetilde{\mathbb{E}}[x_k x_\ell]$.

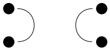
Cancellations for Ideal Constraint: $i = j = k = \ell$



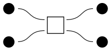
+



+



− 2



$$M_{ik}M_{j\ell} \qquad M_{i\ell}M_{jk} \qquad M_{ij}M_{k\ell} \qquad \sum_{a=1}^{N} M_{ia}M_{ja}M_{ka}M_{\ell a}$$

Cancellations for Ideal Constraint: $i = j = k = \ell$

$1 \qquad 1 \qquad 1 \qquad \sum_{a=1}^{N} M_{ia}^4$

Cancellations for Ideal Constraint: $i = j = k = \ell$

$1 \qquad 1 \qquad 1 \qquad \sum_{a=1}^{N} M_{ia}^4$

Therefore, $\widetilde{\mathbb{E}}[x_i^4] \approx 1$.

Remarkably, $\mu(F)$ is a *Möbius function* of a partial ordering of forests, meaning such things always happen!

Actual Results

- If instead of $\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top] \approx$ random *low-rank* projection want *high-rank* $(N - o(N))$ projection, then can build $\widetilde{\mathbb{E}}$ of degree $\sim \log N / \log\log N$.

Actual Results

- If instead of $\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top] \approx$ random *low-rank* projection want *high-rank* $(N - o(N))$ projection, then can build $\widetilde{\mathbb{E}}$ of degree $\sim \log N / \log\log N$.

- For low-rank projections (lower bounds for SK model), can improve from degree 4 to degree 6, then blocked by technical difficulties stemming from approximation in conditioning calculations.

Actual Results

- If instead of $\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top] \approx$ random *low-rank* projection want *high-rank* $(N - o(N))$ projection, then can build $\widetilde{\mathbb{E}}$ of degree $\sim \log N / \log\log N$.

- For low-rank projections (lower bounds for SK model), can improve from degree 4 to degree 6, then blocked by technical difficulties stemming from approximation in conditioning calculations.

- Driving the proofs are the combinatorics relating the conditional distribution of $\boldsymbol{G}^{(d)}$ (positivity constraints) with the Möbius function (ideal constraints).

Thank You!

Sketch of Conditioning Calculations

If $\boldsymbol{P}$ is projection to top eigenspace, $P_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle$, then amounts to calculating projection to the subspace

$$\mathsf{span}(\{\boldsymbol{v}_1 \odot \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N \odot \boldsymbol{v}_N\}) \odot \mathsf{Sym}^{d-2}(\mathbb{R}^N) \subset \mathsf{Sym}^{d}(\mathbb{R}^N)$$

Sketch of Conditioning Calculations

If $\boldsymbol{P}$ is projection to top eigenspace, $P_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle$, then amounts to calculating projection to the subspace

$$\mathsf{span}(\{\boldsymbol{v}_1 \odot \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N \odot \boldsymbol{v}_N\}) \odot \mathsf{Sym}^{d-2}(\mathbb{R}^N) \subset \mathsf{Sym}^{d}(\mathbb{R}^N)$$

In homogeneous polynomials, this is the *ideal* of $\langle \boldsymbol{v}_i, \boldsymbol{z} \rangle^2$.

Sketch of Conditioning Calculations

If $\boldsymbol{P}$ is projection to top eigenspace, $P_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle$, then amounts to calculating projection to the subspace

$$\mathsf{span}(\{\boldsymbol{v}_1 \odot \boldsymbol{v}_1, \dots, \boldsymbol{v}_N \odot \boldsymbol{v}_N\}) \odot \mathsf{Sym}^{d-2}(\mathbb{R}^N) \subset \mathsf{Sym}^{d}(\mathbb{R}^N)$$

In homogeneous polynomials, this is the *ideal* of $\langle \boldsymbol{v}_i, \boldsymbol{z} \rangle^2$.

The Frobenius inner product becomes the *apolar inner product*, and, remarkably, the orthogonal complement is the *multiharmonic polynomials* $\langle \boldsymbol{v}_i, \boldsymbol{\partial} \rangle^2 p(\boldsymbol{z}) = 0$.

Sketch of Conditioning Calculations

If $\boldsymbol{P}$ is projection to top eigenspace, $P_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle$, then amounts to calculating projection to the subspace

$$\mathsf{span}(\{\boldsymbol{v}_1 \odot \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N \odot \boldsymbol{v}_N\}) \odot \mathsf{Sym}^{d-2}(\mathbb{R}^N) \subset \mathsf{Sym}^{d}(\mathbb{R}^N)$$

In homogeneous polynomials, this is the *ideal* of $\langle \boldsymbol{v}_i, \boldsymbol{z} \rangle^2$.

The Frobenius inner product becomes the *apolar inner product*, and, remarkably, the orthogonal complement is the *multiharmonic polynomials* $\langle \boldsymbol{v}_i, \boldsymbol{\partial} \rangle^2 p(\boldsymbol{z}) = 0$.

Using an analogy with harmonic polynomials, we can heuristically calculate by suggesting a “Green’s function” for this system.