Low-Degree Hardness of Random Optimization Problems

Alex Wein
Courant Institute, New York University

Joint work with:

David Gamarnik
MIT

Aukosh Jagannath
Waterloo

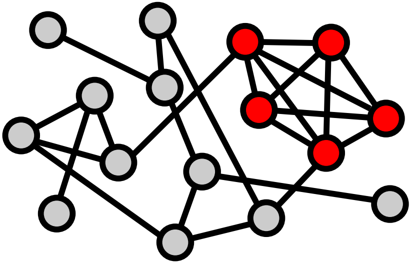
Random Optimization Problems

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 <u>algorithmically</u> (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 <u>algorithmically</u> (ALG)?

Q: In cases where it seems hard to reach a particular objective value, can we understand why? In a <u>unified</u> 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 $G(n, d/n)$,

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

Max Independent Set

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

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

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad (d \text{ large constant})$$

[Frieze '90]

Max Independent Set

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

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

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad \mathrm{ALG} = \frac{\log d}{d}\,n \qquad (d \text{ large constant})$$

[Frieze '90]

Max Independent Set

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

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

$$\text{OPT} = 2\,\frac{\log d}{d} n \qquad \text{ALG} = \frac{\log d}{d} n \qquad (d \text{ large constant})$$

[Frieze '90]

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

Max Independent Set

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

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

$$\mathrm{OPT} = 2\,\frac{\log d}{d}\,n \qquad \mathrm{ALG} = \frac{\log d}{d}\,n \qquad (d \text{ large constant})$$

[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 $G(n, d/n)$,

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

$$\text{OPT} = 2\,\frac{\log d}{d} n \qquad \text{ALG} = \frac{\log d}{d} n \qquad (d \text{ large constant})$$

[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 $Y \in (\mathbb{R}^n)^{\otimes p}$ i.i.d. $\mathcal{N}(0,1)$,

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

(maximize random degree-p polynomial over the sphere)

Spherical Spin Glass

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

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

(maximize random degree-p polynomial over the sphere)

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

Spherical Spin Glass

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

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

(maximize random degree-p polynomial over the sphere)

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

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

Spherical Spin Glass

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

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

(maximize random degree-p polynomial over the sphere)

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

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

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

Spherical Spin Glass

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

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

(maximize random degree-p polynomial over the sphere)

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

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

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

Approximate message passing (AMP) algorithms achieve value ALG_p **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 <u>not</u> 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 <u>not</u> 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. $v \in \mathbb{R}^n$

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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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. $v \in \mathbb{R}^n$
- "Low" degree 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

Low-degree algorithms are already well-studied for problems with a planted signal

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

[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

Planted Problems

Low-degree algorithms are already well-studied for problems with a planted signal

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

[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

For a wide range of planted problems, $O(\log n)$-degree polynomials are as powerful as the best known poly-time algorithms

Planted Problems

Low-degree algorithms are already well-studied for problems with a planted signal

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

[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

For a wide range of planted problems, $O(\log n)$-degree polynomials are as powerful as the best known poly-time algorithms

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

[BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19,S**W**20,...]

Planted Problems

Low-degree algorithms are already well-studied for problems with a planted signal

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

[Hopkins, Steurer '17]

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

[Hopkins '18] (PhD thesis)

For a wide range of planted problems, $O(\log n)$-degree polynomials are as powerful as the best known poly-time algorithms

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

[BHKKMP16,HS17,HKPRSS17,Hop18,BK**W**19,K**W**B19,DK**W**B19,S**W**20,...]

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

Results: Spherical Spin Glass

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

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

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

Results: Spherical Spin Glass

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

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

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

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

Results: Spherical Spin Glass

Example (spherical p-spin model): for $Y \in (\mathbb{R}^n)^{\otimes p}$ 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 $f : (\mathbb{R}^n)^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

Results: Spherical Spin Glass

Example (spherical p-spin model): for $Y \in (\mathbb{R}^n)^{\otimes p}$ 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 $f : (\mathbb{R}^n)^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

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

Results: Spherical Spin Glass

Example (spherical p-spin model): for $Y \in (\mathbb{R}^n)^{\otimes p}$ 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 $f : (\mathbb{R}^n)^{\otimes p} \to \mathbb{R}^n$ of degree $\mathrm{polylog}(n)$ achieves both of the following with probability $1 - \exp(-n^{\Omega(1)})$:

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

Results: Max Independent Set

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

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

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

Results: Max Independent Set

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

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

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

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

Results: Max Independent Set

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

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

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

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

Theorem [Gamarnik, Jagannath, W. '20; 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 $1 - \exp(-n^{\Omega(1)})$:

Results: Max Independent Set

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

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

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

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

Theorem [Gamarnik, Jagannath, W. '20; 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

Results: Max Independent Set

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

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

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

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

Theorem [Gamarnik, Jagannath, W. '20; 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 $(1+\epsilon)\frac{\log d}{d} n$

Proof Techniques

How to prove failure of low-degree polynomials?

Proof Techniques

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]

Proof Techniques

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:

Proof Techniques

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

Proof Techniques

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]
  [Rahman, Virág '14]
  [Chen, Gamarnik, Panchenko, Rahman '17]
  [Gamarnik, Jagannath '19]

Low-Degree Polynomials are Stable

Low-Degree Polynomials are Stable

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

Low-Degree Polynomials are Stable

$Y \sim$ 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

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

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

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

Low-Degree Polynomials are Stable

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

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

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

Definition: Step 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

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

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

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

Definition: Step i is "c-bad" if

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

Theorem

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

Low-Degree Polynomials are Stable

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

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

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

Definition: Step 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 [Gamarnik, Sudan '13]

Overlap gap property (OGP): with high probability over $Y \sim G(n, d/n)$, there does not exist $S, T \subseteq [n]$ such that

- S, T independent sets
- $|S|, |T| \geq (1 + \frac{1}{\sqrt{2}})\Phi \qquad \Phi := \frac{\log d}{d} n$
- $|S \cap T| \approx \Phi$

Overlap Gap Property [Gamarnik, Sudan '13]

Overlap gap property (OGP): with high probability over $Y \sim G(n, d/n)$, there does not exist $S, T \subseteq [n]$ such that

- S, T independent sets
- $|S|, |T| \geq (1 + \frac{1}{\sqrt{2}})\Phi$ $\qquad \Phi := \frac{\log d}{d} n$
- $|S \cap T| \approx \Phi$

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

Ensemble OGP [CGPR'17, GJ'19]

Ensemble OGP: with high probability over

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

there does not exist $S, T \subseteq [n]$ such that

- S independent set in some $Y^{(i)}$
- T independent set in some $Y^{(j)}$
- $|S|, |T| \geq (1 + \frac{1}{\sqrt{2}})\Phi \qquad \Phi := \frac{\log d}{d} n$
- $|S \cap T| \approx \Phi$

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1 + \frac{1}{\sqrt{2}})\Phi$:

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1 + \frac{1}{\sqrt{2}})\Phi$:

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

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1 + \frac{1}{\sqrt{2}})\Phi$:

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

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

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1 + \frac{1}{\sqrt{2}})\Phi$:

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"

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1 + \frac{1}{\sqrt{2}})\Phi$:

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"

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

Proof [Gamarnik, Jagannath, W. '20]

Proof that low-degree polynomials cannot reach $(1+\frac{1}{\sqrt{2}})\Phi$:

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

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Forbidden structure: with high probability over

$$Y^{(0)} \quad Y^{(1)} \quad \dots \quad Y^{(Lm)} \qquad (L \approx 1/\epsilon^2)$$

there does not exist $S_0, \dots, S_L \subseteq [n]$ such that

- S_k independent set in some $Y^{(t_k)}$
- $|S_k| \geq (1+\epsilon)\Phi$
- $|S_k \setminus (\cup_{i<k} S_i)| \in [\frac{\epsilon}{4}\Phi, \frac{\epsilon}{2}\Phi]$

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Forbidden structure: with high probability over

$$Y^{(0)} \quad Y^{(1)} \quad \ldots \quad Y^{(Lm)} \qquad (L \approx 1/\epsilon^2)$$

there does not exist $S_0, \ldots, S_L \subseteq [n]$ such that

- S_k independent set in some $Y^{(t_k)}$
- $|S_k| \geq (1+\epsilon)\Phi$
- $|S_k \setminus (\cup_{i<k} S_i)| \in [\frac{\epsilon}{4}\Phi, \frac{\epsilon}{2}\Phi]$

Proof (low-degree cannot achieve $(1+\epsilon)\Phi$)

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Forbidden structure: with high probability over

$$Y^{(0)} \quad Y^{(1)} \quad \dots \quad Y^{(Lm)} \qquad\qquad (L \approx 1/\epsilon^2)$$

there does not exist $S_0, \dots, S_L \subseteq [n]$ such that

- S_k independent set in some $Y^{(t_k)}$
- $|S_k| \geq (1+\epsilon)\Phi$
- $|S_k \setminus (\cup_{i<k} S_i)| \in [\frac{\epsilon}{4}\Phi, \frac{\epsilon}{2}\Phi]$

Proof (low-degree cannot achieve $(1+\epsilon)\Phi$)

- Suppose we had a stable algorithm that finds sets of size $\geq (1+\epsilon)\Phi$

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Forbidden structure: with high probability over

$$Y^{(0)} \quad Y^{(1)} \quad \dots \quad Y^{(Lm)} \qquad\qquad (L \approx 1/\epsilon^2)$$

there does not exist $S_0, \dots, S_L \subseteq [n]$ such that

- S_k independent set in some $Y^{(t_k)}$
- $|S_k| \geq (1+\epsilon)\Phi$
- $|S_k \setminus (\cup_{i<k} S_i)| \in [\frac{\epsilon}{4}\Phi, \frac{\epsilon}{2}\Phi]$

Proof (low-degree cannot achieve $(1+\epsilon)\Phi$)

- Suppose we had a stable algorithm that finds sets of size $\geq (1+\epsilon)\Phi$
- Use this to construct forbidden structure $\Rightarrow$ contradiction

Optimal Threshold $(1+\epsilon)\Phi$ [W. '20]

Forbidden structure: with high probability over

$$Y^{(0)} \quad Y^{(1)} \quad \ldots \quad Y^{(Lm)} \qquad (L \approx 1/\epsilon^2)$$

there does not exist $S_0, \ldots, S_L \subseteq [n]$ such that

- S_k independent set in some $Y^{(t_k)}$
- $|S_k| \geq (1+\epsilon)\Phi$
- $|S_k \setminus (\cup_{i<k} S_i)| \in [\frac{\epsilon}{4}\Phi, \frac{\epsilon}{2}\Phi]$

Proof (low-degree cannot achieve $(1+\epsilon)\Phi$)

- Suppose we had a stable algorithm that finds sets of size $\geq (1+\epsilon)\Phi$
- Use this to construct forbidden structure $\Rightarrow$ contradiction
  - After m steps, algorithm's output must change

Comments

Comments

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

Comments

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

- Langevin dynamics

Comments

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

- Langevin dynamics

- Connections between heuristics
  - OGP $\rightarrow$ 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