Sum-of-Squares and Spectral Algorithms

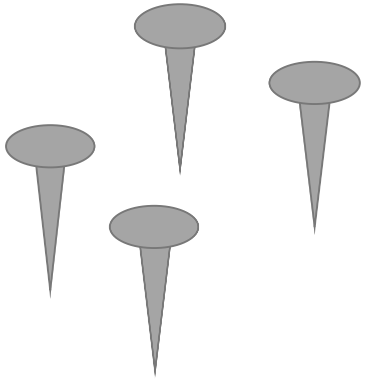
Tselil Schramm

June 23, 2017

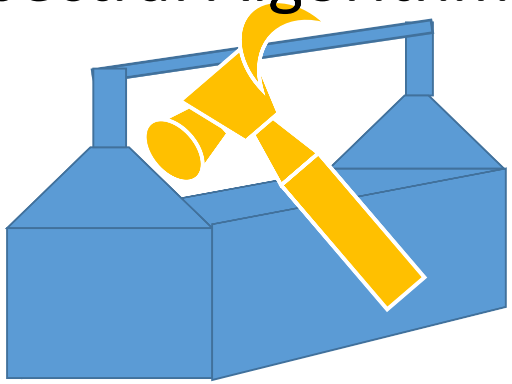
Workshop on SoS @ STOC 2017

Spectral algorithms as a tool for analyzing SoS.

SoS Semidefinite Programs

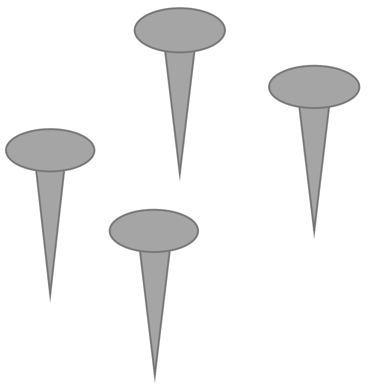


Spectral Algorithms

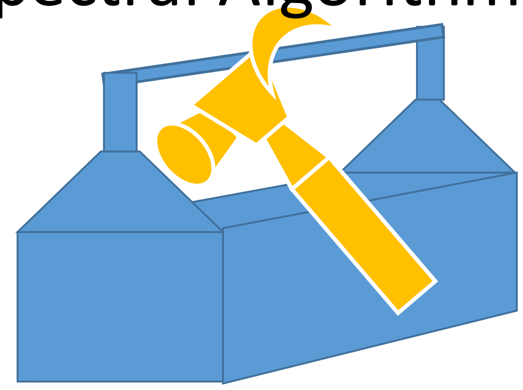


SoS suggests a new family of spectral algorithms!

SoS Semidefinite Programs



Spectral Algorithms



Average-Case & Structured Instances

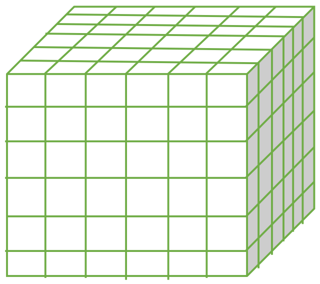
Average Case SoS/Spectral Algorithms

- Tensor Decomposition/Dictionary Learning [Barak-Kelner-Steurer'14, Ge-Ma'15, Ma-Shi-Steurer'16]
- Planted Sparse Vector [Barak-Brandão-Harrow-Kelner-Steurer-Zhou'12, Barak-Kelner-Steurer'14]
- Tensor Completion [Barak-Moitra'16, Potechin-Steurer'17]
- Refuting Random CSPs [Allen-O'Donnell-Witmer'15, Raghavendra-Rao-S'17]
- Tensor Principal Components Analysis [Hopkins-Shi-Steurer'15,Bhattiprolu-Guruswami-Lee'16, Raghavendra-Rao-S'17]

Average Case SoS/Spectral Algorithms

- Tensor Decomposition/Dictionary Learning
  [Barak-Kelner-Steurer'14, Ge-Ma'15, Ma-Shi-Steurer'16]
- Planted Sparse Vector [Barak-Brandão-Harrow-Kelner-Steurer-Zhou'12, Barak-Kelner-Steurer'14]
- Tensor Completion [Barak-Moitra'16, Potechin-Steurer'17]
- Refuting Random CSPs [Allen-O'Donnell-Witmer'15, Raghavendra-Rao-S'17]
- Tensor Principal Components Analysis
  [Hopkins-Shi-Steurer'15,Bhattiprolu-Guruswami-Lee'16, Raghavendra-Rao-S'17]

Tensor Principle Components Analysis (TPCA)

$T =$ 

$n \times n \times n$

Want ``max tensor singular value/vector'':

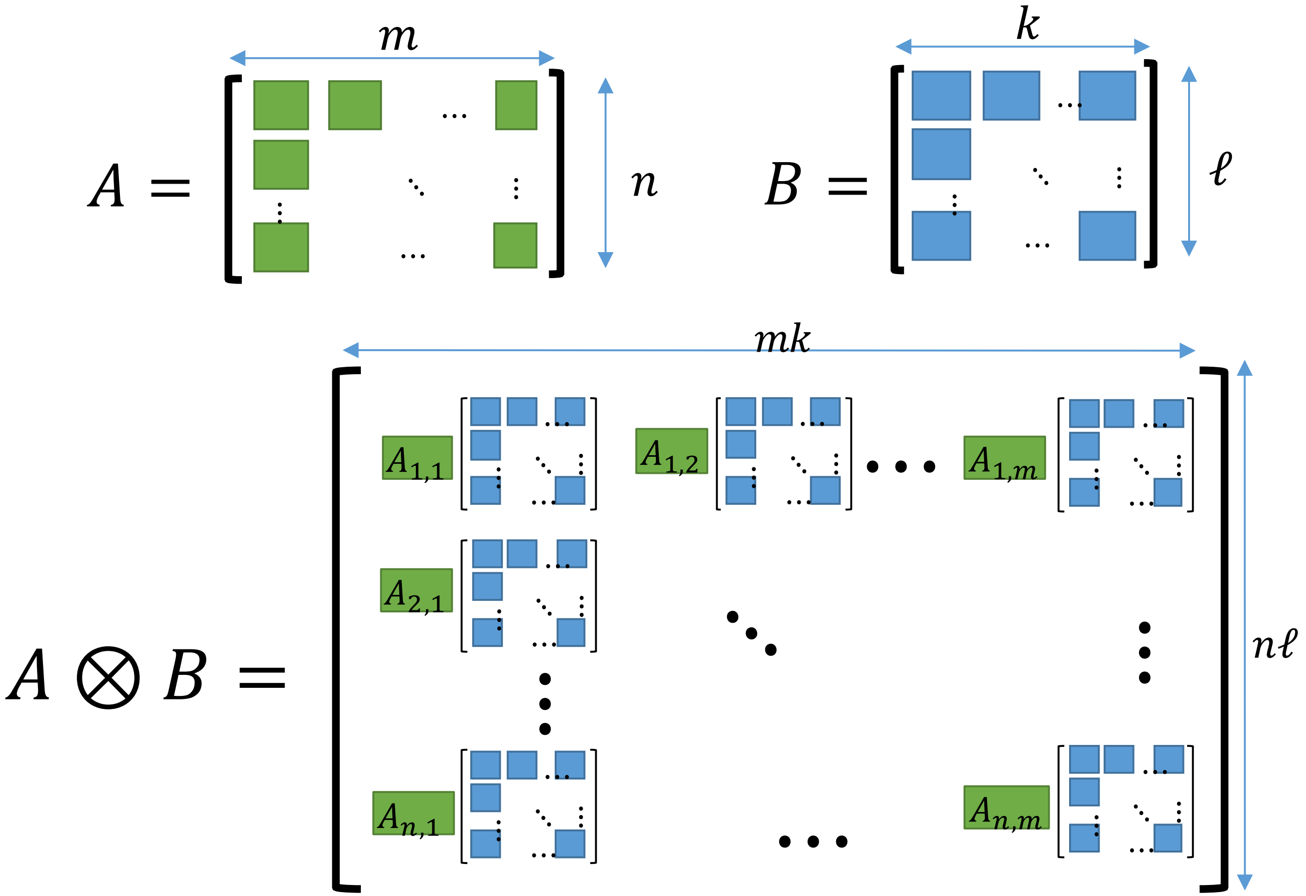
$$\sigma^* = \max_{x \in \mathbb{S}_{n-1}} \langle T, x^{\otimes 3} \rangle \quad \text{and} \quad x^* = \operatorname*{argmax}_{x \in \mathbb{S}_{n-1}} \langle T, x^{\otimes 3} \rangle$$

NP-hard in worst case.

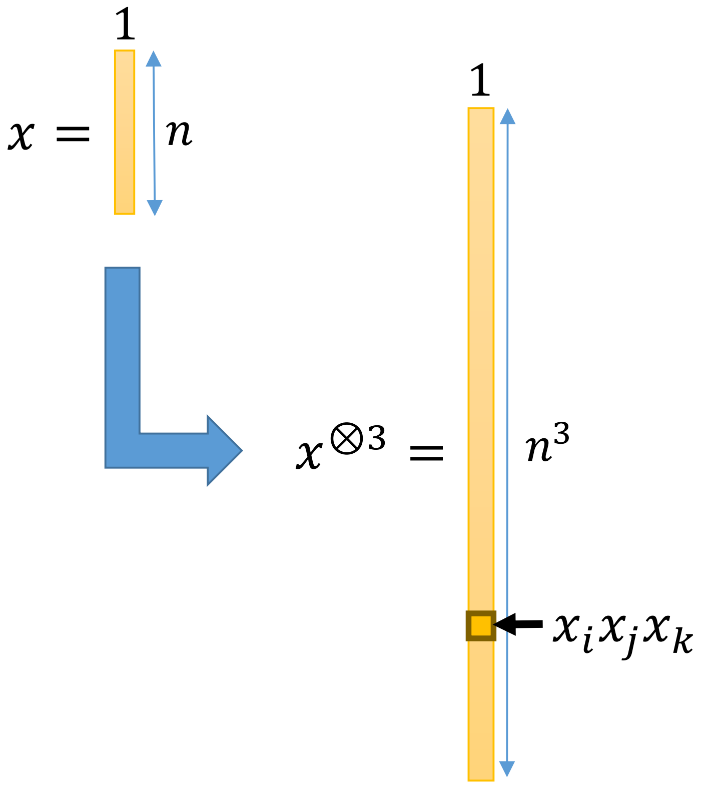
This ⊗ notation…

Definition

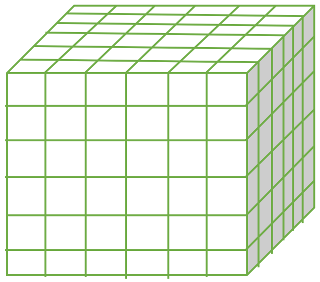
Kronecker/tensor product:



e.g. tensor power of x:



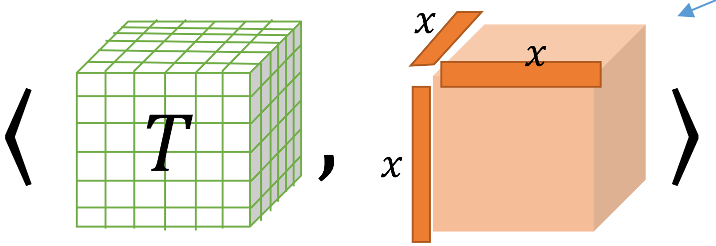
Tensor Principle Components Analysis (TPCA)

$T =$ 

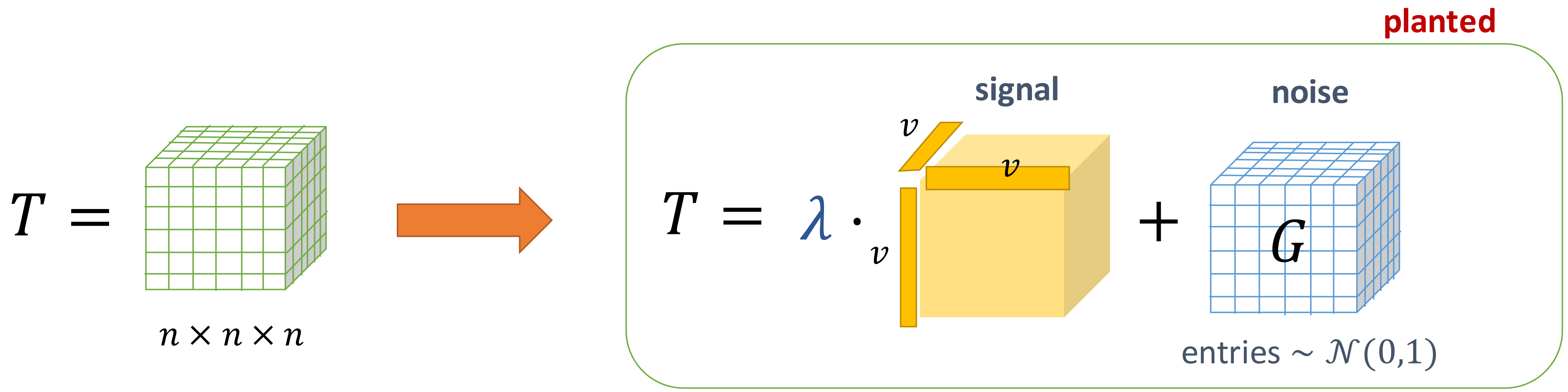
$n \times n \times n$

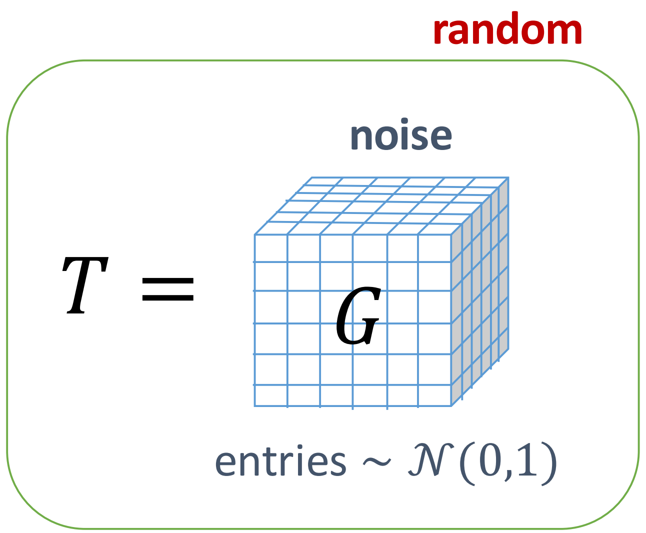
Want ``max tensor singular value/vector'': NP-hard in worst case.

$$\sigma^* = \max_{x \in \mathbb{S}_{n-1}} \langle T, x^{\otimes 3} \rangle \quad \text{and} \quad x^* = \operatorname{argmax}_{x \in \mathbb{S}_{n-1}} \langle T, x^{\otimes 3} \rangle$$



“Spiked” tensor model for TPCA [Montanari-Richard’14]





Search: find v in planted case

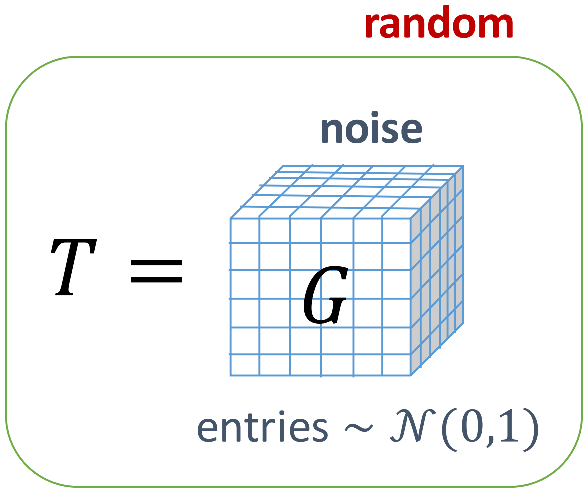
Distinguishing: planted or random case?

Refutation: certify upper bound on $\max_x \langle T, x^{\otimes 3} \rangle$ in random case

The Plan

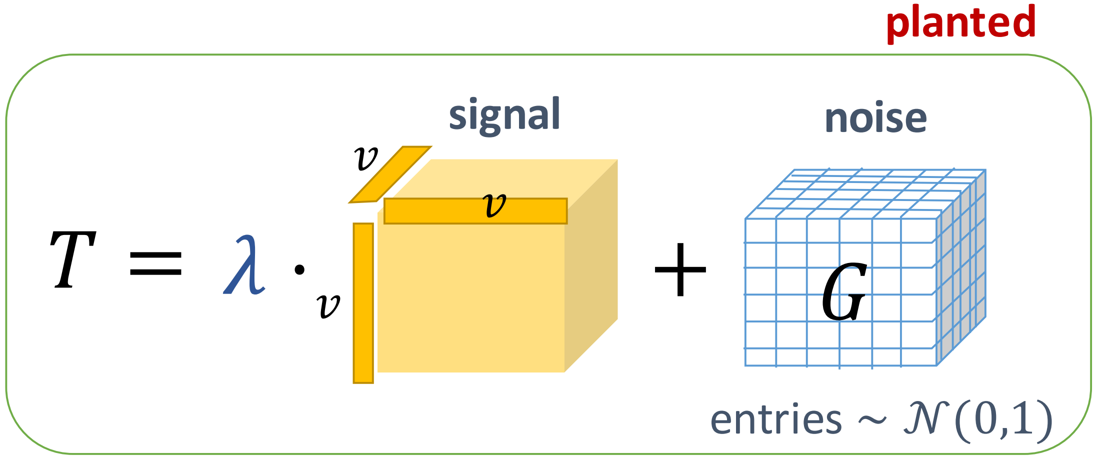
Refutation: certify upper bound on $\max_x \langle T, x^{\otimes 3} \rangle$ in random case

1. SoS suggests a family of spectral algorithms
2. Naïve spectral algorithm
3. Improving with SoS spectral algorithms



Search: find v in planted case

4. Use SoS analysis to get fast algorithms



Degree-D SoS

Solve for

$\deg(p) \leq D$

$$\widetilde{\mathbb{E}}: p(x) \to \mathbb{R}$$

n variables

Linearity:

$$\widetilde{\mathbb{E}}[a \cdot p(x) + b \cdot q(x)] = a \cdot \widetilde{\mathbb{E}}[p(x)] + b \cdot \widetilde{\mathbb{E}}[q(x)]$$

Fixed Scalars:

$$\widetilde{\mathbb{E}}[1] = 1$$

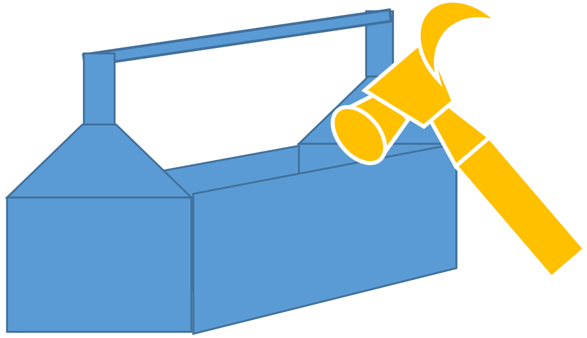
Non-negative squares:

$$\widetilde{\mathbb{E}}[q(x)^2] \geq 0 \quad \deg(q) \leq \frac{D}{2}$$

+ problem-specific constraints, e.g.

$$\widetilde{\mathbb{E}}[\|x\|^2] = 1$$

SoS suggests spectral algorithms



If we want to bound $f(x)$… associate some matrix with f and then

Rearrange entries along "monomial symmetries"

Apply degree-D SoS polynomial inequalities

Cauchy-Schwarz,
Jensen's Inequality (for squares), …

Use problem-specific constraints (e.g. $x_i^2 = 1$)

SoS captures spectral algorithms

Theorem

$$\widetilde{\mathbb{E}}[f(x)] \leq \lambda_{\max}(f)$$

Definition

$$\lambda_{\max}(f) = \operatorname{argmin}\{\lambda_{\max}(F)\}$$

F symmetric

$f(x) = \langle F, x^{\otimes 2d} \rangle$

symmetric

matrix representation

of $f(x)$

$$f(x) = \langle F, x^{\otimes 2d} \rangle$$

e.g. $$x_1^2 + 4x_1x_2 + x_2^2 = \left\langle \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{bmatrix} \right\rangle$$

SoS captures spectral algorithms

Theorem

$$\widetilde{\mathbb{E}}[f(x)] \le \lambda_{\max}(f)$$

Proof

symmetric
matrix representation
of $f(x)$

$$\widetilde{\mathbb{E}}[f(x)] = \widetilde{\mathbb{E}}\left[\langle F, x^{\otimes 2d}\rangle\right]$$

$\lambda = \lambda_{\max}(F)$ $\quad$ ≥ 0

$$0 \preccurlyeq \lambda \cdot Id - F = \sum \sigma_i u_i u_i^{\top}$$

if $2d \le D$

$$\widetilde{\mathbb{E}}\left\langle \lambda \cdot Id - F, x^{\otimes 2d}\right\rangle = \widetilde{\mathbb{E}} \underbrace{\sum\left(\sqrt{\sigma_i}\cdot\langle u_i, x^{\otimes d}\rangle\right)^2}_{\text{sum of degree-}d\text{ squares}} \geq 0$$

SoS captures spectral algorithms

Theorem

$$\boxed{\widetilde{\mathbb{E}}[f(x)] \le \lambda_{\max}(f)}$$

Proof

$$\widetilde{\mathbb{E}}[f(x)] = \widetilde{\mathbb{E}}\left[\langle F, x^{\otimes 2d}\rangle\right]$$

sum of degree-d squares

if $2d \le D$ $$\le \widetilde{\mathbb{E}}\left[\langle F, x^{\otimes 2d}\rangle\right] + \widetilde{\mathbb{E}}\left[\langle \lambda \cdot Id - F, x^{\otimes 2d}\rangle\right]$$

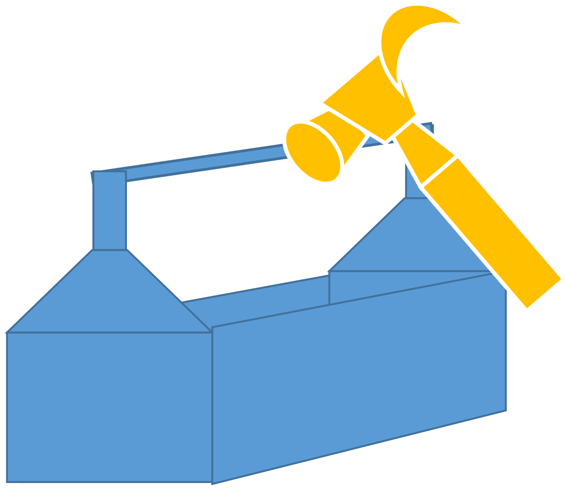
By linearity $$= \widetilde{\mathbb{E}}\left[\langle \lambda \cdot Id, x^{\otimes 2d}\rangle\right]$$

$$= \lambda \cdot \widetilde{\mathbb{E}}\left[\|x\|^{2d}\right]$$

squares on diagonal

$$\begin{bmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{bmatrix}$$

What kind of spectral algorithms?



Choose best matrix representation F by:

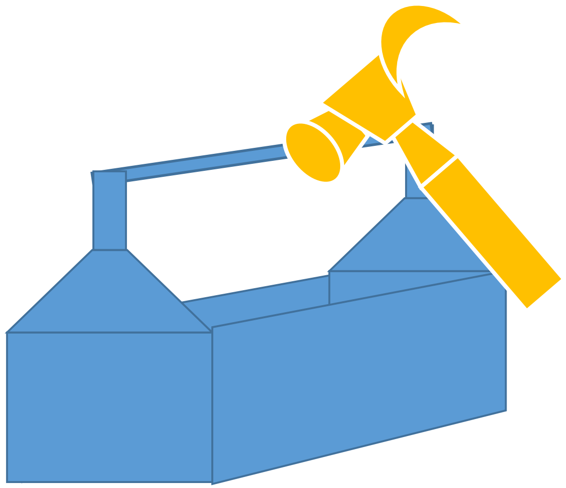
Rearranging entries along “symmetries” of $x^{\otimes d}$

Applying degree-D SoS polynomial inequalities

Cauchy-Schwarz,
Jensen’s Inequality (for squares), …

Problem-specific constraints (e.g. $x_i^2 = 1$)

What kind of spectral algorithms?



Choose best matrix representation F by:

Rearranging entries along "symmetries" of $x^{\otimes d}$

Applying degree-D SoS polynomial inequalities

Cauchy-Schwarz,
Jensen's Inequality (for squares), …

Problem-specific constraints (e.g. $x_i^2 = 1$)

SoS suggests *several* spectral algorithms

matrix representation
of $f(x)$

$$\widetilde{\mathbb{E}}[f(x)] = \widetilde{\mathbb{E}}\left[\langle F, x^{\otimes 2d}\rangle\right] \leq \lambda \cdot \widetilde{\mathbb{E}}[\|x\|^{2d}]$$

choice of F may affect λ!

SoS suggests *several* spectral algorithms

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d} \rangle = \langle F_2, x^{\otimes d} \rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

unit vector

$$f(x) = \mathbb{E}_{g \sim \mathcal{N}(0, Id)}[\underbrace{\langle x, g \rangle}_{\sim \mathcal{N}(0,1)}{}^4] = 3$$

SoS suggests *several* spectral algorithms

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d} \rangle = \langle F_2, x^{\otimes d} \rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

$$f(x) = \mathbb{E}_{g \sim \mathcal{N}(0,Id)}[\underbrace{\langle x, \overset{\substack{\text{unit vector}\\ \downarrow}}{g} \rangle}_{\sim \mathcal{N}(0,1)}{}^4] = \langle \mathbb{E}[g^{\otimes 4}], x^{\otimes 4} \rangle$$

$$\mathbb{E}[(g \otimes g)(g \otimes g)^\top]_{ij,k\ell} = \mathbb{E}[g_i g_j g_k g_\ell] = \begin{cases} 3 & i = j = k = \ell \\ 1 & i, j, k, \ell \text{ two distinct pairs} \\ 0 & \text{any index with odd multiplicity} \end{cases}$$

SoS suggests *several* spectral algorithms

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d}\rangle = \langle F_2, x^{\otimes d}\rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

$$f(x) = \mathbb{E}_{g\sim\mathcal{N}(0,Id)}[\underbrace{\langle x, \overset{\text{unit vector}}{g}\rangle}_{\sim \mathcal{N}(0,1)}{}^4] = \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle$$

$$\mathbb{E}[(g\otimes g)(g\otimes g)^\top] = \begin{array}{c|cc|cc|} & ii & jj & ij & ji \\ \hline ii & 3 & 1 & & \\ & & \ddots & & 0 \\ jj & 1 & 3 & & \\ \hline & & & \ddots & \\ ij & & & 1 & 1 \\ & & 0 & & \ddots \\ ji & & & 1 & 1 \\ \hline \end{array}$$

eigenvalue is $n \gg 3$

Rearranging entries along "symmetries" of $x^{\otimes d}$

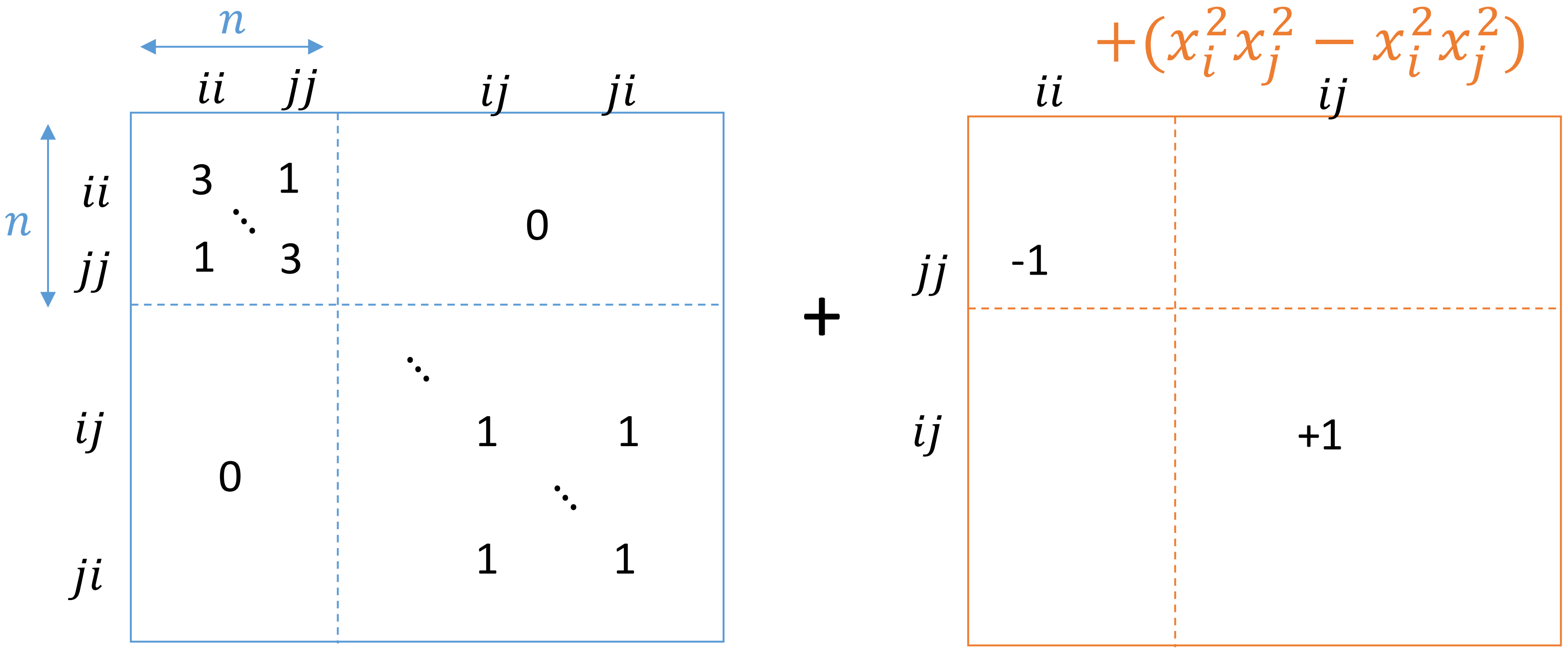
Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d} \rangle = \langle F_2, x^{\otimes d} \rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

unit vector

$$f(x) = \mathbb{E}_{g \sim \mathcal{N}(0, Id)}[\underbrace{\langle x, g \rangle}_{\sim \mathcal{N}(0,1)}{}^4]$$

$$= \langle \mathbb{E}[g^{\otimes 4}], x^{\otimes 4} \rangle$$



Rearranging entries along “symmetries” of $x^{\otimes d}$

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d}\rangle = \langle F_2, x^{\otimes d}\rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

$$f(x) = \mathbb{E}_{g\sim\mathcal{N}(0,Id)}[\underbrace{\langle x, \overset{\text{unit vector}}{\overset{\downarrow}{g}}\rangle}_{\sim\mathcal{N}(0,1)}{}^4]$$

$$= \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle$$

$$\begin{array}{c|cc|cc}
 & \overset{n}{\overleftrightarrow{ii \quad jj}} & & ij & ji \\
\hline
n\ \ ii & 3 \ \ddots & 1 & & 0 \\
\ \ jj & 0 & 3 & & \\
\hline
ij & & & \ddots\ 2 & 1 \\
ji & 0 & & 1 & \ddots\ 1
\end{array}$$

Rearranging entries along "symmetries" of $x^{\otimes d}$

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d} \rangle = \langle F_2, x^{\otimes d} \rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

$$f(x) = \mathbb{E}_{g \sim \mathcal{N}(0, Id)}[\underbrace{\langle \overset{\text{unit vector}}{x}, g \rangle}_{\sim \mathcal{N}(0,1)}{}^4]$$

$$= \langle \mathbb{E}[g^{\otimes 4}], x^{\otimes 4} \rangle$$

$$\begin{array}{c|cc|cc|}
 & \overset{\longleftarrow n \longrightarrow}{ii} & jj & ij & ji \\
\hline
ii & 3 & 1 & & \\
 & \ddots & & 0 & \\
jj & 0 & 3 & & \\
\hline
 & & & \ddots & \\
ij & & & 2 & 1 \\
 & 0 & & & \ddots \\
ji & & & 1 & 1 \\
\hline
\end{array}$$

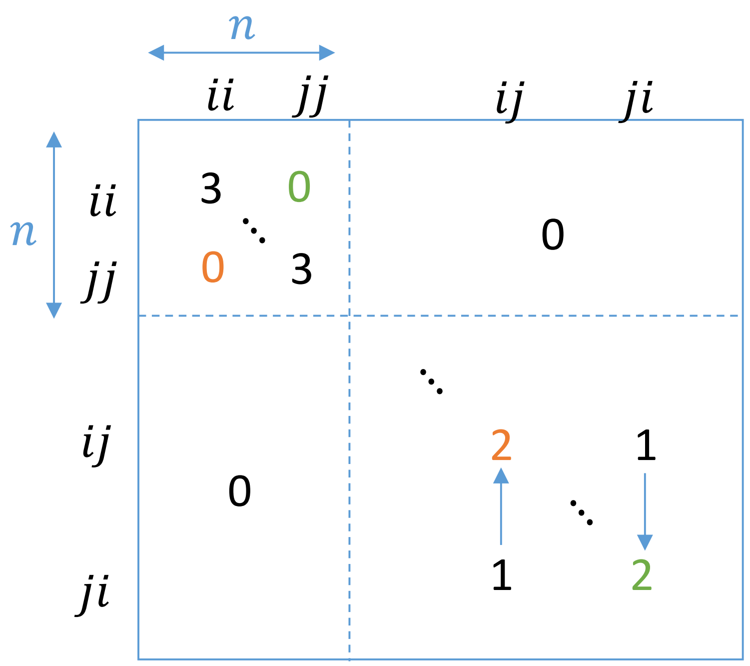
Rearranging entries along "symmetries" of $x^{\otimes d}$

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d}\rangle = \langle F_2, x^{\otimes d}\rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

$$f(x) = \mathbb{E}_{g\sim\mathcal{N}(0,Id)}[\underbrace{\langle \overset{\text{unit vector}}{x}, g\rangle}_{\sim\,\mathcal{N}(0,1)}{}^4]$$

$$= \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle$$



Rearranging entries along "symmetries" of $x^{\otimes d}$

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d}\rangle = \langle F_2, x^{\otimes d}\rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

unit vector

$$f(x) = \mathbb{E}_{g\sim\mathcal{N}(0,Id)}\Big[\underbrace{\langle x, g\rangle}_{\sim\mathcal{N}(0,1)}{}^4\Big]$$

$$= \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle$$

$$\begin{array}{c|cc|cc|}
 & ii & jj & ij & ji \\
\hline
ii & 3 & 0 & & \\
 & \ddots & & 0 & \\
jj & 0 & 3 & & \\
\hline
 & & & \ddots & \\
ij & & & 3 & 0 \\
 & 0 & & & \ddots \\
ji & & & 0 & 3 \\
\hline
\end{array} = 3 \cdot Id$$

The index blocks $ii, \dots, jj$ are each of size n.

eigenvalues are 3!

Rearranging entries along “symmetries” of $x^{\otimes d}$

Claim

There exist $f(x)$ with representations F_1, F_2 such that $f(x) = \langle F_1, x^{\otimes d}\rangle = \langle F_2, x^{\otimes d}\rangle$ but $\lambda(F_1) \gg \lambda(F_2)$.

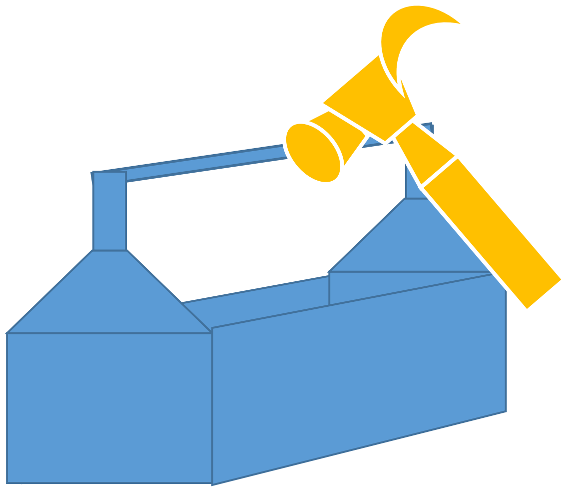
unit vector

$$f(x) = \mathbb{E}_{g\sim\mathcal{N}(0,Id)}[\underbrace{\langle x, g\rangle}_{\sim \mathcal{N}(0,1)}{}^4]$$

$$= \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle$$

$$f(x) = \langle\, \mathbb{E}[g^{\otimes 4}], x^{\otimes 4}\rangle = \langle\, 3 \cdot Id, x^{\otimes 4}\rangle = 3$$

What kind of spectral algorithms?



Choose best matrix representation by:

☑ Rearranging entries along “symmetries” of $x^{\otimes d}$

Applying degree-D SoS polynomial inequalities

Cauchy-Schwarz,
Jensen’s Inequality (for squares), …

Tensor Norm Refutation

random case, noise only

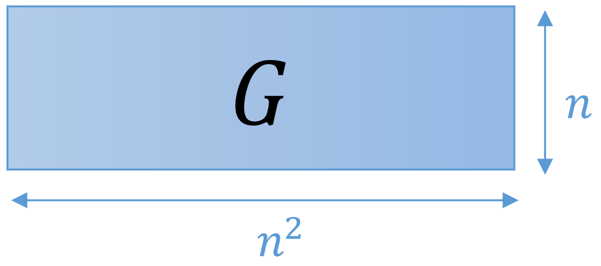
$$\max_{x \in \mathbb{S}_n} \langle G, x^{\otimes 3} \rangle \leq O(\sqrt{n}) \text{ with high probability over } G$$

Claim

"Simple" spectral algorithm can only certify $O(n)$.

Proof:

$$\langle G, x^{\otimes 3} \rangle \leq \sigma_{\max}(F_G)$$



Representations all the same because G is symmetric with iid entries

Gordon's Theorem $\rightarrow \sigma_{\max}(G) \approx n$

SoS Cauchy-Schwarz

Claim

$$\widetilde{\mathbb{E}}\langle f, g\rangle \leq \left(\widetilde{\mathbb{E}}\|g\|^2\right)^{1/2}\left(\widetilde{\mathbb{E}}\|f\|^2\right)^{1/2} \text{ if } D \geq 2\deg(f), 2\deg(g).$$

Proof:

$$\mathbb{E}\langle f, g\rangle \leq \frac{1}{2}\mathbb{E}\|f\|^2 + \frac{1}{2}\mathbb{E}\|g\|^2$$

$$\widetilde{\mathbb{E}}\langle f, g\rangle \leq \frac{1}{2}\widetilde{\mathbb{E}}\|f\|^2 + \frac{1}{2}\widetilde{\mathbb{E}}\|g\|^2$$

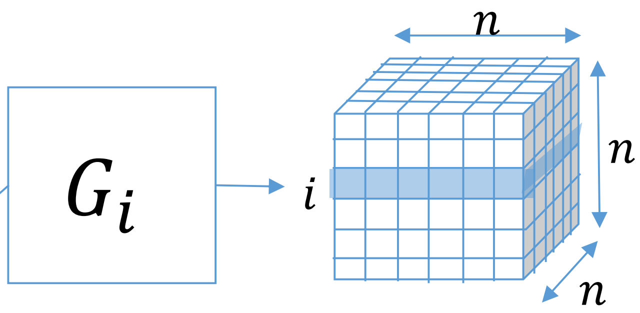
degree $\leq D$ square

$$0 \leq \frac{1}{2}\|f - g\|^2$$

Cauchy-Schwarz for Tensor PCA Refutation

Theorem

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, \overset{\text{unit vector}}{x}{}^{\otimes 3}\rangle \leq n^{3/4}$$



Proof:

$$\langle G, x^{\otimes 3}\rangle = \sum_i x_i \cdot x^\top G_i x$$

Cauchy-Schwarz

$$\leq \|x\| \sqrt{\sum_i (x^\top G_i x)^2} \leq \|x\| \sqrt{\sum_i (x^{\otimes 2})^\top (G_i \otimes G_i) x^{\otimes 2}}$$

Cauchy-Schwarz for Tensor PCA Refutation

Theorem

unit vector

$$\tilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, x^{\otimes 3}\rangle \le n^{3/4}$$

SoS analysis → spectral algorithm for refutation!

Proof:

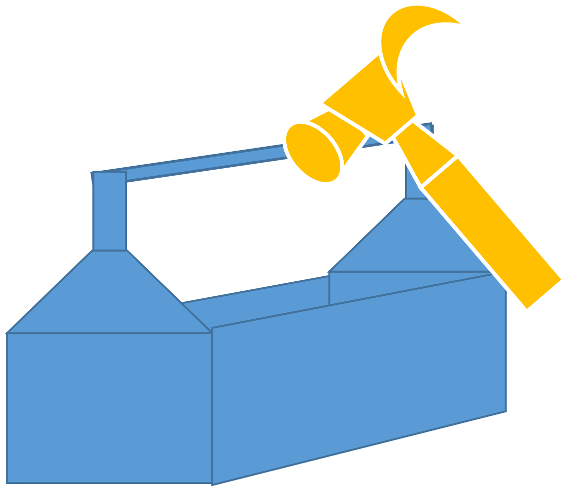
$$\langle G, x^{\otimes 3}\rangle = \sum_i x_i \cdot x^\top G_i x$$

Cauchy-Schwarz

eigenvalues are $\approx n^{3/2}$

$$\le \|x\| \sqrt{\left\langle \sum_i G_i \otimes G_i , x^{\otimes 4} \right\rangle} \le \|x\|^3 \left\| \sum_i G_i \otimes G_i \right\|^{1/2}$$

What kind of spectral algorithms?



Choose best matrix representation by:

☑ Rearranging entries along "symmetries" of $x^{\otimes d}$

Applying degree-D SoS polynomial inequalities

☑ Cauchy-Schwarz,
Jensen's Inequality (for squares), …

Better Approx (in more time)

Theorem

unit vector

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, x^{\otimes 3}\rangle \leq n^{3/4}$$

Theorem

unit vector

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, x^{\otimes 3}\rangle \leq O\left(\frac{n^{3/4}}{D^{1/4}}\right)$$ (in time n^D)

But actually, $\max_{x \in \mathbb{S}_n} \langle G, x^{\otimes 3}\rangle \leq O(\sqrt{n})$.

Information-theoretically, can certify $\leq O(\sqrt{n})$ in time 2^n (epsilon net).

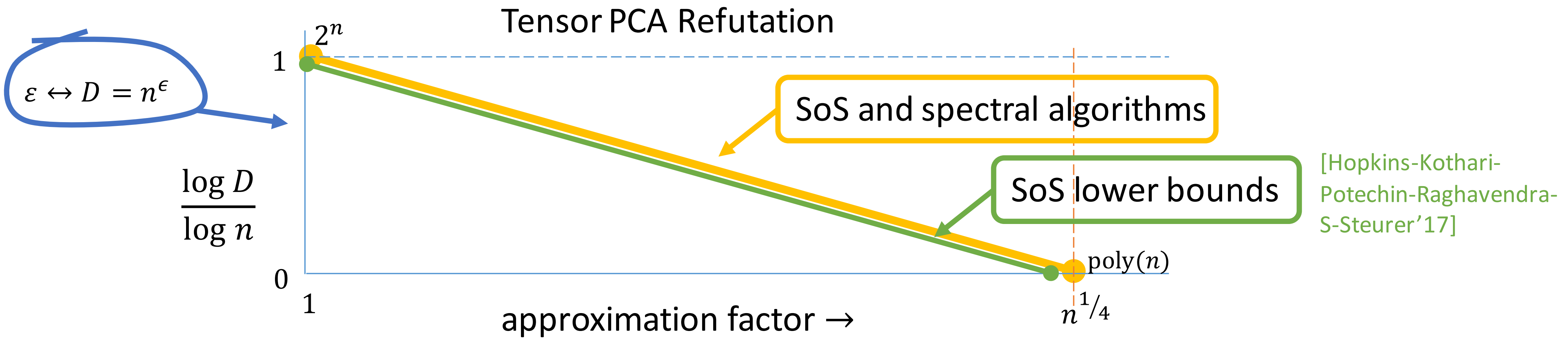
Better Approx (in more time)

Theorem

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, \overset{\text{unit vector}}{x}{}^{\otimes 3}\rangle \le n^{3/4}$$

Theorem

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, \overset{\text{unit vector}}{x}{}^{\otimes 3}\rangle \le O\left(\frac{n^{3/4}}{D^{1/4}}\right)$$ (in time n^D)



Jensen's Inequality

For d a power of 2, $D \geq d \cdot \deg(f)$

$$\boxed{\widetilde{\mathbb{E}}[f(x)] \leq \left(\widetilde{\mathbb{E}}[f(x)^d]\right)^{1/d}}$$

Proof by induction on d...

$$0 \leq \widetilde{\mathbb{E}}\left(f(x)^d - \widetilde{\mathbb{E}}[f(x)^d]\right)^2$$

$$\widetilde{\mathbb{E}}[f(x)^d]^2 \leq \widetilde{\mathbb{E}}[f(x)^{2d}]$$

$$\widetilde{\mathbb{E}}[f(x)^d] \leq \left(\widetilde{\mathbb{E}}[f(x)^{2d}]\right)^{1/2}$$ apply inductive hypothesis

Jensen’s Inequality

For d a power of 2, $D \geq d \cdot \deg(f)$

$$\boxed{\widetilde{\mathbb{E}}[f(x)] \leq \left(\widetilde{\mathbb{E}}[f(x)^d]\right)^{1/d}}$$

We can take advantage of increased symmetry
in higher-degree polynomials
(more matrix representations)

Better Approximation

Theorem

$$\tilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, \overset{\text{unit vector}}{x}{}^{\otimes 3}\rangle \le O\left(\frac{n^{3/4}}{D^{1/4}}\right)$$

(in time n^D)

Proof:

$$\tilde{\mathbb{E}}\langle G, x^{\otimes 3}\rangle = \tilde{\mathbb{E}} \sum_i x_i \cdot x^\top G_i x \le \left(\tilde{\mathbb{E}}\left\langle x^{\otimes 4}, \sum_i G_i \otimes G_i\right\rangle\right)^{1/2}$$

Jensen's inequality for d some power of 2

$$\le \left(\tilde{\mathbb{E}}\left\langle x^{\otimes 4}, \sum_i G_i \otimes G_i\right\rangle^d\right)^{1/2d}$$

Better Approximation

Theorem

unit vector

$$\tilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, x^{\otimes 3}\rangle \le O\left(\frac{n^{3/4}}{D^{1/4}}\right)$$

(in time n^D)

Proof:

$$\tilde{\mathbb{E}}\langle G, x^{\otimes 3}\rangle = \tilde{\mathbb{E}} \sum_i x_i \cdot x^\top G_i x \le \left(\tilde{\mathbb{E}}\left\langle x^{\otimes 4}, \sum_i G_i \otimes G_i \right\rangle\right)^{1/2}$$

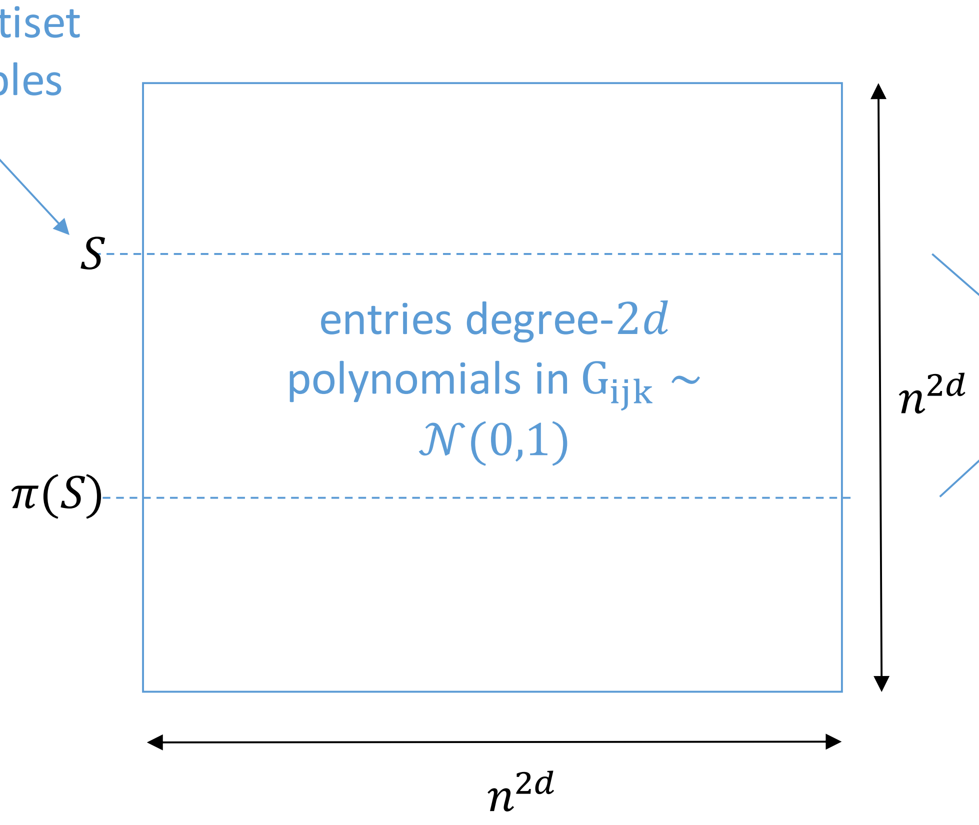
Jensen's inequality for d some power of 2

$$\le \left(\tilde{\mathbb{E}}\left\langle x^{\otimes 4d}, \left(\sum_i G_i \otimes G_i\right)^{\otimes d} \right\rangle\right)^{1/2d}$$

Symmetrize to improve eigenvalue

Taking the average of row S and $\pi(S)$ fixes the polynomial

$$\left(\sum_i G_i \otimes G_i\right)^{\otimes d} =$$

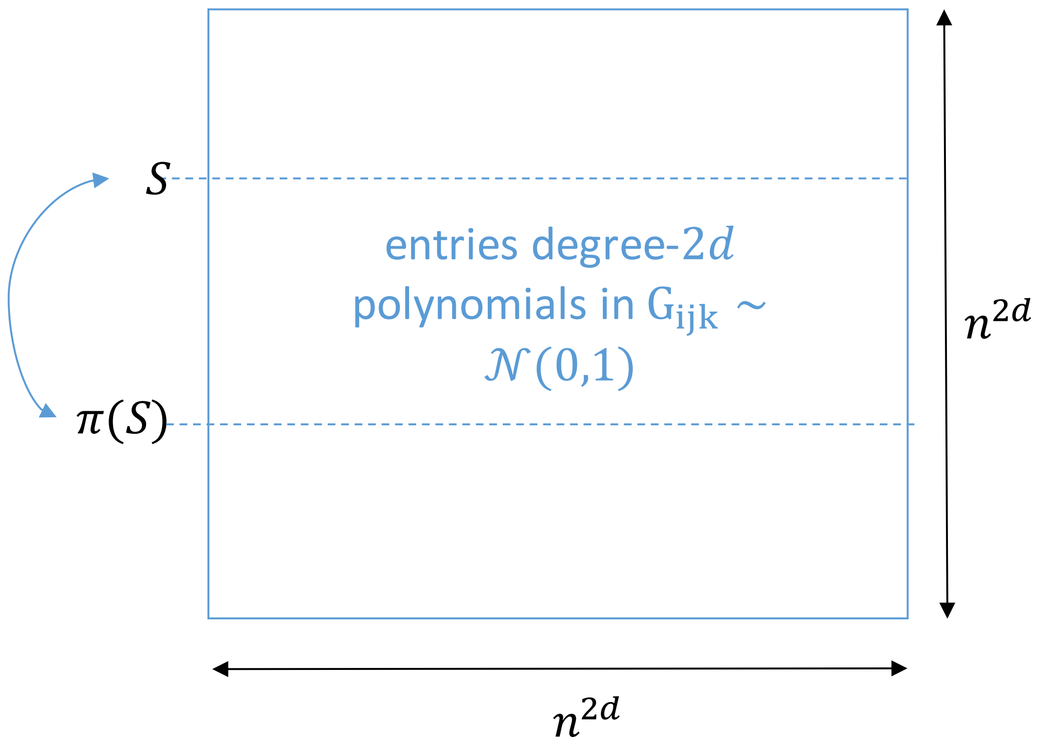


$$\prod_{i \in S} x_i = \prod_{j \in \pi(S)} x_j$$

Symmetrizing to improve eigenvalue

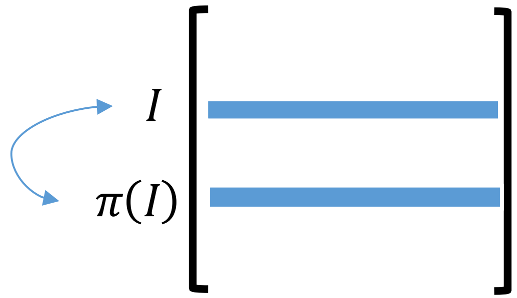
Taking the average of row S and $\pi(S)$ fixes the polynomial

$$M_{S,T}^{(d)} = \mathrm{Avg}_{\pi}\left[\left(\sum_i G_i \otimes G_i\right)^{\otimes d}_{\pi(S),T}\right]$$



$$M^{(2)}_{abcd,ijk\ell} := \frac{1}{4!}\Big((\textstyle\sum G_i \otimes G_i)_{abij}(\sum G_i \otimes G_i)_{cdk\ell} + (\sum G_i \otimes G_i)_{acij}(\sum G_i \otimes G_i)_{bdk\ell} + \cdots\Big)$$

Heuristic spectral norm calculation



Spectral norm?

$$M^{(d)} = \mathrm{Avg}_\pi\left[\left(\left(\sum G_i \otimes G_i\right)^{\otimes d}\right)_\pi\right]$$

$$\left\|\left(\sum G_i \otimes G_i\right)^{\otimes d}\right\| = \left(n^{3/2}\right)^d$$

each entry is average of $\sim d!$ “i.i.d. uniform” randomly signed variables

$\Longrightarrow$

avg. entry magnitude
$\mathrm{m} \to \approx \frac{m}{\sqrt{d!}}$

$\Longrightarrow$

$$\left\|M^{(d)}\right\| \leq \frac{n^{3d/2}}{d^{d/2}}$$

Improving Tensor PCA noise parameter

Theorem

$$\widetilde{\mathbb{E}}\langle \underbrace{G}_{\text{noise} \sim \mathcal{N}(0,1)}, \overset{\text{unit vector}}{x}{}^{\otimes 3}\rangle \le O\left(\frac{n^{3/4}}{D^{1/4}}\right)$$

$$\widetilde{\mathbb{E}}\langle G, x^{\otimes 3}\rangle \le \left(\widetilde{\mathbb{E}}\left\langle x^{\otimes 4}, \sum_i G_i \otimes G_i \right\rangle\right)^{1/2}$$

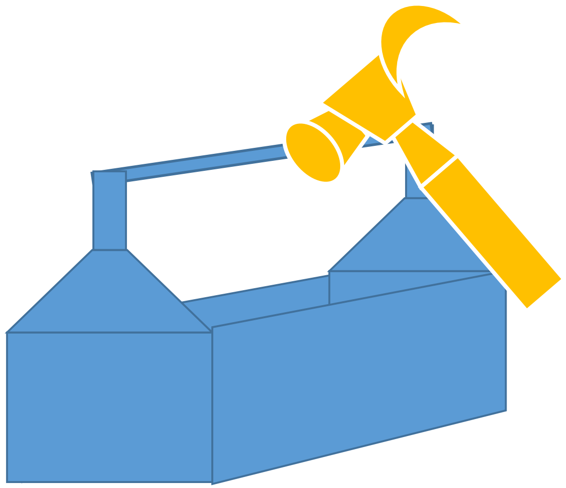
Jensen's inequality for d some power of 2 (if $D \ge 4d$)

$$\le \left(\widetilde{\mathbb{E}}\left\langle x^{\otimes 4d}, \left(\sum_i G_i \otimes G_i\right)^{\otimes d} \right\rangle\right)^{1/2d}$$

Average over symmetries of $x^{\otimes 2d}$ to reduce matrix representation eigenvalues

$$= \left(\widetilde{\mathbb{E}}\langle x^{\otimes 4d}, M^{(d)}\rangle\right)^{1/2d} \le \left\|M^{(d)}\right\|^{1/2d} \le \frac{n^{3/4}}{d^{1/4}} \quad \text{w.h.p.}$$

What kind of spectral algorithms?



Choose best matrix representation by:

- [x] Rearranging entries along "symmetries" of $x^{\otimes d}$

Applying degree-D SoS polynomial inequalities

- [x] Cauchy-Schwarz,
- [x] Jensen's Inequality (for squares), …

Other SoS (via Spectral) Algorithms

- Tensor Decomposition: symmetry, Cauchy-Schwarz, constant-d Jensen's
- Dictionary Learning: symmetry + tensor decomposition

  [Barak-Kelner-Steurer'14, Ge-Ma'15, Ma-Shi-Steurer'16]

- Planted Sparse Vector: symmetry

  [Barak-Brandão-Harrow-Kelner-Steurer-Zhou'12, Barak-Kelner-Steurer'14]

- Tensor Completion: symmetry, Cauchy-Schwarz [Barak-Moitra'16, Potechin-Steurer'17]

- Refuting Random CSPs: symmetry, Cauchy-Schwarz, Jensen's, $(x_i^2 = 1)$ constraints

  [Allen-O'Donnell-Witmer'15, Raghavendra-Rao-S'17]

- Polynomial Maximization over $\mathbb{S}_n$: symmetry, Cauchy-Schwarz, Jensen's, *worst case*

  [Bhattiprolu-Ghosh-Guruswami-E.Lee-Tulsiani'16]

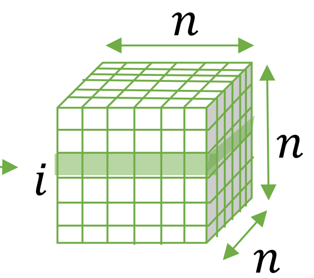
Fast spectral algorithms from SoS Analyses

[Hopkins-S-Shi-Steurer'16]

SoS Gives Spectral Search Algorithm

$$T = G + \lambda \cdot v^{\otimes 3}$$

T_i



$$T_i = G_i + \lambda v_i \cdot vv^\top$$

$$\sum_{i \in [n]} \boxed{T_i \otimes T_i}$$

n^2

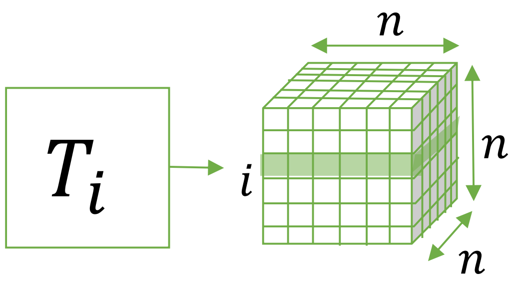
n^2

$\sum v_i^2 = 1$

$$= \underbrace{\sum G_i \otimes G_i}_{\text{eigenvalue} \leq n^{3/2}} + \text{cross-terms} + \underbrace{\sum \lambda^2 v_i^2 \cdot vv^\top \otimes vv^\top}_{\text{eigenvalue} = \lambda^2}$$

Running in $O(n^5)$....

$$T = G + \lambda \cdot v^{\otimes 3}$$



$$\sum_{i \in [n]} \underbrace{T_i \otimes T_i}_{n^2 \times n^2}$$

practical spectral algorithm?

sum of n matrices of size $n^2 \times n^2$ $\Rightarrow$ time = $n^5 + n^4 \log n$

build matrix

compute top eigenvalue

Theorem

Can compress to get an $O(n^3)$-time algorithm.

“Compressing” the matrix

Theorem

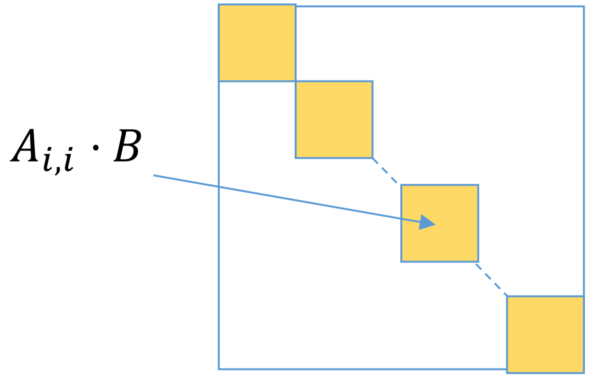
There is an $O(n^3)$-time algorithm.

$$\sum T_i \otimes T_i = \underbrace{\sum G_i \otimes G_i}_{\text{eigenvalue} \leq n^{3/2}} + \text{crossterms} + \underbrace{\lambda^2 \cdot vv^\top \otimes vv^\top}_{\text{eigenvalue} = \lambda^2}$$

How to reduce dimension but preserve signal-to-noise ratio?

A, B are $n \times n$ matrices

Partial Trace: $\mathrm{Tr}_{par}(A \otimes B) = \mathrm{Tr}(A) \cdot B$



$$\mathrm{Tr}_{par}(\lambda^2 vv^\top \otimes vv^\top) = \lambda^2 \|v\|^2 \cdot vv^\top = \lambda^2 \cdot vv^\top$$

$$\mathrm{Tr}_{par}(\sum G_i \otimes G_i) = \sum \underbrace{\mathrm{Tr}(G_i)}_{\approx \pm n^{1/2}} \cdot \underbrace{G_i}_{\text{eigenvalues} \approx \pm n^{1/2}} \rightarrow \text{eigs } n^{3/2}$$

signal-to-noise ratio preserved!

"Compressing" the matrix

Theorem

There is an $O(n^3)$-time algorithm.

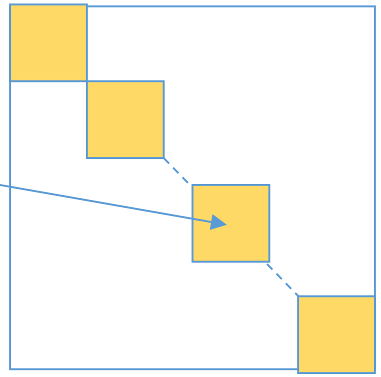
$$\underbrace{\sum T_i \otimes T_i = \sum G_i \otimes G_i}_{\text{eigenvalue} \leq n^{3/2}} + \text{crossterms} + \underbrace{\lambda^2 \cdot vv^\top \otimes vv^\top}_{\text{eigenvalue} = \lambda^2}$$

How to reduce dimension but preserve signal-to-noise ratio?

A, B are $n \times n$ matrices

Partial Trace: $\mathrm{Tr}_{par}(A \otimes B) = \mathrm{Tr}(A) \cdot B$

$A_{i,i} \cdot B$



$$\mathrm{Tr}_{par}(\sum T_i \otimes T_i) = \sum \mathrm{Tr}(T_i) \cdot T_i$$

runtime?

computing all $\mathrm{Tr}(T_i)$: n^2 time

each of the n^2 entries is sum of n numbers: n^3 time

linear in input!

computing top eigenvector/eigenvalue of $n \times n$ matrix: $n^2 \log n$ time

Fast Spectral Algorithms via SoS

Secret Sauce: apply partial trace to SoS matrix (in a way that enables fast power iteration)

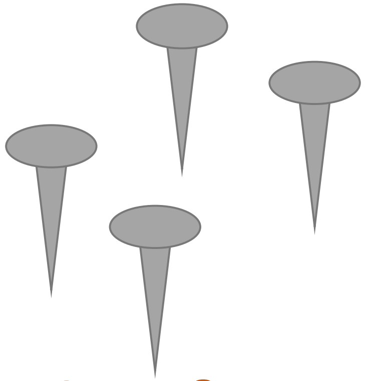
- Tensor PCA [Hopkins-S-Shi-Steurer'16]
- Tensor decomposition [Hopkins-S-Shi-Steurer'16, S-Steurer'17]
- Planted Sparse Vector [Hopkins-S-Shi-Steurer'16]
- Tensor Completion [Montanari-Sun'17]

SoS perspective gives new spectral algorithms

Spectral techniques let us analyze SoS

Worst-case problems?

Spectral Algorithms



Average-Case &
Structured Instances

Sum-of-Squares Algorithms