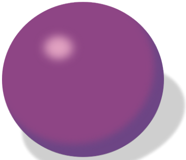
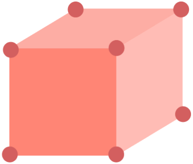
On maximizing polynomials



$$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$$



$$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$$

based on joint works with

Tim Hsieh

CMU

Chris Jones

Bocconi

Pravesh Kothari

Princeton

Lucas Pesenti

Bocconi

Luca Trevisan

Bocconi

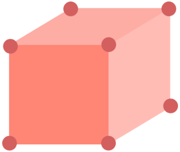
Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

$$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$$

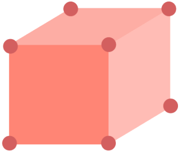


$$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$$

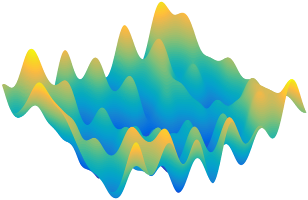
Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

$$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$$



$$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$$

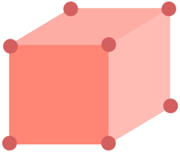


Nonconvex optimization

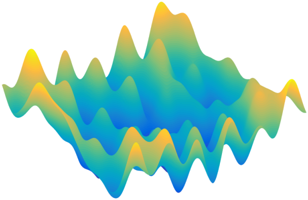
Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$

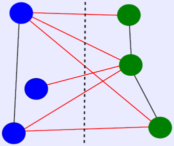


$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$



Nonconvex optimization

Ex: Max-Cut

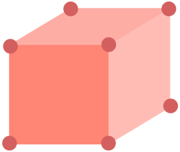


Fraction of edges cut by $\mathbf{x}$:
$\frac{1}{2} - \frac{1}{2}\sum_{i\sim j} x_i x_j$

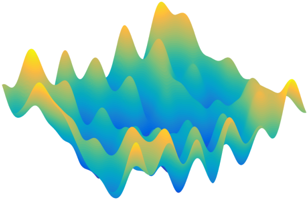
Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$

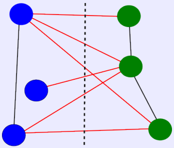


$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$



Nonconvex optimization

Ex: Max-Cut



Fraction of edges cut by $\mathbf{x}$:

$\frac{1}{2} - \frac{1}{2}\sum_{i\sim j} x_i x_j$

Ex: Principal Component Analysis

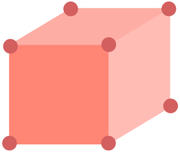
$$p(\mathbf{x}) = \sum c_{ijk} x_i x_j x_k$$

with $c_{ijk} = x_i^* x_j^* x_k^* + \text{noise}$

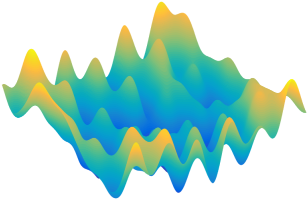
Motivation (1/2)

Maximize over x: $p(x_1, \ldots, x_n) := \sum_i c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \sum_{i<j<k} c_{ijk} x_i x_j x_k$

$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$

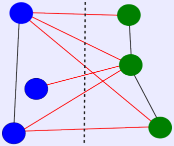


$\max_{\mathbf{x}\in\{-1,1\}^n} p(\mathbf{x})$



Nonconvex optimization

Ex: Max-Cut



Fraction of edges cut by $\mathbf{x}$:
$\frac{1}{2} - \frac{1}{2}\sum_{i\sim j} x_i x_j$

Ex: Principal Component Analysis

$$p(\mathbf{x}) = \sum c_{ijk} x_i x_j x_k$$

with $c_{ijk} = x_i^* x_j^* x_k^* + \text{noise}$

Ex: Max-3SAT

$(x_1 \vee \neg x_3 \vee x_4)$

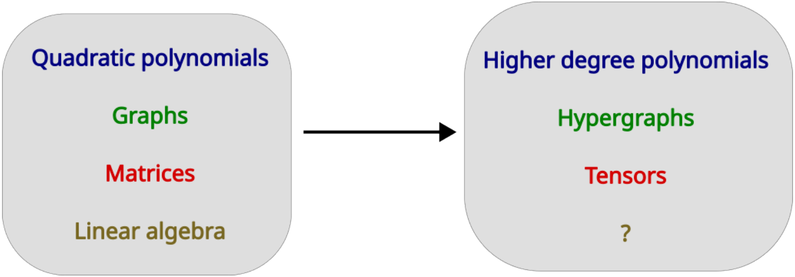
$\wedge\ (\neg x_2 \vee x_3 \vee \neg x_5)$

$\wedge\ \ldots$

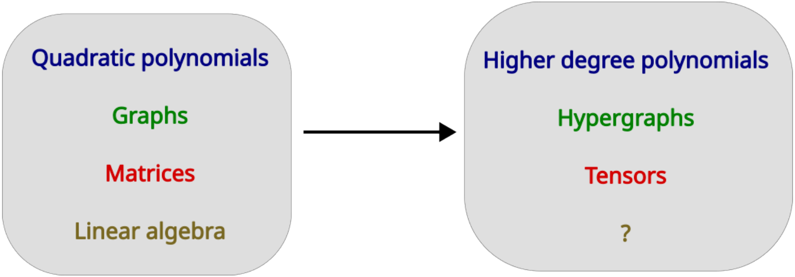
Fraction of clauses satisfied by $\mathbf{x}$:

$\frac{7}{8} + p_1(\mathbf{x}) + p_2(\mathbf{x}) + p_3(\mathbf{x})$

Motivation (2/2)



Motivation (2/2)

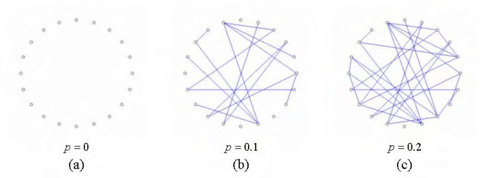


Today:

1. Spectral theory for hypergraphs
2. Approximation algorithms
3. Optimizing random polynomials

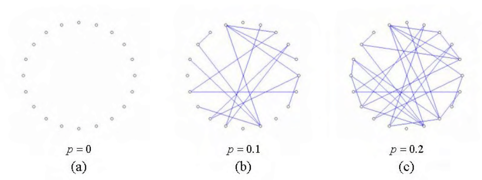
1. Spectral theory for hypergraphs

Spectral bound for random hypergraphs



$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound for random hypergraphs

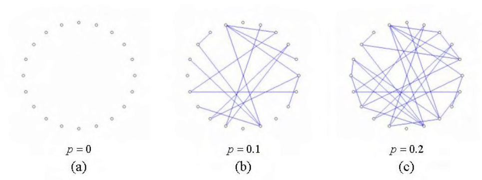


$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound (max cut, max independent set, ...):

$$\|\mathbf{A}\| := \max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} \sum_{i,j=1}^{n} A_{ij} x_i y_j$$

Spectral bound for random hypergraphs



$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound (max cut, max independent set, ...):

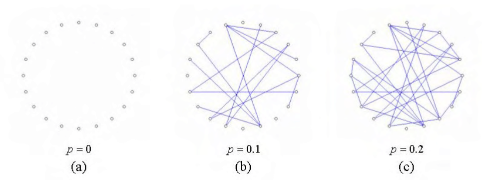
$$\|\mathbf{A}\| := \max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} \sum_{i,j=1}^{n} A_{ij} x_i y_j$$

Thm: $\|\mathbf{A}\| = O(\sqrt{np})$ (if $np \gg \log n$)

Trace method:

$$\|\mathbf{A}\| \asymp \mathrm{Tr}\left(\mathbf{A}^{2k}\right)^{\frac{1}{2k}} \text{ if } k \sim \log n\,.$$

Spectral bound for random hypergraphs



$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound (max cut, max independent set, ...):

$$\|\mathbf{A}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1} \sum_{i,j=1}^{n} A_{ij} x_i y_j$$

Thm: $\|\mathbf{A}\| = O(\sqrt{np})$ (if $np \gg \log n$)

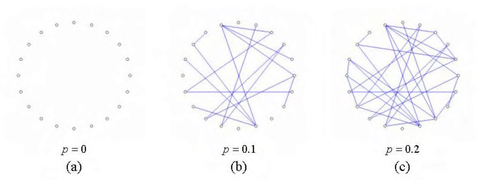
Trace method:

$$\|\mathbf{A}\| \asymp \mathrm{Tr}\left(\mathbf{A}^{2k}\right)^{\frac{1}{2k}} \text{ if } k \sim \log n\,.$$

[Friedman-Wigderson'95] for tensors?

$$\|\mathbf{T}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = \|\mathbf{z}\|_2 = 1} \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$$

Spectral bound for random hypergraphs



$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound (max cut, max independent set, ...):

$$\|\mathbf{A}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1} \sum_{i,j=1}^{n} A_{ij} x_i y_j$$

Thm: $\|\mathbf{A}\| = O(\sqrt{np})$ (if $np \gg \log n$)

Trace method:

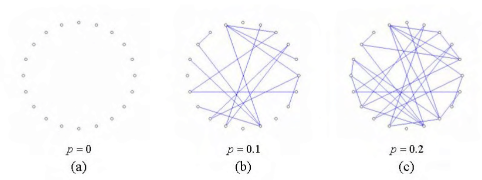
$$\|\mathbf{A}\| \asymp \mathrm{Tr}\left(\mathbf{A}^{2k}\right)^{\frac{1}{2k}} \text{ if } k \sim \log n\,.$$

[Friedman-Wigderson'95] for tensors?

$$\|\mathbf{T}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = \|\mathbf{z}\|_2 = 1} \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$$

[P.-Trevisan] $\|\mathbf{T}\| = O(1)$ if $p \ll 1/n$

Spectral bound for random hypergraphs



$$A_{ij} \underset{\text{i.i.d.}}{\sim} \begin{cases} 1-p & \text{w. prob. } p \\ -p & \text{w. prob. } 1-p \end{cases}$$

Spectral bound (max cut, max independent set, ...):

$$\|\mathbf{A}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1} \sum_{i,j=1}^{n} A_{ij} x_i y_j$$

Thm: $\|\mathbf{A}\| = O(\sqrt{np})$ (if $np \gg \log n$)

Trace method:

$$\|\mathbf{A}\| \asymp \operatorname{Tr}\left(\mathbf{A}^{2k}\right)^{\frac{1}{2k}} \text{ if } k \sim \log n\,.$$

✗ no tensor analog!

[Friedman-Wigderson'95] for tensors?

$$\|\mathbf{T}\| := \max_{\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = \|\mathbf{z}\|_2 = 1} \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$$

[P.-Trevisan] $\|\mathbf{T}\| = O(1)$ if $p \ll 1/n$

✗ Flattenings

✓ Union bounds/Chaining

2. Approximation algorithms

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1, 1\}^n$?

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij} x_i x_j$

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij}x_i x_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ijk}x_i x_j x_k$

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1, 1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij} x_i x_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ijk} x_i x_j x_k$

Posted on **October 12, 2021**

The Khot-Naor Approximation Algorithm for 3-XOR

Today I would like to discuss the Khot-Naor approximation algorithm for the 3-XOR problem, and an open question related to it.

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij} x_i x_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ijk} x_i x_j x_k$

Posted on October 12, 2021

The Khot-Naor Approximation Algorithm for 3-XOR

Today I would like to discuss the Khot-Naor approximation algorithm for the 3-XOR problem, and an open question related to it.

New SDP Roundings and Certifiable Approximation for Cubic Optimization

Jun-Ting Hsieh* Pravesh K. Kothari† Lucas Pesenti‡ Luca Trevisan§

October 3, 2023

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij} x_i x_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ijk} x_i x_j x_k$

Posted on **October 12, 2021**

The Khot-Naor Approximation Algorithm for 3-XOR

Today I would like to discuss the Khot-Naor approximation algorithm for the 3-XOR problem, and an open question related to it.

New SDP Roundings and Certifiable Approximation for Cubic Optimization

Jun-Ting Hsieh* Pravesh K. Kothari† Lucas Pesenti‡ Luca Trevisan§

October 3, 2023

Open: getting $\frac{1}{n^{1/4}}$?

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ij} x_i x_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{\mathbf{x}\in\{-1,1\}^n} \sum c_{ijk} x_i x_j x_k$

Posted on October 12, 2021

The Khot-Naor Approximation Algorithm for 3-XOR

Today I would like to discuss the Khot-Naor approximation algorithm for the 3-XOR problem, and an open question related to it.

New SDP Roundings and Certifiable Approximation for Cubic Optimization

Jun-Ting Hsieh* Pravesh K. Kothari† Lucas Pesenti‡ Luca Trevisan§

October 3, 2023

Open: getting $\frac{1}{n^{1/4}}$?

Open: given a satisfiable 3SAT formula, what fraction of clauses can we efficiently satisfy?

Approximation algorithms

Given a polynomial with **arbitrary** coefficients, how well can **efficient** algorithms maximize it over $\{-1,1\}^n$?

[Charikar-Wirth'04] $\widetilde{\Omega}(1)$-approximation to $\max_{x\in\{-1,1\}^n} \sum c_{ij}x_ix_j$

[Khot-Naor'07] $\widetilde{\Omega}(\frac{1}{\sqrt{n}})$-approximation to $\max_{x\in\{-1,1\}^n} \sum c_{ijk}x_ix_jx_k$

Posted on October 12, 2021

The Khot-Naor Approximation Algorithm for 3-XOR

Today I would like to discuss the Khot-Naor approximation algorithm for the 3-XOR problem, and an open question related to it.

New SDP Roundings and Certifiable Approximation for Cubic Optimization

Jun-Ting Hsieh* Pravesh K. Kothari† Lucas Pesenti‡ Luca Trevisan§

October 3, 2023

Open: getting $\frac{1}{n^{1/4}}$?

Open: given a satisfiable 3SAT formula, what fraction of clauses can we efficiently satisfy?

✗ [Hastad'01] $\frac{7}{8}+\epsilon$ is NP-hard.

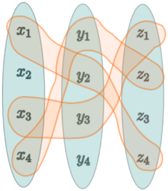
✓ [Hsieh-Kothari-P.-Trevisan'24] $\frac{7}{8}+\widetilde{\Omega}(\frac{1}{n^{3/4}})$

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?

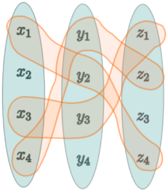


Decoupling:

✓ degree-3: $\max_{\mathbf{x}} \sum c_{ijk} x_i x_j x_k \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum c_{ijk} x_i y_j z_k$

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?



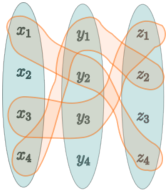
Decoupling:

✓ degree-3: $\max_{\mathbf{x}} \sum c_{ijk} x_i x_j x_k \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum c_{ijk} x_i y_j z_k$

✗ degree-2: $p(\mathbf{x}) = -\sum x_i x_j$

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?



Decoupling:

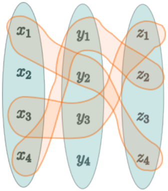
✓ degree-3: $\max_{\mathbf{x}} \sum c_{ijk} x_i x_j x_k \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum c_{ijk} x_i y_j z_k$

✗ degree-2: $p(\mathbf{x}) = -\sum x_i x_j$

Open: What is the worst *value* of non-homogeneous degree-3 polynomials?

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?



Decoupling:

✓ degree-3: $\max_{\mathbf{x}} \sum c_{ijk} x_i x_j x_k \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum c_{ijk} x_i y_j z_k$

✗ degree-2: $p(\mathbf{x}) = -\sum x_i x_j$

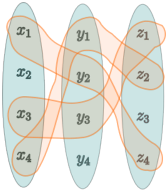
Open: What is the worst *value* of non-homogeneous degree-3 polynomials?

For all p,

$$\frac{1}{\|\mathbf{c}\|_1} \max_{\mathbf{x} \in \{-1,1\}^n} p(\mathbf{x}) \gtrsim \frac{1}{n^{1.5}}$$

Non-homogeneous polynomials

What goes wrong when $p(\mathbf{x}) = \sum c_{ij} x_i x_j + \sum c_{ijk} x_i x_j x_k$?



Decoupling:

✓ degree-3: $\max_{\mathbf{x}} \sum c_{ijk} x_i x_j x_k \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum c_{ijk} x_i y_j z_k$

✗ degree-2: $p(\mathbf{x}) = -\sum x_i x_j$

Open: What is the worst *value* of non-homogeneous degree-3 polynomials?

For all p,

$$\frac{1}{\|\mathbf{c}\|_1} \max_{\mathbf{x} \in \{-1,1\}^n} p(\mathbf{x}) \gtrsim \frac{1}{n^{1.5}}$$

There exists p such that

$$\frac{1}{\|\mathbf{c}\|_1} \max_{\mathbf{x} \in \{-1,1\}^n} p(\mathbf{x}) \lesssim \frac{1}{n}$$

3. Optimizing random polynomials

Optimizing random polynomials

$$\max_{\mathbf{x}\in\{-1,1\}^n} \sum_{i,j=1}^{n} c_{ij} x_i x_j$$

 Typical/random $c_{ij} \sim \{-1, 1\}$?

Optimizing random polynomials

$$\max_{\mathbf{x}\in\{-1,1\}^n} \sum_{i,j=1}^{n} c_{ij} x_i x_j$$

 Typical/random $c_{ij} \sim \{-1, 1\}$?

[Parisi'79, Talagrand'06] $\max p(\mathbf{x}) \approx (P_* \pm o(1))n^{1.5}$ for most (c_{ij})

Optimizing random polynomials

$$\max_{\mathbf{x}\in\{-1,1\}^n} \sum_{i,j=1}^{n} c_{ij} x_i x_j$$

 Typical/random $c_{ij} \sim \{-1, 1\}$?

[Parisi'79, Talagrand'06] $\max p(\mathbf{x}) \approx (P_* \pm o(1))n^{1.5}$ for most (c_{ij})

[..., Montanari'19] efficiently find $\mathbf{x}$ s.t. $p(\mathbf{x}) \geqslant (P_* - o(1))n^{1.5}$ for most (c_{ij})

Optimizing random polynomials

$$\max_{\mathbf{x}\in\{-1,1\}^n} \sum_{i,j=1}^{n} c_{ij} x_i x_j$$

 Typical/random $c_{ij} \sim \{-1,1\}$?

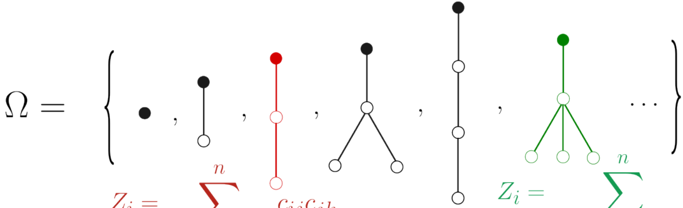
[Parisi'79, Talagrand'06] $\max p(\mathbf{x}) \approx (P_* \pm o(1))n^{1.5}$ for most (c_{ij})

[..., Montanari'19] efficiently find $\mathbf{x}$ s.t. $p(\mathbf{x}) \geqslant (P_* - o(1))n^{1.5}$ for most (c_{ij})

Physics: ✓ simple ✗ non-rigorous

Maths: ✓ rigorous ✗ technical

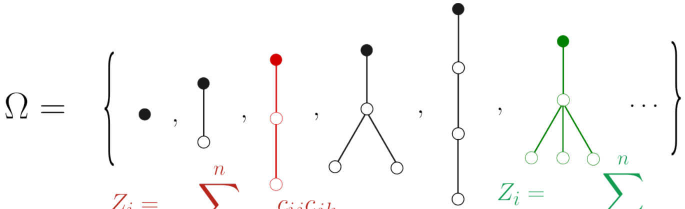
The Fourier tree basis



$$Z_i = \sum_{\substack{j,k=1 \\ i,j,k \text{ distinct}}}^{n} c_{ij}c_{jk}$$

$$Z_i = \sum_{\substack{j,k,\ell,m=1 \\ i,j,k,\ell,m \text{ distinct}}}^{n} c_{ij}c_{jk}c_{j\ell}c_{jm}$$

The Fourier tree basis



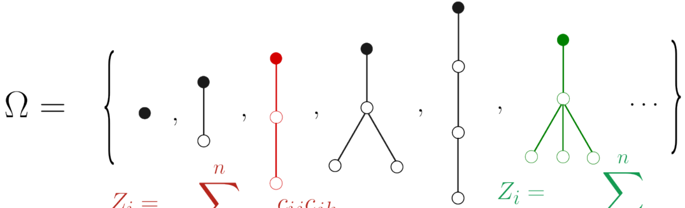
$$Z_i = \sum_{\substack{j,k=1 \\ i,j,k \text{ distinct}}}^{n} c_{ij}c_{jk}$$

$$Z_i = \sum_{\substack{j,k,\ell,m=1 \\ i,j,k,\ell,m \text{ distinct}}}^{n} c_{ij}c_{jk}c_{j\ell}c_{jm}$$

[Jones-P.'24+]

1. Ω consists of *asympotically independent Gaussian* vectors.

The Fourier tree basis



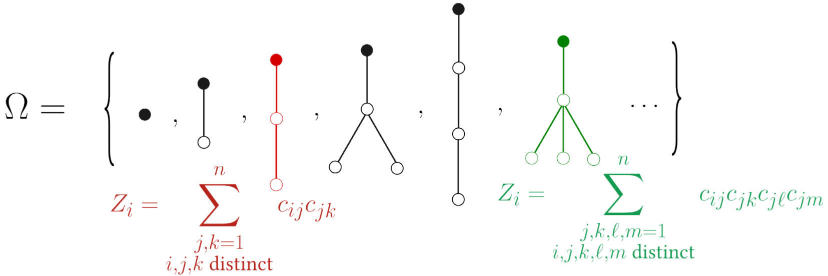
$$Z_i = \sum_{\substack{j,k=1\\ i,j,k \text{ distinct}}}^{n} c_{ij}c_{jk}$$

$$Z_i = \sum_{\substack{j,k,\ell,m=1\\ i,j,k,\ell,m \text{ distinct}}}^{n} c_{ij}c_{jk}c_{j\ell}c_{jm}$$

[Jones-P.'24+]

1. Ω consists of *asympotically independent Gaussian* vectors.
2. Ω asymptotically spans all iterates of *nonlinear iterative algorithms*.

The Fourier tree basis



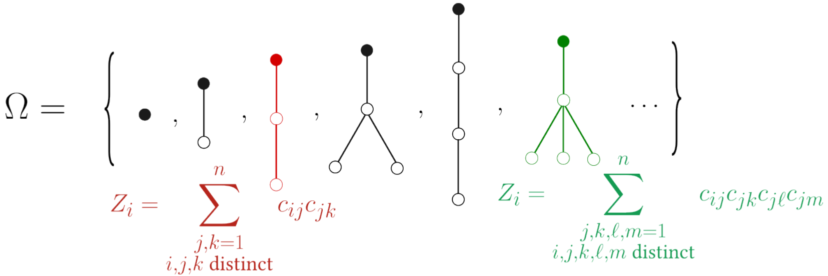
[Jones-P.'24+]

1. Ω consists of *asympotically independent Gaussian* vectors.
2. Ω asymptotically spans all iterates of *nonlinear iterative algorithms*.

✓ Simple analysis of algorithms for maximizing random polynomials

[state evolution]

The Fourier tree basis



[Jones-P.'24+]

1. Ω consists of *asympotically independent Gaussian* vectors.
2. Ω asymptotically spans all iterates of *nonlinear iterative algorithms*.

✓ Simple analysis of algorithms for maximizing random polynomials [state evolution]

✓ New expression for the value that they achieve [extended Parisi formula]

Conclusion

1. Beyond trace method & flattenings $\implies$ spectral theory for hypergraphs.
2. Maximizing non-homogeneous polynomials?
3. Full power of the Fourier tree basis?

Conclusion

1. Beyond trace method & flattenings $\implies$ spectral theory for hypergraphs.
2. Maximizing non-homogeneous polynomials?
3. Full power of the Fourier tree basis?

Thank you, Luca.