The Power of Low-Degree Polynomials for Solving Statistical Problems

Alex Wein

Courant Institute, New York University

Based on joint work with:

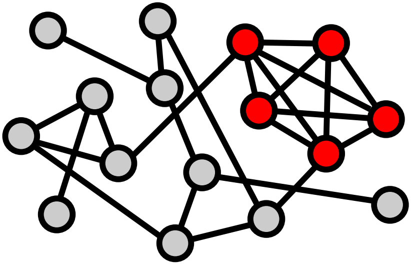
David Gamarnik (MIT)

Aukosh Jagannath (Waterloo)

Tselil Schramm (Stanford)

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph



Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique
- **Recovery**: given a graph with a planted clique, find the clique

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique
- **Recovery**: given a graph with a planted clique, find the clique
- **Optimization**: given a random graph (with no planted clique), find as large a clique as possible

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique
- **Recovery**: given a graph with a planted clique, find the clique
- **Optimization**: given a random graph (with no planted clique), find as large a clique as possible

Common to have information-computation gaps

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique
- **Recovery**: given a graph with a planted clique, find the clique
- **Optimization**: given a random graph (with no planted clique), find as large a clique as possible

Common to have information-computation gaps

E.g. planted k-clique (either detection or recovery)

Problems in High-Dimensional Statistics

Example: finding a large clique in a random graph

- **Detection**: distinguish between a random graph and a graph with a planted clique
- **Recovery**: given a graph with a planted clique, find the clique
- **Optimization**: given a random graph (with no planted clique), find as large a clique as possible

Common to have information-computation gaps

E.g. planted k-clique (either detection or recovery)

What makes problems easy vs hard?

The Low-Degree Polynomial Method

A framework for understanding computational complexity

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

[Barak, Hopkins, Kelner, Kothari, Moitra, Potechin '16]

[Hopkins, Steurer '17]

[Hopkins, Kothari, Potechin, Raghavendra, Schramm, Steurer '17]

[Hopkins '18 (PhD thesis)]

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$
- "Low" means $O(\log n)$ where n is dimension

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$
- "Low" means $O(\log n)$ where n is dimension

Some low-degree algorithms:

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$
- "Low" means $O(\log n)$ where n is dimension

Some low-degree algorithms:

- Spectral methods (power iteration)

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$
- "Low" means $O(\log n)$ where n is dimension

Some low-degree algorithms:

- Spectral methods (power iteration)
- Approximate message passing (AMP) [DMM09]

The Low-Degree Polynomial Method

A framework for understanding computational complexity

Originated from sum-of-squares literature (for detection)

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^N \to \mathbb{R}^M$
- "Low" means $O(\log n)$ where n is dimension

Some low-degree algorithms:

- Spectral methods (power iteration)
- Approximate message passing (AMP) [DMM09]

Low-degree algorithms are as powerful as the best known polynomial-time algorithms for many problems: planted clique, sparse PCA, community detection, tensor PCA, constraint satisfaction, spiked matrix [BHKKMP16,HS17,HKPRSS17,Hop18,BKW19,KWB19,DKWB19]

Overview

This talk: techniques to prove that **all** low-degree polynomials fail

Overview

This talk: techniques to prove that **all** low-degree polynomials fail

Constitutes evidence for computational hardness

Overview

This talk: techniques to prove that **all** low-degree polynomials fail

Constitutes evidence for computational hardness

Settings:

- Detection (prior work)
- Recovery
  Schramm, W. "Computational Barriers to Estimation from Low-Degree Polynomials", arXiv, 2020.
- Optimization
  Gamarnik, Jagannath, W. "Low-Degree Hardness of Random Optimization Problems", FOCS 2020.

Detection (e.g. [Hopkins, Steurer '17])

Goal: hypothesis test with error probability $o(1)$ between:

- Null model $Y \sim \mathbb{Q}_n$ e.g. $G(n, 1/2)$
- Planted model $Y \sim \mathbb{P}_n$ e.g. $G(n, 1/2) \cup \{\text{random } k\text{-clique}\}$

Detection (e.g. [Hopkins, Steurer '17])

Goal: hypothesis test with error probability $o(1)$ between:

- Null model $Y \sim \mathbb{Q}_n$ e.g. $G(n, 1/2)$
- Planted model $Y \sim \mathbb{P}_n$ e.g. $G(n, 1/2) \cup \{\text{random } k\text{-clique}\}$

Look for a degree-D (multivariate) polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ that distinguishes $\mathbb{P}$ from $\mathbb{Q}$

Detection (e.g. [Hopkins, Steurer '17])

Goal: hypothesis test with error probability $o(1)$ between:

- Null model $Y \sim \mathbb{Q}_n$ e.g. $G(n, 1/2)$
- Planted model $Y \sim \mathbb{P}_n$ e.g. $G(n, 1/2) \cup \{\text{random } k\text{-clique}\}$

Look for a degree-D (multivariate) polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ that distinguishes $\mathbb{P}$ from $\mathbb{Q}$

- In the sense that $f(Y)$ is "big" when $Y \sim \mathbb{P}$ and "small" when $Y \sim \mathbb{Q}$

Detection (e.g. [Hopkins, Steurer '17])

Goal: hypothesis test with error probability $o(1)$ between:

- Null model $Y \sim \mathbb{Q}_n$ e.g. $G(n, 1/2)$
- Planted model $Y \sim \mathbb{P}_n$ e.g. $G(n, 1/2) \cup \{\text{random } k\text{-clique}\}$

Look for a degree-D (multivariate) polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ that distinguishes $\mathbb{P}$ from $\mathbb{Q}$

- In the sense that $f(Y)$ is "big" when $Y \sim \mathbb{P}$ and "small" when $Y \sim \mathbb{Q}$

Compute $$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y \sim \mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y \sim \mathbb{Q}}[f(Y)^2]}} \qquad \frac{\text{mean in } \mathbb{P}}{\text{fluctuations in } \mathbb{Q}}$$

Detection (e.g. [Hopkins, Steurer '17])

Goal: hypothesis test with error probability $o(1)$ between:

- Null model $Y \sim \mathbb{Q}_n$ e.g. $G(n, 1/2)$
- Planted model $Y \sim \mathbb{P}_n$ e.g. $G(n, 1/2) \cup \{\text{random } k\text{-clique}\}$

Look for a degree-D (multivariate) polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ that distinguishes $\mathbb{P}$ from $\mathbb{Q}$

- In the sense that $f(Y)$ is "big" when $Y \sim \mathbb{P}$ and "small" when $Y \sim \mathbb{Q}$

Compute $$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y \sim \mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y \sim \mathbb{Q}}[f(Y)^2]}} \qquad \frac{\text{mean in } \mathbb{P}}{\text{fluctuations in } \mathbb{Q}}$$

$$= \begin{cases} \omega(1) & \text{degree-}D \text{ polynomial succeed} \\ O(1) & \text{degree-}D \text{ polynomials fail} \end{cases}$$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}}$$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}}$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}}$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\langle L, f\rangle}{\|f\|}$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y \sim \mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y \sim \mathbb{Q}}[f(Y)^2]}}$$

$$= \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y \sim \mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y \sim \mathbb{Q}}[f(Y)^2]}}$$

$$= \max_{f \text{ deg } D} \frac{\langle L, f \rangle}{\|f\|}$$

$$\langle f, g \rangle = \mathbb{E}_{Y \sim \mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f \rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Maximizer: $f = L^{\leq D}$:= projection of L onto degree-D subspace

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}}$$
$$= \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}}$$
$$= \max_{f \text{ deg } D} \frac{\langle L, f\rangle}{\|f\|}$$
$$= \|L^{\leq D}\|$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Maximizer: $f = L^{\leq D} :=$ projection of L onto degree-D subspace

Norm of low-degree likelihood ratio

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\langle L, f\rangle}{\|f\|} = \|L^{\leq D}\|$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Maximizer: $f = L^{\leq D} :=$ projection of L onto degree-D subspace

Norm of low-degree likelihood ratio

To project: expand L in orthogonal polynomials w.r.t. $\mathbb{Q}$

Detection (e.g. [Hopkins, Steurer '17])

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{P}}[f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\mathbb{E}_{Y\sim\mathbb{Q}}[L(Y)f(Y)]}{\sqrt{\mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)^2]}} = \max_{f \text{ deg } D} \frac{\langle L, f\rangle}{\|f\|} = \|L^{\leq D}\|$$

$$\langle f, g\rangle = \mathbb{E}_{Y\sim\mathbb{Q}}[f(Y)g(Y)]$$

$$\|f\| = \sqrt{\langle f, f\rangle}$$

Likelihood ratio:
$L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$

Maximizer: $f = L^{\leq D} :=$ projection of L onto degree-D subspace

Norm of low-degree likelihood ratio

To project: expand L in orthogonal polynomials w.r.t. $\mathbb{Q}$

- Works if $\mathbb{Q}$ has independent entries

Recovery [Schramm, W. '20]

Example (planted submatrix): observe $n \times n$ matrix $Y = X + Z$

- Signal: $X = \lambda v v^\top$ where $\lambda > 0$ and $v_i \sim \mathrm{Bernoulli}(\rho)$
- Noise: Z i.i.d. $\mathcal{N}(0,1)$

Recovery [Schramm, W. '20]

Example (planted submatrix): observe $n \times n$ matrix $Y = X + Z$

- Signal: $X = \lambda v v^\top$ where $\lambda > 0$ and $v_i \sim \mathrm{Bernoulli}(\rho)$
- Noise: Z i.i.d. $\mathcal{N}(0, 1)$

Goal: given Y, estimate v_1 via polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$

Recovery [Schramm, W. '20]

Example (planted submatrix): observe $n \times n$ matrix $Y = X + Z$

- Signal: $X = \lambda v v^\top$ where $\lambda > 0$ and $v_i \sim \mathrm{Bernoulli}(\rho)$
- Noise: Z i.i.d. $\mathcal{N}(0,1)$

Goal: given Y, estimate v_1 via polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$

Low-degree minimum mean squared error:

$$\mathrm{MMSE}_{\leq D} = \min_{f \text{ deg } D} \mathbb{E}(f(Y) - v_1)^2$$

Recovery [Schramm, W. '20]

Example (planted submatrix): observe $n \times n$ matrix $Y = X + Z$

- Signal: $X = \lambda v v^\top$ where $\lambda > 0$ and $v_i \sim \mathrm{Bernoulli}(\rho)$
- Noise: Z i.i.d. $\mathcal{N}(0,1)$

Goal: given Y, estimate v_1 via polynomial $f : \mathbb{R}^{n \times n} \to \mathbb{R}$

Low-degree minimum mean squared error:

$$\mathrm{MMSE}_{\leq D} = \min_{f \text{ deg } D} \mathbb{E}(f(Y) - v_1)^2$$

Equivalent to low-degree maximum correlation:

$$\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$$

Fact: $\mathrm{MMSE}_{\leq D} = \mathbb{E}[v_1^2] - \mathrm{Corr}_{\leq D}^2$

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\text{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Same proof as detection?

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Same proof as detection?

Issue: would need orthogonal polynomials for planted distribution

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Same proof as detection?

Issue: would need orthogonal polynomials for planted distribution

Trick: bound denominator via Jensen's inequality

$$\mathbb{E}[f(Y)^2] = \mathbb{E}_Z \mathbb{E}_X [f(X+Z)^2] \geq \mathbb{E}_Z \left(\mathbb{E}_X f(X+Z)\right)^2$$

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Same proof as detection?

Issue: would need orthogonal polynomials for planted distribution

Trick: bound denominator via Jensen's inequality

$$\mathbb{E}[f(Y)^2] = \mathbb{E}_Z \mathbb{E}_X [f(X+Z)^2] \geq \mathbb{E}_Z \left(\mathbb{E}_X f(X+Z)\right)^2$$

- This simplifies expression enough to find closed form

Recovery [Schramm, W. '20]

For hardness, want upper bound on $\mathrm{Corr}_{\leq D} = \max_{f \text{ deg } D} \frac{\mathbb{E}[f(Y) \cdot v_1]}{\sqrt{\mathbb{E}[f(Y)^2]}}$

Same proof as detection?

Issue: would need orthogonal polynomials for planted distribution

Trick: bound denominator via Jensen's inequality

$$\mathbb{E}[f(Y)^2] = \mathbb{E}_Z \mathbb{E}_X [f(X+Z)^2] \geq \mathbb{E}_Z \left(\mathbb{E}_X f(X+Z) \right)^2$$

- This simplifies expression enough to find closed form
- Yields tight bounds for planted submatrix problem

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}} \langle Y, v^{\otimes 3} \rangle$

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}} \langle Y, v^{\otimes 3} \rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}} \langle Y, v^{\otimes 3} \rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Our result: no constant-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}}\langle Y, v^{\otimes 3}\rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Our result: no constant-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

Theorem (GJW'20)

For some $\epsilon > 0$, no degree-$O(1)$ polynomial $f : \mathbb{R}^{n \times n \times n} \to \mathbb{R}^n$ achieves both of the following with probability $1 - o(1)$:

- *Objective:* $H(f(Y)) \geq \mathrm{OPT} - \epsilon$
- *Normalization:* $\|f(Y)\| \approx 1$

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}} \langle Y, v^{\otimes 3} \rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Our result: no constant-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

- Best known algorithms are constant-degree [Sub18,Mon18,EMS20]

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}}\langle Y, v^{\otimes 3}\rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Our result: no constant-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

- Best known algorithms are constant-degree [Sub18,Mon18,EMS20]
- Proof:
  - Low-degree polynomials are stable
  - Overlap gap property [GS13,CGPR17,GJ19]

Optimization [Gamarnik, Jagannath, W. '20]

Example (spherical spin glass): for $Y \in \mathbb{R}^{n \times n \times n}$ i.i.d. $\mathcal{N}(0,1)$, find unit vector v maximizing $H(v) = \frac{1}{\sqrt{n}} \langle Y, v^{\otimes 3} \rangle$

Optimum value: $\mathrm{OPT} = \max_{\|v\|=1} H(v) = \Theta(1)$

Our result: no constant-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

- Best known algorithms are constant-degree [Sub18,Mon18,EMS20]
- Proof:
  - Low-degree polynomials are stable
  - Overlap gap property [GS13,CGPR17,GJ19]
- Open: show that no low-degree polynomial can achieve the precise objective value achieved by [Sub18]

References

- Detection (survey article)
  Kunisky, W., Bandeira. "Notes on Computational Hardness of Hypothesis Testing: Predictions using the Low-Degree Likelihood Ratio", arXiv:1907.11636

- Recovery
  Schramm, W. "Computational Barriers to Estimation from Low-Degree Polynomials", arXiv:2008.02269

- Optimization
  Gamarnik, Jagannath, W. "Low-Degree Hardness of Random Optimization Problems", arXiv:2004.12063