Lower Bounds Against the Sum of Squares Algorithm

Sam Hopkins
Simons + UC Berkeley

Agenda

1. What is SoS?
2. SoS, learning, and crypto
3. Lower bounds: what we know
4. Lower bounds: what we hope for

What is SoS?

Alg. for deciding:

Does $p_1(x_1 \ldots x_n) \geq 0, \ldots, p_m(x_1 \ldots x_n) \geq 0$

have a solution in $\mathbb{R}^n$?

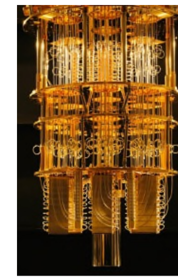
[Parrilo '01, Lasserre '01]

Why Polynomial Systems?

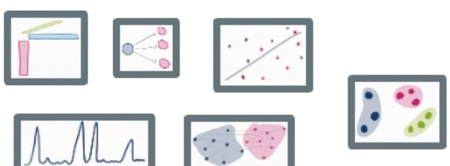
Polynomials are highly expressive

Combinatorial optimization

robotics / optimal control



quantum info.



Statistics/learning/ avg. case algos

Crypto

What is SoS?

Degree-d __SoS proof__ that

$p_1(x) \geq 0 \ldots p_m(x) \geq 0$: $\{q_S = \sum q_{S,i}(x)^2\}_{S \subseteq [m]}$

$$-1 = \underbrace{\sum_{S \subseteq [m]} q_S(x) \cdot \prod_{j \in S} p_j(x)}_{\text{degree} \leq d}$$

What is SoS?

- Hilbert, ..., Krivine-Stengle: every unsatisfiable $p_1 \geq 0 \ldots p_m \geq 0$ has an SoS proof (!)
- Parrilo-Lasserre: can decide existence of degree d proof in $(n \cdot m)^{O(d)}$ time

What is SoS?

Proving power vs. running time

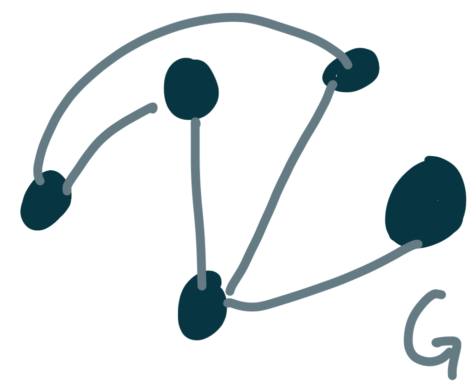
<u>tradeoff</u>

Seems to capture best known algs for many problems (sometimes provably optimal)

<u>Ex:</u> min-cut$(G) = c$

$\Updownarrow$

degree $O(1)$ SoS proof that no cut $< c$.



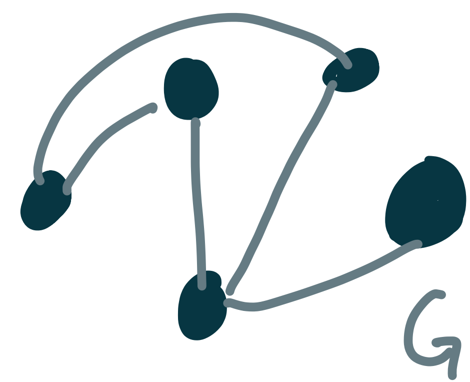
Seems to capture best known algs for many problems (sometimes provably optimal)

<u>Ex:</u> $\text{max-cut}(G) = c$

$\Downarrow$

degree-2 SoS proof that no cut $> \frac{c}{0.878}$

[Goemans - Williamson '95]



Seems to capture best known algs for many problems (sometimes provably optimal)

Ex: max-cut $(G) = c$

$\Downarrow$

degree-2 SoS proof that no cut $> \frac{c}{0.878}$

[Goemans - Williamson '95]

optimal under UGC

generalizes to all CSPs!

[KKMO, Raghavendra]

SoS + Avg. Case / Stat. inference

SoS gives best known algorithms for canonical avg. case problems

- planted clique
- refuting random CSPs
- ind. set in $G(n,p)$
- etc.

SoS + Avg. Case / Stat. inference

SoS gives best known algorithms for canonical avg. case problems

- planted clique
- refuting random CSPs
- ind. set in $G(n,p)$
- etc.

Aside: recent revolution in SoS for high-dimensional Statistics:

- robust stats
- latent variable models
- tensor problems, & more

Example: Planted Clique

Goal: distinguish $G(n, 1/2)$ from $G(n, \frac{1}{2})$
+ k-clique

$k \geq \Omega(\sqrt{n})$: poly-time algs known [AKS '97]

$k \ll \sqrt{n}$: conjecturally hard

Many consequences, incl. private-key crypto

[Juels-Peinado '00]

SoS to detect a k-clique:

Given G, check if $\exists$ SoS proof of:

"all cliques in G have $< k$ vertices"

vars $x_1 \dots x_n$

$x_i^2 = x_i$

$x_i x_j = 0 \quad \forall i \nsim j$

$\sum x_i = k$

look for SoS proof of unsatisfiability

Fact: degree-d proof exists for $k \approx \frac{\sqrt{n}}{2^d}$ (w.h.p.)

Proof for $d=2$:

$x_i^2 = x_i$
$x_i x_j = 0 \quad \forall i \not\sim j$
$\sum x_i = k$

① $k^2 = (\sum x_i)^2 = \sum x_i x_j$

centered adj. matrix

$= \sum_{i \sim j} x_i x_j - \sum_{i \not\sim j} x_i x_j + \sum_{i \not\sim j} x_i x_j = x^T A x + \sum_{i \not\sim j} x_i x_j$

② $x^T A x = \underbrace{\sqrt{n}\|x\|^2 - \sum \langle x, a_i \rangle^2}_{\text{eigenvalue bound on } A} = \sqrt{n}k - \sum \langle x, a_i \rangle^2 - \sqrt{n} \sum x_i - x_i^2$
$- \sqrt{n}(k - \sum x_i)$

①+②: $\sqrt{n}k - k^2 = \sum \langle x, a_i \rangle^2 + \text{pink stuff}$ ∎

SoS + Crypto

Before Jain-Lin-Sahai iO construction, many iO candidates: Structured PRGs $\Rightarrow$ iO

Lin-Tessaro,
Ananth-Jain-Sahai
Lin-Matt,
Agrawal

BUT: Structured PRG candidates REPEATEDLY BROKEN

[Barak-Brakerski-Komargodski-Kothari '17, Barak-H.-Jain-Kothari-Sahai '18]

In light of algorithmic power of SoS why lower bounds?

narrow view: good to know when SoS algs attacks don't work

broad view: lens on hardness, esp for avg. case.

"SoS is optimal" (??)

hardness in stats

Source for crypto?

Agenda

1. ~~What is SoS?~~
2. ~~SoS, learning, and crypto~~
3. Lower bounds: what we know
4. Lower bounds: what we hope for

Today: focus on average case

- CSPs
- Planted Clique
- Boolean vec. in random subspace
- Ind. set in sparse graphs
- Sparse PCA
- Tensor PCA

[KMOW16, BHKKMP16, GJJPR20, JPRTX21, HKPRSS'17, P'21]

Ex. Planted Clique

Thm. [BHKKMP + Peng]: There is no degree $d = o(\log n)$ SoS refutation of $x_i^2 = x_i$, $x_i x_j = 0$ for $i \not\sim_G j$, $\sum x_i = n^{\frac{1}{2} - o(1)}$ w.h.p for $G \sim G(n, 1/2)$.

Beyond poly-time →

Thm. [BHKKMP+ Peng]: There is no degree $d = o(\log n)$ SoS refutation of $x_i^2 = x_i$, $x_i x_j = 0$ for $i \sim_G j$, $\sum x_i = n^{\frac{1}{2} - o(1)}$ w.h.p for $G \sim G(n, 1/2)$.

Suggests refuting is hard (for SoS?) in poly time.

Often interpreted to mean distinguishing hard

How to prove an SoS lower bd.

$$-1 = \sum_{S \subseteq [m]} q_S(x) \prod_{j \in S} p_j(x)$$

$\sum x_i = n^{\frac{1}{2} - o(1)}$

$x_i^2 = x_i$,

$x_i x_j = 0$ for $i \sim_G j$,

Necessary & Sufficient: Pseudoexpectation

$\tilde{\mathbb{E}} : \mathbb{R}[x]_{\leq d} \to \mathbb{R}$

$\tilde{\mathbb{E}}\, 1 = 1$

$\tilde{\mathbb{E}}\, p^2 \geq 0$

$\tilde{\mathbb{E}}\, p(x)\,(x_i^2 - x_i) = 0, \forall p$

$\tilde{\mathbb{E}}\, p(x)\left(\sum x_i - k\right) = 0, \forall p$

$\tilde{\mathbb{E}}\, p(x)\, x_i x_j = 0, \forall p$

Need map $G \to \tilde{\mathbb{E}}_G$

Pseudocalibration: Construction of such a map from hard distinguishing problem e.g. $G(n, 1/2)$ vs $G(n, 1/2)$ + k-clique

H_0 $\quad$ H_1

Idea: low-degree moments of H_1 define $\tilde{\mathbb{E}}$

All known avg. case SoS l.b.'s can (or should be) proved via pseudocalibration

- CSPs
- Planted Clique
- Boolean vec. in random subspace
- Ind. set in sparse graphs
- Sparse PCA
- Tensor PCA

Analysis remains case-by-case:

- random matrices
- random graphs, expansion

All known avg. case SoS l.b.'s can (or should be) proved via pseudocalibration

- CSPs
- Planted clique
- Boolean vec. in random subspace
- Ind. set in sparse graphs
- Sparse PCA
- Tensor PCA

<u>Analysis</u> remains case-by-case:

- random matrices
- random graphs, expansion

numerous open problems

- understand random matrices w/ dependent entries

3. Lower bounds: what we know

- a bunch of l.b.'s for avg-case refutation problems, with matching algos
- a canonical technique which we still struggle to analyze

Agenda

1. ~~What is SoS?~~
2. ~~SoS, learning, and crypto~~
3. ~~Lower bounds: what we know~~
4. Lower bounds: what we hope for

Things we don't understand

Refutation vs Distinguishing

For stats & crypto, usually want hardness for distinguishing 2 distn's

SoS natively does <u>refutation</u>

How do we interpret SoS l.b.'s for distinguishing?

Ex. Planted Clique

Could try to refute

$x_i^2 = x_i$, $x_i x_j = 0$ for $i \nsim j$, $\sum x_i = k$,

and $p(G, x) \geq 0$

for any low-degree p s.t. $p(G, x) \geq 0$

whp for $G, x \sim G(n, 1/2) + k$-clique

Ex. Distinguishing random CSPs from CSPs w/ planted solutions

- Usual refutation problem: for random CSP, refute existence of satisfying assignment
- Kothari-O'Donnell-Schramm: balanced assign.?

$\sum x_i = 0$

A very partial resolution:

refutation lower bounds proved using pseudocalibration "should not" have this problem

$\tilde{\mathbb{E}}$ from pseudocalibration should rule out SoS for all low-degree refutation problems derived from a distinguishing problem

The Role of Noise

(More) serious roadblock to using SoS to understand complexity of avg-case problems:

There are SoS lower bounds for easy problems

!!

Ex. 3XOR is easy by Gaussian elimination

BUT: degree $\Omega(n)$ SoS cannot refute random 3XOR!

SoS is inherently noise-robust

$\Rightarrow$ cannot accurately capture the complexity of "non-robust" problems

Can hope to use SoS lower bounds to reason about complexity of distinguishing problems when:

- all corresponding refutation problems have suitable SoS l.b.'s
- everything is a little noisy

4. Lower bounds: what we hope for

Low-Degree Conjecture: H_0 vs H_1 on $\{\pm 1\}^{\binom{n}{k}}$

H_0 product, H_1 S_n-Symmetric.

Not distinguishable via $\omega(\log n)$ moments

$$\Downarrow$$

H_0 vs $\underbrace{\tilde{H}_1}_{\text{Noisy } H_1}$ not distinguishable in poly time

4. Lower bounds: what we hope for

SoS

Low-Degree Conjecture: H_0 vs H_1 on $\{\pm 1\}^{\binom{n}{k}}$

H_0 product, H_1 S_n-symmetric.

Not distinguishable via $\omega(\log n)$ moments

$\Downarrow$

H_0 vs $\tilde{H}_1$ not distinguishable by $O(1)$-deg.

SoS

A few open problems:

- SoS low degree conjecture

  Special case: "Kesten-Stigum" threshold for sparse graph models

- Non-product distributions

  important for crypto applications

- Random matrices with dependent entries