Low-Degree Hardness of Random Optimization Problems

Alex Wein
Courant Institute, New York University

Joint work with:

David Gamarnik
MIT

Aukosh Jagannath
Waterloo

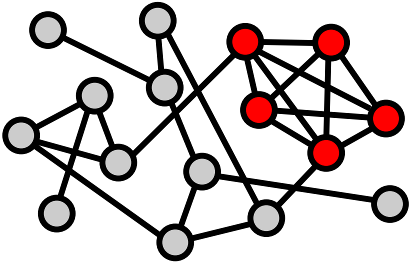
Random Optimization Problems

Examples:

Random Optimization Problems

Examples:

- Max clique in a random graph



Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Note: no planted solution

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Note: no planted solution

Q: What is the typical value of the optimum (OPT)?

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Note: no planted solution

Q: What is the typical value of the optimum (OPT)?

Q: What objective value can be reached algorithmically (ALG)?

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Note: no planted solution

Q: What is the typical value of the optimum (OPT)?

Q: What objective value can be reached algorithmically (ALG)?

Q: In cases where it seems hard to reach a particular objective value, can we understand why?

Random Optimization Problems

Examples:

- Max clique in a random graph
- Max-k-SAT on a random formula
- Maximizing a random degree-p polynomial over the sphere

Note: no planted solution

Q: What is the typical value of the optimum (OPT)?

Q: What objective value can be reached algorithmically (ALG)?

Q: In cases where it seems hard to reach a particular objective value, can we understand why? In a unified way?

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n$$

[Frieze '90]

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

[Frieze '90]

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

[Frieze '90]

[Karp '76]: Is there a better algorithm?

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad \mathrm{ALG} = \frac{\log d}{d}\,n$$

[Frieze '90]

[Karp '76]: Is there a better algorithm?

Structural evidence suggests no!

[Achlioptas, Coja-Oghlan '08; Coja-Oghlan, Efthymiou '10]

Max Independent Set

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

[Frieze '90]

[Karp '76]: Is there a better algorithm?

Structural evidence suggests no!

[Achlioptas, Coja-Oghlan '08; Coja-Oghlan, Efthymiou '10]

Local algorithms achieve value ALG **and no better**

[Gamarnik, Sudan '13; Rahman, Virág '14]

Spherical Spin Glass

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

Spherical Spin Glass

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathrm{OPT} = \Theta(1)$ [Auffinger, Ben Arous, Černý '13]

Spherical Spin Glass

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathrm{OPT} = \Theta(1)$ [Auffinger, Ben Arous, Černý '13]

$\mathrm{ALG} = \Theta(1)$ [Subag '18]

Spherical Spin Glass

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\text{OPT} = \Theta(1)$ [Auffinger, Ben Arous, Černý '13]

$\text{ALG} = \Theta(1)$ [Subag '18]

$\text{ALG} < \text{OPT}$ (for $p \geq 3$)

Spherical Spin Glass

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathrm{OPT} = \Theta(1)$ [Auffinger, Ben Arous, Černý '13]

$\mathrm{ALG} = \Theta(1)$ [Subag '18]

$\mathrm{ALG} < \mathrm{OPT}$ (for $p \geq 3$)

Approximate message passing (AMP) algorithms achieve value ALG **and no better** [El Alaoui, Montanari, Sellke '20]

What's Missing?

How to give the best "evidence" that there are no better algorithms?

What's Missing?

How to give the best “evidence” that there are no better algorithms?

Prior work rules out certain classes of algorithms (local, AMP), but do we expect these to be optimal?

What's Missing?

How to give the best "evidence" that there are no better algorithms?

Prior work rules out certain classes of algorithms (local, AMP), but do we expect these to be optimal?

- AMP is not optimal for tensor PCA [Montanari, Richard '14]

What's Missing?

How to give the best "evidence" that there are no better algorithms?

Prior work rules out certain classes of algorithms (local, AMP), but do we expect these to be optimal?

- AMP is not optimal for tensor PCA [Montanari, Richard '14]

Would like a unified framework for lower bounds

What's Missing?

How to give the best "evidence" that there are no better algorithms?

Prior work rules out certain classes of algorithms (local, AMP), but do we expect these to be optimal?

- AMP is not optimal for tensor PCA [Montanari, Richard '14]

Would like a unified framework for lower bounds

- Local algorithms only make sense on sparse graphs

What's Missing?

How to give the best "evidence" that there are no better algorithms?

Prior work rules out certain classes of algorithms (local, AMP), but do we expect these to be optimal?

- AMP is not optimal for tensor PCA [Montanari, Richard '14]

Would like a unified framework for lower bounds

- Local algorithms only make sense on sparse graphs

Solution: lower bounds against a larger class of algorithms (low-degree polynomials) that contains both local and AMP algorithms

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

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

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms:

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms: input $Y \in \mathbb{R}^{n \times n}$

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms: input $Y \in \mathbb{R}^{n \times n}$

- Power iteration: $Y^k \mathbf{1}$

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms: input $Y \in \mathbb{R}^{n \times n}$

- Power iteration: $Y^k \mathbf{1}$
- Approximate message passing

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms: input $Y \in \mathbb{R}^{n \times n}$

- Power iteration: $Y^k \mathbf{1}$
- Approximate message passing
- Local algorithms on sparse graphs

The Low-Degree Polynomial Framework

Study a restricted class of algorithms: low-degree polynomials

- Multivariate polynomial $f : \mathbb{R}^M \to \mathbb{R}^N$
  - Input: e.g. graph $Y \in \{0,1\}^{\binom{n}{2}}$
  - Output: e.g. $b \in \{0,1\}$ or $v \in \mathbb{R}^n$
- "Low" means $O(\log n)$ where n is dimension

Examples of low-degree algorithms: input $Y \in \mathbb{R}^{n \times n}$

- Power iteration: $Y^k \mathbf{1}$
- Approximate message passing
- Local algorithms on sparse graphs
- Or any of the above applied to $\tilde{Y} = g(Y)$

Planted Problems

For problems with a planted signal, the low-degree framework is already well-established

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

[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

Planted Problems

For problems with a planted signal, the low-degree framework is already well-established

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

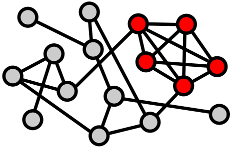
[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

Example (planted clique): $G(n, 1/2)$ with planted k-clique

- Detection
- Recovery



Planted Problems (Continued)

For all of these planted problems...

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

- the best known poly-time algorithms are captured by $O(\log n)$-degree polynomials (spectral/AMP)

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

- the best known poly-time algorithms are captured by $O(\log n)$-degree polynomials (spectral/AMP)
- low-degree polynomials fail in the "hard" regime
  [BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19]

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

- the best known poly-time algorithms are captured by $O(\log n)$-degree polynomials (spectral/AMP)
- low-degree polynomials fail in the "hard" regime
  [BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19]

<u>This work</u>: extend low-degree framework to non-planted setting

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

- the best known poly-time algorithms are captured by $O(\log n)$-degree polynomials (spectral/AMP)
- low-degree polynomials fail in the "hard" regime
  [BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19]

This work: extend low-degree framework to non-planted setting

Other frameworks: sum-of-squares, statistical query model

Planted Problems (Continued)

For all of these planted problems...

planted clique, sparse PCA, community detection, tensor PCA, spiked Wigner/Wishart, planted submatrix, planted dense subgraph, ...

...it is the case that

- the best known poly-time algorithms are captured by $O(\log n)$-degree polynomials (spectral/AMP)
- low-degree polynomials fail in the "hard" regime
  [BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19]

This work: extend low-degree framework to non-planted setting

Other frameworks: sum-of-squares, statistical query model

"Robustness": Gaussian elimination for XOR-SAT

Spherical Spin Glass: Results

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathrm{ALG} < \mathrm{OPT}$ (for $p \geq 3$)

Spherical Spin Glass: Results

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathrm{ALG} < \mathrm{OPT}$ (for $p \geq 3$)

Result: no low-degree polynomial can achieve value $\mathrm{OPT} - \epsilon$

Spherical Spin Glass: Results

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

ALG $<$ OPT (for $p \geq 3$)

Result: no low-degree polynomial can achieve value OPT $- \epsilon$

Theorem [Gamarnik, Jagannath, W. '20] Let $p \geq 4$ be even.
For some $\epsilon > 0$, no $f : \mathbb{R}^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

Spherical Spin Glass: Results

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

ALG $<$ OPT (for $p \geq 3$)

Result: no low-degree polynomial can achieve value OPT $- \epsilon$

Theorem [Gamarnik, Jagannath, W. '20] Let $p \geq 4$ be even.
For some $\epsilon > 0$, no $f : \mathbb{R}^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

- Objective: $H(f(Y)) \geq \text{OPT} - \epsilon$

Spherical Spin Glass: Results

Example (spherical p-spin model): for $Y \in \mathbb{R}^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

$$\max_{\|v\|=1} \frac{1}{\sqrt{n}} \langle Y, v^{\otimes p} \rangle$$

$\mathsf{ALG} < \mathsf{OPT} \quad$ (for $p \geq 3$)

Result: no low-degree polynomial can achieve value $\mathsf{OPT} - \epsilon$

Theorem [Gamarnik, Jagannath, W. '20] Let $p \geq 4$ be even.
For some $\epsilon > 0$, no $f : \mathbb{R}^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

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

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n$$

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad \mathrm{ALG} = \frac{\log d}{d}\,n$$

Result: no low-degree polynomial can achieve $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d}n$

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad \mathrm{ALG} = \frac{\log d}{d}\,n$$

Result: no low-degree polynomial can achieve $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d}n$

Theorem [Gamarnik, Jagannath, W. '20]
No polynomial $f : \{0,1\}^{\binom{n}{2}} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

Result: no low-degree polynomial can achieve $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d} n$

Theorem [Gamarnik, Jagannath, W. '20]
No polynomial $f : \{0,1\}^{\binom{n}{2}} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

- $f_i(Y) \in [0, 1/3] \cup [2/3, 1]$ for most i

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d}\,n \qquad \text{ALG} = \frac{\log d}{d}\,n$$

Result: no low-degree polynomial can achieve $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d}n$

Theorem [Gamarnik, Jagannath, W. '20]

No polynomial $f : \{0,1\}^{\binom{n}{2}} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

- $f_i(Y) \in [0, 1/3] \cup [2/3, 1]$ for most i
- $\{i : f_i(Y) \in [2/3, 1]\}$ is a near-indep set of size $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d}n$

Max Independent Set: Results

Example (max independent set): given sparse graph $G(n, d/n)$,

$$\max_{S \subseteq [n]} |S| \quad s.t. \quad S \text{ independent}$$

$$\text{OPT} = 2\,\frac{\log d}{d} n \qquad \text{ALG} = \frac{\log d}{d} n$$

Result: no low-degree polynomial can achieve $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d} n$

Theorem [Gamarnik, Jagannath, W. '20]
No polynomial $f : \{0,1\}^{\binom{n}{2}} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

- $f_i(Y) \in [0, 1/3] \cup [2/3, 1]$ for most i
- $\{i : f_i(Y) \in [2/3, 1]\}$ is a near-indep set of size $(1 + \frac{1}{\sqrt{2}})\frac{\log d}{d} n$

Forthcoming: improve $1 + \frac{1}{\sqrt{2}} \to 1 + \epsilon$ (optimal)

Optimization [Gamarnik, Jagannath, W. '20]

How to prove failure of low-degree polynomials?

Optimization [Gamarnik, Jagannath, W. '20]

How to prove failure of low-degree polynomials?

For problems with a planted signal:

- Detection: linear algebra [BHKKMP'16; HS'17; HKPRSS'17]
- Recovery: Jensen + linear algebra [Schramm, W. '20]

Optimization [Gamarnik, Jagannath, W. '20]

How to prove failure of low-degree polynomials?

For problems with a planted signal:

- Detection: linear algebra [BHKKMP'16; HS'17; HKPRSS'17]
- Recovery: Jensen + linear algebra [Schramm, W. '20]

For random optimization problems, need different approach:

Optimization [Gamarnik, Jagannath, W. '20]

How to prove failure of low-degree polynomials?

For problems with a planted signal:

- Detection: linear algebra [BHKKMP'16; HS'17; HKPRSS'17]
- Recovery: Jensen + linear algebra [Schramm, W. '20]

For random optimization problems, need different approach:

- Stability of low-degree polynomials

Optimization [Gamarnik, Jagannath, W. '20]

How to prove failure of low-degree polynomials?

For problems with a planted signal:

- Detection: linear algebra [BHKKMP'16; HS'17; HKPRSS'17]
- Recovery: Jensen + linear algebra [Schramm, W. '20]

For random optimization problems, need different approach:

- Stability of low-degree polynomials
- Overlap gap property (OGP)
  [Gamarnik, Sudan '13]
  [Chen, Gamarnik, Panchenko, Rahman '17]
  [Gamarnik, Jagannath '19]

Low-Degree Polynomials are Stable

Low-Degree Polynomials are Stable

Theorem

Let Y, Y' be ρ-correlated samples from $\mathcal{N}(0, \mathrm{I}_m)$

Low-Degree Polynomials are Stable

Theorem

Let Y, Y' be ρ-correlated samples from $\mathcal{N}(0, \mathrm{I}_m)$

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ have degree $\leq D$

Low-Degree Polynomials are Stable

Theorem

Let Y, Y' be ρ-correlated samples from $\mathcal{N}(0, \mathrm{I}_m)$

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ have degree $\leq D$

Normalization $\mathbb{E}_Y \|f(Y)\|^2 = 1$

Low-Degree Polynomials are Stable

Theorem

Let Y, Y' be ρ-correlated samples from $\mathcal{N}(0, \mathrm{I}_m)$

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ have degree $\leq D$

Normalization $\mathbb{E}_Y \|f(Y)\|^2 = 1$

Then for any $t \geq (6e)^D$,

$$\Pr\left[\|f(Y) - f(Y')\|^2 \geq 2t(1 - \rho^D)\right] \leq \exp\left(-\frac{D}{3e} t^{1/D}\right)$$

Low-Degree Polynomials are Stable

Theorem

Let Y, Y' be ρ-correlated samples from $\mathcal{N}(0, \mathrm{I}_m)$

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ have degree $\le D$

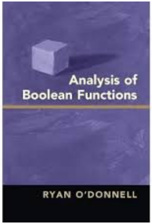
Normalization $\mathbb{E}_Y \|f(Y)\|^2 = 1$

Then for any $t \ge (6e)^D$,

$$\Pr\left[\|f(Y) - f(Y')\|^2 \ge 2t(1-\rho^D)\right] \le \exp\left(-\frac{D}{3e} t^{1/D}\right)$$

Proof: low-degree polynomials have

- Low noise sensitivty
- Low total influence
- Hypercontractivity



Low-Degree Polynomials are Stable (Binary Case)

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Interpolation path: $Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Interpolation path: $Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$

$f : \{0,1\}^m \to \mathbb{R}^n$ degree D

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Interpolation path: $Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$

$f : \{0,1\}^m \to \mathbb{R}^n$ degree D

Definition: Index i is "c-bad" if

$$\|f(Y^{(i)}) - f(Y^{(i-1)})\|^2 > c \mathop{\mathbb{E}}_{Y} \|f(Y)\|^2$$

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Interpolation path: $Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$

$f : \{0,1\}^m \to \mathbb{R}^n$ degree D

Definition: Index i is "c-bad" if

$$\|f(Y^{(i)}) - f(Y^{(i-1)})\|^2 > c \mathop{\mathbb{E}}_{Y} \|f(Y)\|^2$$

Theorem

$$\Pr_{Y^{(0)},\ldots,Y^{(m)}} [\nexists\, c\text{-bad } i] \geq p^{4D/c}$$

Low-Degree Polynomials are Stable (Binary Case)

$Y \sim$ i.i.d. Bernoulli(p)

Interpolation path: $Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$

$f : \{0,1\}^m \to \mathbb{R}^n$ degree D

Definition: Index i is "c-bad" if

$$\|f(Y^{(i)}) - f(Y^{(i-1)})\|^2 > c \mathop{\mathbb{E}}_{Y} \|f(Y)\|^2$$

Theorem

$$\Pr_{Y^{(0)},\ldots,Y^{(m)}} [\nexists\, c\text{-bad } i] \geq p^{4D/c}$$

With non-trivial probability (over path), f's output is "smooth"

Overlap Gap Property

Overlap gap property (OGP): with high probability,
$Y \sim G(n, d/n)$ has no occurrence of

Overlap Gap Property

Overlap gap property (OGP): with high probability, $Y \sim G(n, d/n)$ has no occurrence of

- S, T independent sets

Overlap Gap Property

Overlap gap property (OGP): with high probability, $Y \sim G(n, d/n)$ has no occurrence of

- S, T independent sets
- $|S|, |T| \approx (1 + \frac{1}{\sqrt{2}})\Phi$

Overlap Gap Property

Overlap gap property (OGP): with high probability, $Y \sim G(n, d/n)$ has no occurrence of

- S, T independent sets
- $|S|, |T| \approx (1 + \frac{1}{\sqrt{2}})\Phi$
- $|S \cap T| \approx \Phi$

Overlap Gap Property

Overlap gap property (OGP): with high probability, $Y \sim G(n, d/n)$ has no occurrence of

- S, T independent sets
- $|S|, |T| \approx (1 + \frac{1}{\sqrt{2}})\Phi$
- $|S \cap T| \approx \Phi$

Proof: first moment method [Gamarnik, Sudan '13]

Ensemble OGP

Ensemble OGP: with high probability, $\forall i, j$ on the interpolation path

$$Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \dots \quad Y^{(m-1)} \quad Y^{(m)}$$

there is no occurrence of

- S independent set in $Y^{(i)}$
- T independent set in $Y^{(j)}$
- $|S|, |T| \approx (1 + \frac{1}{\sqrt{2}})\Phi$
- $|S \cap T| \approx \Phi$

Putting it Together

Proof that low-degree polynomials fail:

Putting it Together

Proof that low-degree polynomials fail:

Suppose $f(Y)$ outputs independent sets of size $(1 + \frac{1}{\sqrt{2}})\Phi$

Putting it Together

Proof that low-degree polynomials fail:

Suppose $f(Y)$ outputs independent sets of size $(1 + \frac{1}{\sqrt{2}})\Phi$

$$Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \dots \quad Y^{(m-1)} \quad Y^{(m)}$$

Putting it Together

Proof that low-degree polynomials fail:

Suppose $f(Y)$ outputs independent sets of size $(1 + \frac{1}{\sqrt{2}})\Phi$

$$Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \dots \quad Y^{(m-1)} \quad Y^{(m)}$$

Separation: $f(Y^{(0)})$ and $f(Y^{(m)})$ are "far apart"

Putting it Together

Proof that low-degree polynomials fail:

Suppose $f(Y)$ outputs independent sets of size $(1+\frac{1}{\sqrt{2}})\Phi$

$$Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$$

<u>Separation</u>: $f(Y^{(0)})$ and $f(Y^{(m)})$ are "far apart"

<u>Stability</u>: with probability $\gtrsim n^{-D}$, there are no big "jumps"

$$f(Y^{(i)}) \to f(Y^{(i+1)})$$

Putting it Together

Proof that low-degree polynomials fail:

Suppose $f(Y)$ outputs independent sets of size $(1+\frac{1}{\sqrt{2}})\Phi$

$$Y^{(0)} \quad Y^{(1)} \quad Y^{(2)} \quad \ldots \quad Y^{(m-1)} \quad Y^{(m)}$$

Separation: $f(Y^{(0)})$ and $f(Y^{(m)})$ are “far apart”

Stability: with probability $\gtrsim n^{-D}$, there are no big “jumps”

$$f(Y^{(i)}) \to f(Y^{(i+1)})$$

Contradicts OGP

Comments

Comments

- Improvement to $(1+\epsilon)\frac{\log d}{d}n$
  - Inspired by [Rahman, Virág '14]

Comments

- Improvement to $(1+\epsilon)\frac{\log d}{d} n$
  - Inspired by [Rahman, Virág '14]

- Proof of OGP for p-spin (for $p \geq 4$ even)
  [Chen, Sen '15; Auffinger, Chen '17]

Comments

- Improvement to $(1+\epsilon)\frac{\log d}{d} n$
  - Inspired by [Rahman, Virág '14]
- Proof of OGP for p-spin (for $p \geq 4$ even)
  [Chen, Sen '15; Auffinger, Chen '17]
- Langevin dynamics

Comments

- Improvement to $(1+\epsilon)\frac{\log d}{d}n$
  - Inspired by [Rahman, Virág '14]
- Proof of OGP for p-spin (for $p \geq 4$ even)
  [Chen, Sen '15; Auffinger, Chen '17]
- Langevin dynamics
- Connections between heuristics
  - OGP $\to$ Low-Degree

References for the Low-Degree Framework

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

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

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