Low-Degree Hardness of Random Optimization Problems

David Gamarnik

MIT

Probabilistic Combinatorics. September 2020

Joint work with Aukosh Jagannath (U Waterloo) and Alex Wein (NYU)

September 23, 2020

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

- Consider $\mathbb{G}(N, p)$.

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

- Consider $\mathbb{G}(N, p)$.
- The largest clique (fully connected subgraph) is $\sim 2\log_{\frac{1}{p}} N$.

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

- Consider $\mathbb{G}(N, p)$.
- The largest clique (fully connected subgraph) is $\sim 2\log_{\frac{1}{p}} N$.
- A trivial greedy algorithm finds a clique of size $\sim \log_{\frac{1}{p}} N$.

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

- Consider $\mathbb{G}(N, p)$.
- The largest clique (fully connected subgraph) is $\sim 2 \log_{\frac{1}{p}} N$.
- A trivial greedy algorithm finds a clique of size $\sim \log_{\frac{1}{p}} N$.
- Karp [1976] Find a better polynomial time algorithm.

(Arguably) Most Embarrassing Algorithmic Problem in Random Graphs

- Consider $\mathbb{G}(N, p)$.
- The largest clique (fully connected subgraph) is $\sim 2 \log_{\frac{1}{p}} N$.
- A trivial greedy algorithm finds a clique of size $\sim \log_{\frac{1}{p}} N$.
- Karp [1976] Find a better polynomial time algorithm.
- Still open. This is embarrassing...

Sparse graphs – similar story

Sparse graphs – similar story

- Consider $\mathbb{G}(N, d/N)$.

Sparse graphs – similar story

- Consider $\mathbb{G}(N, d/N)$.
- The largest independent set is $\sim 2(1 + o_d(1))\frac{\log d}{d}N$ Frieze [1990].

Sparse graphs – similar story

- Consider $\mathbb{G}(N, d/N)$.
- The largest independent set is $\sim 2(1 + o_d(1))\frac{\log d}{d}N$ Frieze [1990].
- A trivial greedy algorithm finds an independent set of size $\sim (1 + o_d(1))\,\frac{\log d}{d}N$.

Sparse graphs – similar story

- Consider $\mathbb{G}(N, d/N)$.
- The largest independent set is $\sim 2(1 + o_d(1))\frac{\log d}{d}N$ Frieze [1990].
- A trivial greedy algorithm finds an independent set of size $\sim (1 + o_d(1))\frac{\log d}{d}N$.
- Nothing better known.

Algorithmic Barriers in Random Structures

- Many other examples of an *apparent* gap between the optimal values estimated by non-constructive method, and the values achievable by fast (polynomial time) algorithms:

- Many other examples of an *apparent* gap between the optimal values estimated by non-constructive method, and the values achievable by fast (polynomial time) algorithms:

  Random K-Sat, MaxCut on random graphs, proper coloring of a random graph (dense and sparse), optimizing Hamiltonian of a p-spin glass problem, etc, etc.

- Many other examples of an *apparent* gap between the optimal values estimated by non-constructive method, and the values achievable by fast (polynomial time) algorithms:

  Random K-Sat, MaxCut on random graphs, proper coloring of a random graph (dense and sparse), optimizing Hamiltonian of a p-spin glass problem, etc, etc.
- What's the barrier?

- Many other examples of an *apparent* gap between the optimal values estimated by non-constructive method, and the values achievable by fast (polynomial time) algorithms:

  Random K-Sat, MaxCut on random graphs, proper coloring of a random graph (dense and sparse), optimizing Hamiltonian of a *p*-spin glass problem, etc, etc.

- What's the barrier? Intricate geometry of the solution space, the *Overlap Gap Property (OGP)* originating from spin glass theory.

- Many other examples of an *apparent* gap between the optimal values estimated by non-constructive method, and the values achievable by fast (polynomial time) algorithms:

  Random K-Sat, MaxCut on random graphs, proper coloring of a random graph (dense and sparse), optimizing Hamiltonian of a p-spin glass problem, etc, etc.
- What's the barrier? Intricate geometry of the solution space, the *Overlap Gap Property (OGP)* originating from spin glass theory.
- This talk: OGP - obstruction to optimization based on low-degree polynomials for spin glass models and largest ind set in $\mathbb{G}(N, d/N)$.

Overlap Gap Property

Overlap Gap Property

Generic minimization problem with random input $\mathcal{X}$

$$\min_{\theta \in \Theta} \mathcal{L}(\theta, \mathcal{X}).$$

Overlap Gap Property

Generic minimization problem with random input $\mathcal{X}$

$$\min_{\theta \in \Theta} \mathcal{L}(\theta, \mathcal{X}).$$

OGP holds if there exists $\mu > 0, \nu_1 < \nu_2$, such that for every θ_1, θ_2 satisfying

$$\mathcal{L}(\theta_j, \mathcal{X}) \leq \min_{\theta \in \Theta} \mathcal{L}(\theta, \mathcal{X}) + \mu, \qquad j = 1, 2$$

it holds

$$\|\theta_1 - \theta_2\| \in [0, \nu_1] \cup [\nu_2, +\infty)$$

Overlap Gap Property

Generic minimization problem with random input $\mathcal{X}$

$$\min_{\theta \in \Theta} \mathcal{L}(\theta, \mathcal{X}).$$

OGP holds if there exists $\mu > 0, \nu_1 < \nu_2$, such that for every θ_1, θ_2 satisfying

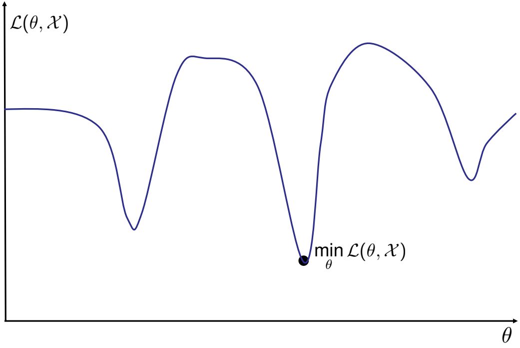
$$\mathcal{L}(\theta_j, \mathcal{X}) \leq \min_{\theta \in \Theta} \mathcal{L}(\theta, \mathcal{X}) + \mu, \qquad j = 1, 2$$

it holds

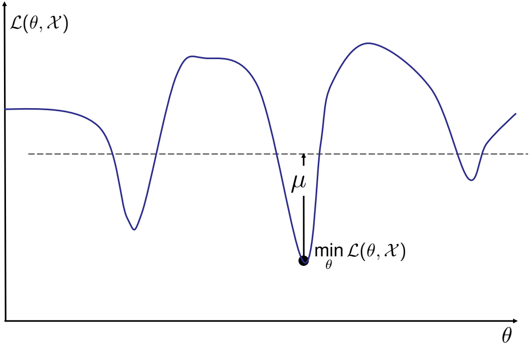
$$\|\theta_1 - \theta_2\| \in [0, \nu_1] \cup [\nu_2, +\infty)$$

That is every two μ-optimal solutions are either ”close” or ”far” from each other.

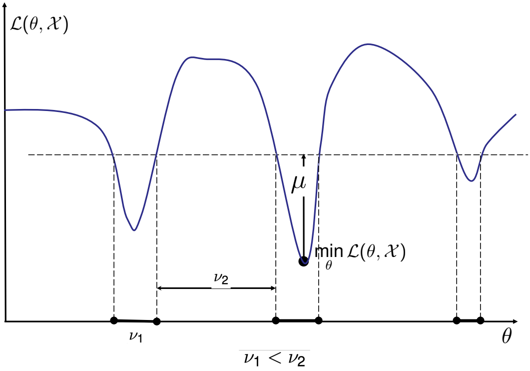
Overlap Gap Property



Overlap Gap Property



Overlap Gap Property



OGP for Largest Independent Set Problem in $\mathbb{G}(N, d/N)$

OGP for Largest Independent Set Problem in $\mathbb{G}(N, d/N)$

Theorem (G & Sudan [2017])

For every sufficiently large d and every $\beta \in \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}, 1\right)$ there exists $0 < \nu_1 < \nu_2 < 1$ such that for every two ind sets I_1, I_2 with size at least $\beta OPT \approx \beta \frac{2 \log g}{d} N$, it is the case that

$$\frac{|I_1 \cap I_2|}{OPT}$$

is either $< \nu_1$ or $> \nu_2$, w.h.p. as $n \to \infty$. Namely OGP holds.

OGP for Largest Independent Set Problem in $\mathbb{G}(N, d/N)$

Theorem (G & Sudan [2017])

For every sufficiently large d and every $\beta \in \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}, 1\right)$ *there exists* $0 < \nu_1 < \nu_2 < 1$ *such that for every two ind sets* I_1, I_2 *with size at least* $\beta OPT \approx \beta \frac{2 \log g}{d} N$*, it is the case that*

$$\frac{|I_1 \cap I_2|}{OPT}$$

is either $< \nu_1$ *or* $> \nu_2$*, w.h.p. as* $n \to \infty$*. Namely OGP holds.*

Remark

Proof: simple first moment method.

OGP for an interpolated family of $\mathbb{G}(N, d/N)$

OGP for an interpolated family of $\mathbb{G}(N, d/N)$

- Enumerate all pairs $j = 1, 2, \ldots, \binom{N}{2}$.

- Enumerate all pairs $j = 1, 2, \ldots, \binom{N}{2}$.
- Create a sequence $\mathbb{G}_j, 0 \leq j \leq \binom{N}{2}$ where $\mathbb{G}_0 = \mathbb{G}(N, d/N)$ and $\mathbb{G}_{j+1}$ is obtained from $\mathbb{G}_j$ by resampling edge $j+1$.

OGP for an interpolated family of $\mathbb{G}(N, d/N)$

- Enumerate all pairs $j = 1, 2, \ldots, \binom{N}{2}$.
- Create a sequence $\mathbb{G}_j, 0 \le j \le \binom{N}{2}$ where $\mathbb{G}_0 = \mathbb{G}(N, d/N)$ and $\mathbb{G}_{j+1}$ is obtained from $\mathbb{G}_j$ by resampling edge $j + 1$.
- Note $\mathbb{G}_j \stackrel{d}{=} \mathbb{G}(N, d/N)$. $\mathbb{G}_0$ and $\mathbb{G}_{\binom{N}{2}}$ are independent.

OGP for an interpolated family of $\mathbb{G}(N, d/N)$

OGP for an interpolated family of $\mathbb{G}(N, d/N)$

Theorem

For every sufficiently large d and every $\beta \in \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}, 1\right)$ *there exists* $0 < \nu_1 < \nu_2 < 1$ *such that for every* $0 \leq j_1 \leq j_2 \leq \binom{N}{2}$ *and every two ind sets* I_1 *in* $\mathbb{G}_{j_1}$ *and* I_2 *in* $\mathbb{G}_{j_2}$ *with size at least* $\beta OPT \approx \beta \frac{2 \log g}{d} N$, *it is the case that*

$$\frac{|I_1 \cap I_2|}{OPT}$$

is either $< \nu_1$ *or* $> \nu_2$. *Furthermore, when* $j_1 = 0$ *and* $j_2 = \binom{N}{2}$, *only the case* $< \nu_1$ *is possible.*

Optimizing a Hamiltonian of a p-spin model. Ising

Optimizing a Hamiltonian of a p-spin model. Ising

- $A = (A_{i_1,\ldots,i_p}, 1 \le i_1, \ldots, i_p \le N) \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.

Optimizing a Hamiltonian of a p-spin model. Ising

- $A = (A_{i_1,\dots,i_p}, 1 \le i_1, \dots, i_p \le N) \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
- Optimization over Hamming cube $\mathcal{B}_N \triangleq \{-1, 1\}^N$

$$\min_{\sigma \in \mathcal{B}_N} \langle A, \sigma^{\otimes p} \rangle = \min_{\sigma \in \mathcal{B}_N} \sum_{1 \le i_1,\dots,i_p \le N} A_{i_1,\dots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}$$

Optimizing a Hamiltonian of a p-spin model. Ising

- $A = (A_{i_1,\dots,i_p}, 1 \le i_1, \dots, i_p \le N) \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
- Optimization over Hamming cube $\mathcal{B}_N \triangleq \{-1, 1\}^N$

$$\min_{\sigma \in \mathcal{B}_N} \langle A, \sigma^{\otimes p} \rangle = \min_{\sigma \in \mathcal{B}_N} \sum_{1 \le i_1,\dots,i_p \le N} A_{i_1,\dots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}$$

- When $p = 2$

$$\min_{\sigma \in \mathcal{B}_{\mathcal{N}}} \sigma^T A \sigma.$$

Optimizing a Hamiltonian of a p-spin model. Spherical

Optimizing a Hamiltonian of a p-spin model. Spherical

- $A = (A_{i_1,\ldots,i_p}, 1 \leq i_1, \ldots, i_p \leq N) \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.

Optimizing a Hamiltonian of a p-spin model. Spherical

- $A = (A_{i_1,\dots,i_p}, 1 \le i_1, \dots, i_p \le N) \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
- Optimization over sphere $\mathcal{S}_N \triangleq \{x \in \mathbb{R}^N : \|x\|_2 = \sqrt{N}\}$.

$$\min_{\sigma \in \mathcal{S}_N} \langle A, \sigma^{\otimes p} \rangle = \min_{\sigma \in \mathcal{S}_N} \sum_{1 \le i_1,\dots,i_p \le N} A_{i_1,\dots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}$$

Optimizing a Hamiltonian of a p-spin model. Spherical

- $A = (A_{i_1,\dots,i_p}, 1 \le i_1,\dots,i_p \le N) \in \mathbb{R}^{N\otimes p}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
- Optimization over sphere $\mathcal{S}_N \triangleq \{x \in \mathbb{R}^N : \|x\|_2 = \sqrt{N}\}$.

$$\min_{\sigma\in\mathcal{S}_N} \langle A, \sigma^{\otimes p}\rangle = \min_{\sigma\in\mathcal{S}_N} \sum_{1\le i_1,\dots,i_p\le N} A_{i_1,\dots,i_p}\sigma_{i_1}\cdots\sigma_{i_p}$$

- When $p = 2$ (easy),

$$\min_{\sigma\in\mathcal{S}_{\mathcal{N}}} \sigma^T A\sigma.$$

Background

Background

- With high probability the limits exists and can be computed Talagrand [2005], confirming prediction by Parisi [1979]. Panchenko's book [2013]

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{B}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\mathrm{Bin}} < 0$$

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{S}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\mathrm{Sp}} < 0.$$

Background

- With high probability the limits exists and can be computed Talagrand [2005], confirming prediction by Parisi [1979]. Panchenko's book [2013]

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{B}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\mathrm{Bin}} < 0$$

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{S}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\mathrm{Sp}} < 0.$$

- $\eta^*_{\mathrm{Bin}} \approx -0.763166$ when $p = 2$.

Background

- With high probability the limits exists and can be computed Talagrand [2005], confirming prediction by Parisi [1979]. Panchenko's book [2013]

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{B}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\text{Bin}} < 0$$

$$\lim_{N\to\infty} \frac{1}{N} \min_{\sigma\in\mathcal{S}_N} \langle A, \sigma^{\otimes p}\rangle = \eta^*_{\text{Sp}} < 0.$$

- $\eta^*_{\text{Bin}} \approx -0.763166$ when $p = 2$.
- **Algorithmic goal:** find $\sigma_{\text{Alg}} \in \mathcal{B}_N$ or $\in \mathcal{S}_N$ (depending on the problem) such that w.h.p. as $N \to \infty$.

$$\frac{1}{N}\langle A, \sigma_{\text{Alg}}^{\otimes p}\rangle \leq -\eta^* + \epsilon,$$

for every $\epsilon > 0$.

Existing algorithms

Existing algorithms

- Montanari [2018] Solved the problem when $p = 2$.
  Assumes unproven conjecture of no OGP.

Existing algorithms

- Montanari [2018] Solved the problem when $p = 2$. Assumes unproven conjecture of no OGP.
- Subag [2018] Same for (mixture of) spherical p-spin: optimization over $\mathcal{S}_N = \{\sigma \in \mathbb{R}^N : \|\sigma\|_2 = \sqrt{N}\}$ when no OGP.

Existing algorithms

- Montanari [2018] Solved the problem when $p = 2$. Assumes unproven conjecture of no OGP.
- Subag [2018] Same for (mixture of) spherical p-spin: optimization over $\mathcal{S}_N = \{\sigma \in \mathbb{R}^N : \|\sigma\|_2 = \sqrt{N}\}$ when no OGP.
- Both motivated by iteration scheme proposed by Bolthausen [2014]

OGP. Definition

OGP. Definition

Fix two independent copies $A, \hat{A} \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
Interpolate the instances $\mathcal{A} = \{\sqrt{1-\tau}A + \sqrt{\tau}\hat{A}\}$.

OGP. Definition

Fix two independent copies $A, \hat{A} \in \mathbb{R}^{N^{\otimes p}}$ i.i.d. $\mathcal{N}(0, N^{-\frac{p-1}{2}})$.
Interpolate the instances $\mathcal{A} = \{\sqrt{1-\tau}A + \sqrt{\tau}\hat{A}\}$.

Definition

The set $\mathcal{A}$ satisfies the Overlap Gap Property (OGP) with domain $\mathcal{X}_N \subset \mathbb{R}^N$, and parameters $\mu > 0, 0 < \nu_1 < \nu_2 < 1$ if for every $0 \leq \tau_1, \tau_2 \leq 1$ and every $\sigma_1, \sigma_2 \in \mathcal{X}_N$ satisfying

$$\frac{1}{N}\langle A_{\tau_j}, \sigma_j^{\otimes p}\rangle \leq \inf_{\sigma \in \mathcal{X}_N} \frac{1}{N}\langle A_{\tau_j}, \sigma^{\otimes p}\rangle + \mu, \qquad j = 1, 2,$$

it holds

$$\frac{|\langle \sigma_1, \sigma_2\rangle|}{N} \in [0, \nu_1] \cup [\nu_2, 1].$$

OGP. Facts and conjectures

OGP. Facts and conjectures

Theorem (Auffinger & Chen [2018], G, Panchenko & Rahman [2019])

OGP holds for every $p \geq 4$ for the domains $\mathcal{B}_N = \{-1, 1\}^N$ and $\mathcal{S}_N = \{\sigma \in \mathbb{R}^N : \|\sigma\|_2 = \sqrt{N}\}$

$p = 2$ conjectured not to exhibit OGP for $\mathcal{B}_N$.

Algorithms: Low-degree polynomials

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1, \dots, x_{N^p}) = \sum_{|S| \leq D, S \subset [N^p]} \beta_{j,S} \prod_{i \in S} x_i, \qquad 1 \leq j \leq N.$$

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1,\ldots,x_{N^p}) = \sum_{|S|\le D, S\subset[N^p]} \beta_{j,S} \prod_{i\in S} x_i, \qquad 1\le j\le N.$$

- The proposed solution is $P(A) = \big(P_j(A), 1\le j\le N\big)$. Round if necessary.

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1, \ldots, x_{N^p}) = \sum_{|S| \leq D, S \subset [N^p]} \beta_{j,S} \prod_{i \in S} x_i, \qquad 1 \leq j \leq N.$$

- The proposed solution is $P(A) = \big(P_j(A), 1 \leq j \leq N\big)$. Round if necessary.
- Algorithms which can be modeled by low degree polynomials.

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1, \ldots, x_{N^p}) = \sum_{|S| \le D, S \subset [N^p]} \beta_{j,S} \prod_{i \in S} x_i, \qquad 1 \le j \le N.$$

- The proposed solution is $P(A) = \big(P_j(A), 1 \le j \le N\big)$. Round if necessary.

- Algorithms which can be modeled by low degree polynomials.
  - Local algorithms on graphs

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1,\ldots,x_{N^p}) = \sum_{|S|\le D, S\subset[N^p]} \beta_{j,S} \prod_{i\in S} x_i, \qquad 1\le j\le N.$$

- The proposed solution is $P(A) = \big(P_j(A), 1 \le j \le N\big)$. Round if necessary.

- Algorithms which can be modeled by low degree polynomials.
  - Local algorithms on graphs
  - Spectral methods

Algorithms: Low-degree polynomials

- Fix N degree-D polynomials

$$P_j(x_1, \ldots, x_{N^p}) = \sum_{|S| \leq D, S \subset [N^p]} \beta_{j,S} \prod_{i \in S} x_i, \qquad 1 \leq j \leq N.$$

- The proposed solution is $P(A) = \big(P_j(A), 1 \leq j \leq N\big)$. Round if necessary.

- Algorithms which can be modeled by low degree polynomials.
  - Local algorithms on graphs
  - Spectral methods
  - Approximate Message Passing

Main result. Failure of low-degree polynomials. Spin glasses

Main result. Failure of low-degree polynomials. Spin glasses

Low-D polymial based methods fails to optimize p-spin spherical model with high probability promise.

Main result. Failure of low-degree polynomials. Spin glasses

Low-D polymial based methods fails to optimize p-spin spherical model with high probability promise.

Theorem (G, Jagannath & Wein [2020], some technical assumptions skipped)

Fix even $p \geq 4$. Suppose $\mathbb{E}\|P(A)\|_2^2 = N$. Let $Q(A) = \sqrt{N} P(A)/\|P(A)\|_2$. If

$$\mathbb{P}\left(N^{-1}\langle A, Q^{\otimes p}(A)\rangle \leq \eta\right) \geq 1 - (1/4)\exp(-2D),$$

then $\eta \geq \eta^* + \mu$, where μ comes from OGP. I.e. degree-D polynomials cannot optimize with "high" promise.

Main result. Failure of low-degree polynomials. Spin glasses

Low-D polymial based methods fails to optimize p-spin spherical model with high probability promise.

Theorem (G, Jagannath & Wein [2020], some technical assumptions skipped)

Fix even $p \geq 4$. Suppose $\mathbb{E}\|P(A)\|_2^2 = N$. Let $Q(A) = \sqrt{N}P(A)/\|P(A)\|_2$. If

$$\mathbb{P}\left(N^{-1}\langle A, Q^{\otimes p}(A)\rangle \leq \eta\right) \geq 1 - (1/4)\exp(-2D),$$

then $\eta \geq \eta^* + \mu$, where μ comes from OGP. I.e. degree-D polynomials cannot optimize with "high" promise.

Similar result holds for Ising model.

Main result. Failure of low-degree polynomials. Largest i.s. in $\mathbb{G}(N, d/N)$

Main result. Failure of low-degree polynomials. Largest i.s. in $\mathbb{G}(N, d/N)$

- Consider any degree-D polynomial $P(A) \in \mathbb{R}^N$ where A – adjacency matrix of $\mathbb{G}(N, d/N)$.

Main result. Failure of low-degree polynomials. Largest i.s. in $\mathbb{G}(N, d/N)$

- Consider any degree-D polynomial $P(A) \in \mathbb{R}^N$ where A – adjacency matrix of $\mathbb{G}(N, d/N)$.
- Rounding: a) select all nodes i with $P_i(A) \geq 1$; b) delete nodes violating ind set constraint.

Main result. Failure of low-degree polynomials. Largest i.s. in $\mathbb{G}(N, d/N)$

- Consider any degree-D polynomial $P(A) \in \mathbb{R}^N$ where A – adjacency matrix of $\mathbb{G}(N, d/N)$.
- Rounding: a) select all nodes i with $P_i(A) \geq 1$; b) delete nodes violating ind set constraint.

Theorem (G, Jagannath & Wein [2020], some technical assumptions skipped)

Consider the independent set produced by the degree-D polynomial P(A) plus rounding. Suppose $|I| = \beta(\log d/d)N$ *with probability at least* $1 - \exp(-\Omega(D \log N))$. *Then for all large enough d* $\beta \leq \beta^*_{OGP}$.

I.e. degree-D polynomials cannot produce an ind set above OGP threshold with "high promise".

Proof idea: stability of algorithms

Proof idea: stability of algorithms

- Proof by contradiction. $\mu > 0, 0 < \nu_1 < \nu_2 < 1$ – parameters of OGP. Suppose $\sigma_{\mathrm{Alg}}(A)$ satisfies $N^{-1}\langle A, \sigma_{\mathrm{Alg}}^{\otimes p}\rangle < \eta^* + \mu$ with "good enough" probability.

Proof idea: stability of algorithms

- Proof by contradiction. $\mu > 0, 0 < \nu_1 < \nu_2 < 1$ – parameters of OGP. Suppose $\sigma_{\mathrm{Alg}}(A)$ satisfies $N^{-1}\langle A, \sigma_{\mathrm{Alg}}^{\otimes p}\rangle < \eta^* + \mu$ with "good enough" probability.
- Key property – **Stability.** Small changes in A result in small changes in σ_{Alg} – most difficult part. Stability is established using noise sensitivity type arguments.

Noise sensitivity

Noise sensitivity

- Suppose $A, \hat{A} \in \mathbb{R}^{N \otimes p}$ are gaussian ρ-correlated. Suppose $f = (f_1, \ldots, f_N) : \mathbb{R}^{N \otimes p} \to \mathbb{R}^N$ consists of degree D polynomials and $\mathbb{E}[\|f(A)\|_2^2] = \mathbb{E}[\|f(\hat{A})\|_2^2] = 1$.

Noise sensitivity

- Suppose $A, \hat{A} \in \mathbb{R}^{N \otimes p}$ are gaussian ρ-correlated. Suppose $f = (f_1, \ldots, f_N) : \mathbb{R}^{N \otimes p} \to \mathbb{R}^N$ consists of degree D polynomials and $\mathbb{E}[\|f(A)\|_2^2] = \mathbb{E}[\|f(\hat{A})\|_2^2] = 1$.

Theorem

For any $t \geq (6e)^D$

$$\mathbb{P}\left(\|f(A) - f(\hat{A})\|_2^2 \geq 2t(1-\rho^D)\right) \leq \exp\left(-\frac{D}{3e} t^{\frac{1}{D}}\right).$$

Proof idea: stability of algorithms

Proof idea: stability of algorithms

- Recall interpolation $A_\tau \triangleq \sqrt{1-\tau}A + \sqrt{\tau}\hat{A}, \tau \in [0,1]$.

Proof idea: stability of algorithms

- Recall interpolation $A_\tau \triangleq \sqrt{1-\tau}A + \sqrt{\tau}\hat{A}, \tau \in [0,1]$.
- If $\tau_2 - \tau_1$ is "small" then $\|\sigma_{\mathrm{Alg}}(A_{\tau_2}) - \sigma_{\mathrm{Alg}}(A_{\tau_1})\|_2$ is small as well – **stability**

Proof idea: stability of algorithms

- Recall interpolation $A_\tau \triangleq \sqrt{1-\tau}A + \sqrt{\tau}\hat{A}, \tau \in [0,1]$.
- If $\tau_2 - \tau_1$ is "small" then $\|\sigma_{\mathrm{Alg}}(A_{\tau_2}) - \sigma_{\mathrm{Alg}}(A_{\tau_1})\|_2$ is small as well – **stability**
- That is $N^{-1}\langle\sigma_{\mathrm{Alg}}(A_\tau), \sigma_{\mathrm{Alg}}(A_0)\rangle$ changes continuously in τ.

Proof idea: stability of algorithms

- Recall interpolation $A_\tau \triangleq \sqrt{1-\tau}A + \sqrt{\tau}\hat{A}, \tau \in [0,1]$.
- If $\tau_2 - \tau_1$ is "small" then $\|\sigma_{\rm Alg}(A_{\tau_2}) - \sigma_{\rm Alg}(A_{\tau_1})\|_2$ is small as well – **stability**
- That is $N^{-1}\langle\sigma_{\rm Alg}(A_\tau), \sigma_{\rm Alg}(A_0)\rangle$ changes continuously in τ.
- On the other hand, when A and $\hat{A}$ are independent $N^{-1}\langle\sigma_{\rm Alg}(A_0), \sigma_{\rm Alg}(A_1)\rangle$ is $o(1)$ and thus $< \nu_1$.

Proof idea: stability of algorithms

- Recall interpolation $A_\tau \triangleq \sqrt{1-\tau}A + \sqrt{\tau}\hat{A}, \tau \in [0,1]$.
- If $\tau_2 - \tau_1$ is "small" then $\|\sigma_{\mathrm{Alg}}(A_{\tau_2}) - \sigma_{\mathrm{Alg}}(A_{\tau_1})\|_2$ is small as well – **stability**
- That is $N^{-1}\langle \sigma_{\mathrm{Alg}}(A_\tau), \sigma_{\mathrm{Alg}}(A_0)\rangle$ changes continuously in τ.
- On the other hand, when A and $\hat{A}$ are independent $N^{-1}\langle \sigma_{\mathrm{Alg}}(A_0), \sigma_{\mathrm{Alg}}(A_1)\rangle$ is $o(1)$ and thus $< \nu_1$.
- Thus for some τ, $N^{-1}\langle \sigma_{\mathrm{Alg}}(A_\tau), \sigma_{\mathrm{Alg}}(A_0)\rangle \in (\nu_1, \nu_2)$ – contradiction to μ-optimality of σ_{Alg}. $\square$

Other problems and algorithms ruled out by OGP

Other problems and algorithms ruled out by OGP

- Local algorithms for $\mathbb{G}(N, d/N)$ are half-optimal at best Rahman & Virag [2017] (improving on G & Sudan [2017]).

Other problems and algorithms ruled out by OGP

- Local algorithms for $\mathbb{G}(N, d/N)$ are half-optimal at best Rahman & Virag [2017] (improving on G & Sudan [2017]).
- Finding a near-satisfying assignment of a random NAE-K-SAT problem G & Sudan [2017] using sequential local algorithms.

Other problems and algorithms ruled out by OGP

- Local algorithms for $\mathbb{G}(N, d/N)$ are half-optimal at best Rahman & Virag [2017] (improving on G & Sudan [2017]).
- Finding a near-satisfying assignment of a random NAE-K-SAT problem G & Sudan [2017] using sequential local algorithms.
- WALKSAT for random NAE-K-SAT problem Coja-Oghlan, Haqshenas & Hetterich [2017]

Other problems and algorithms ruled out by OGP

- Local algorithms for $\mathbb{G}(N, d/N)$ are half-optimal at best Rahman & Virag [2017] (improving on G & Sudan [2017]).
- Finding a near-satisfying assignment of a random NAE-K-SAT problem G & Sudan [2017] using sequential local algorithms.
- WALKSAT for random NAE-K-SAT problem Coja-Oghlan, Haqshenas & Hetterich [2017]
- Finding a large cut of a random sparse hypergraph using local algorithms Chen, G, Panchenko & Rahman [2019]

Other problems and algorithms ruled out by OGP

- Local algorithms for $\mathbb{G}(N, d/N)$ are half-optimal at best Rahman & Virag [2017] (improving on G & Sudan [2017]).
- Finding a near-satisfying assignment of a random NAE-K-SAT problem G & Sudan [2017] using sequential local algorithms.
- WALKSAT for random NAE-K-SAT problem Coja-Oghlan, Haqshenas & Hetterich [2017]
- Finding a large cut of a random sparse hypergraph using local algorithms Chen, G, Panchenko & Rahman [2019]
- Finding a large ind set in $\mathbb{G}(N, d/N)$ using quantum local algorithm (QAOA) G, Farhi & Gutmann [2020]

Problems in high-dimensional inference exhibiting OGP

Problems in high-dimensional inference exhibiting OGP

- High-dimensional regression below LASSO threshold G & Zadik [2017]

Problems in high-dimensional inference exhibiting OGP

- High-dimensional regression below LASSO threshold G & Zadik [2017]
- Planted Clique Problem G & Zadik [2019]

Problems in high-dimensional inference exhibiting OGP

- High-dimensional regression below LASSO threshold G & Zadik [2017]
- Planted Clique Problem G & Zadik [2019]
- Maximum sub-matrix of a random matrix G & Li [2018]

Problems in high-dimensional inference exhibiting OGP

- High-dimensional regression below LASSO threshold G & Zadik [2017]
- Planted Clique Problem G & Zadik [2019]
- Maximum sub-matrix of a random matrix G & Li [2018]
- Sparse PCA Arous, Wein & Zadik [2020], G, Jagannath & Sen [2020]

Thank you