Almost Optimal Sum of Squares Lower Bound for *Planted Clique*

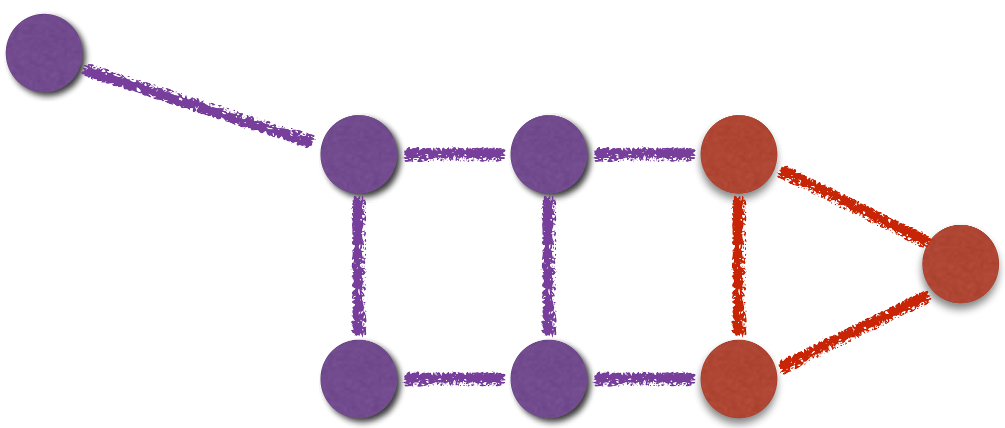
Pravesh K. Kothari
UT Austin

Based on joint works with Boaz Barak, Sam Hopkins, Jon Kelner, Ankur Moitra and Aaron Potechin.

Clique in Random Graphs

find the largest set of vertices all connected to each other

in a random graph $G \sim G(n, \frac{1}{2})$

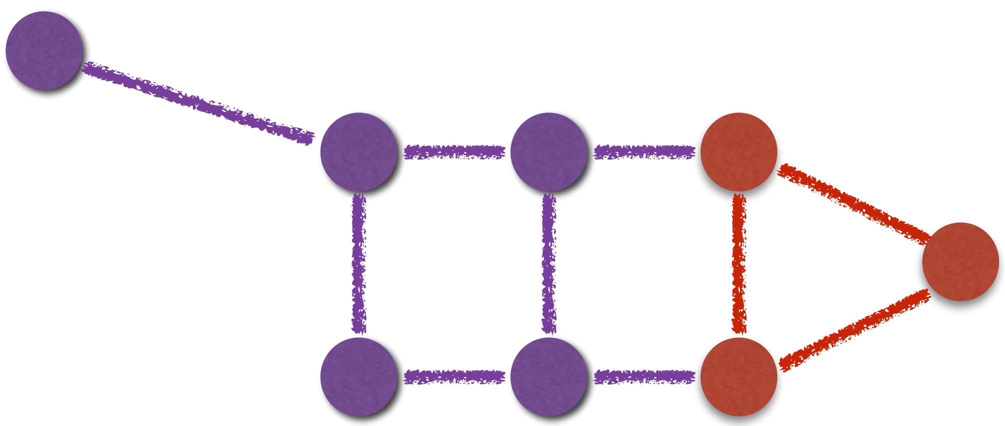


each edge in G with prob. 0.5

Clique in Random Graphs

find the largest set of vertices all connected to each other

in a random graph $G \sim G(n, \frac{1}{2})$



each edge in G with prob. 0.5

w.h.p. max clique $\sim 2\log_2(n)$

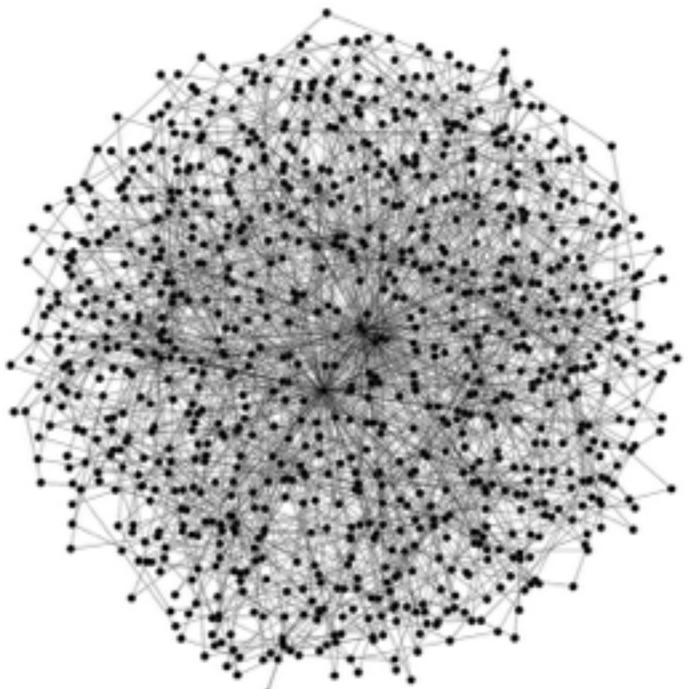
greedy finds cliques of size $\sim \log_2(n)$

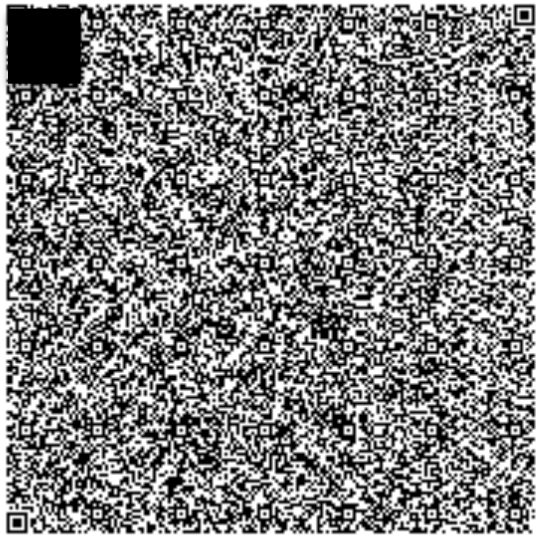
Karp'76: hard to find cliques of size $\sim (1+\epsilon)\log_2(n)$?

Planted Clique

[Jerrum'92, Kucera'95]

find an ω-clique added to a random graph $G \sim G(n, \frac{1}{2})$

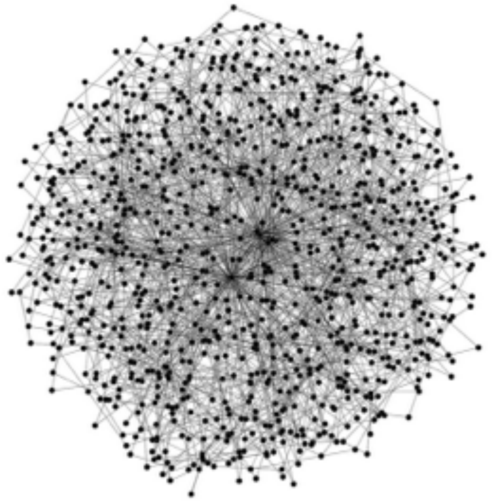
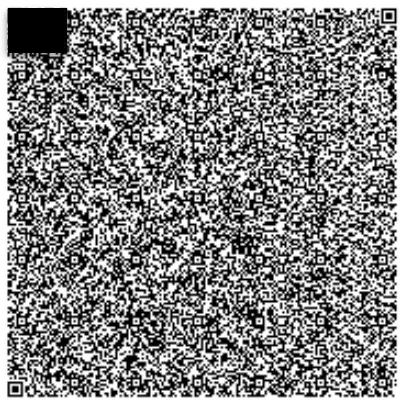




Planted Clique

[Jerrum'92, Kucera'95]

find an ω***-clique added to a random graph*** $G \sim G(n, \frac{1}{2})$

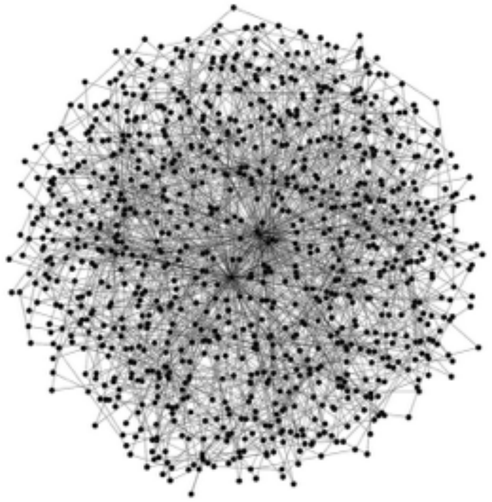
 

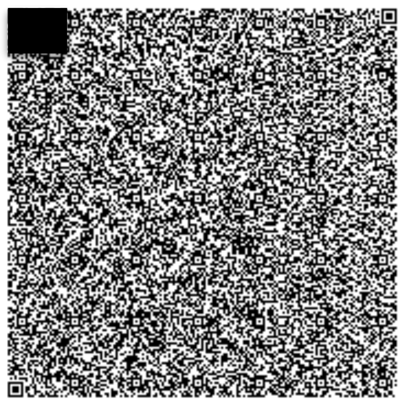
spectral algorithm: $\omega > cn^{1/2}$ **[AKS'98]**

Planted Clique

[Jerrum'92, Kucera'95]

find an ω-clique added to a random graph $G \sim G(n, \frac{1}{2})$





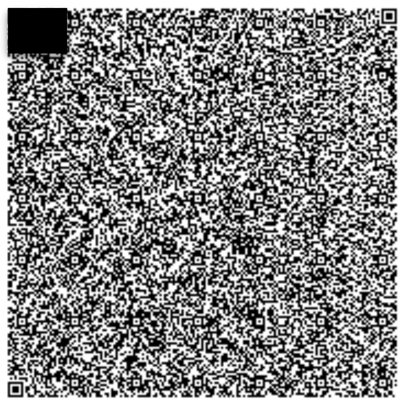
spectral algorithm: $\omega > cn^{1/2}$ **[AKS'98]**

no polynomial time algorithm when $\omega \ll \sqrt{n}$

Planted Clique

[Jerrum'92, Kucera'95]

find an ω***-clique added to a random graph*** $G \sim G(n, \frac{1}{2})$



spectral algorithm: $\omega > cn^{1/2}$ **[AKS'98]**

no polynomial time algorithm when $\omega \ll \sqrt{n}$

"Statistical Algorithms" **[FGRVX'13]**

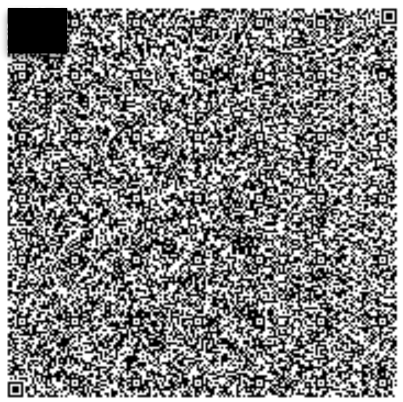
MCMC Methods **[Jer'92]**

Lovász-Schrijver **[FK'00,03]**

Planted Clique

[Jerrum'92, Kucera'95]

find an ω***-clique added to a random graph*** $G \sim G(n, \frac{1}{2})$



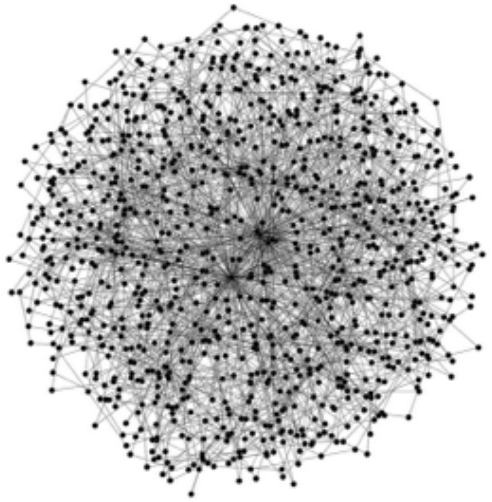
central problem in average case complexity

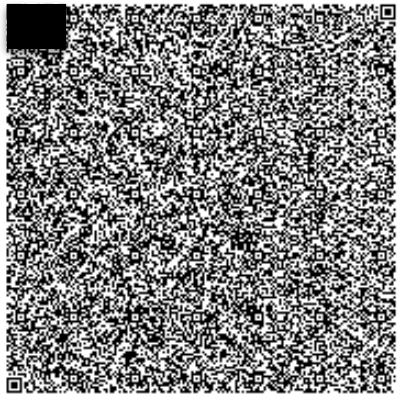
Cryptography	**[Jue'02, ABW'10]**	Sparse PCA	**[BR'13]**
Motifs in Biological Networks	**[Milo et. al. Science'02, Lotem et. al. PNAS'14]**	Nash Equilibrium	**[HK'09]**
Certifying Restricted Isometry Property	**[KL'12]**	Mathematical Finance	**[ABBG'09]**

Planted Clique

[Jerrum'92, Kucera'95]

find an ω-clique added to a random graph $G \sim G(n, \frac{1}{2})$





a **testbed** for algorithmic techniques for detecting "***signal***" (added clique) surrounded by "*random noise*" (random graph)

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

hierarchical strengthening of the basic SDP/spectral algo.

d: "degree" in the hierarchy.

$n^{O(d)}$: running time

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

SPARSEST CUT [ARV04]

UNIQUE GAMES [ABS10, BRS12, GS12]

….breaks known integrality gap instances for weaker hierarchies in **polynomial** time for

UNIQUE GAMES, BALANCED SEPARATOR, MAX CUT

[BBHKSZ12, OZ13,DMN13]

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

SPARSEST CUT **[ARV04]**

UNIQUE GAMES **[ABS10, BRS12, GS12]**

….breaks known integrality gap instances for weaker hierarchies in **polynomial** time for

UNIQUE GAMES, BALANCED SEPARATOR, MAX CUT

[BBHKSZ12, OZ13,DMN13]

…optimal amongst all poly size SDPs for a large class. **[LRS15]**

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

... successful for other *average case* problems.

[BKS'14,15]

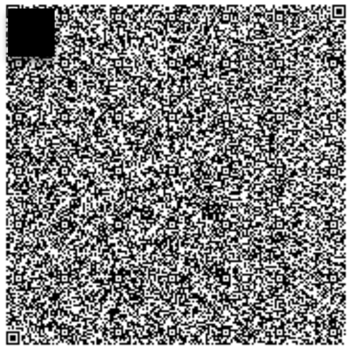
PLANTED SPARSE VECTOR, DICTIONARY LEARNING

SoS vs Planted Clique?

[Meka-Wigderson'13]

?

vs

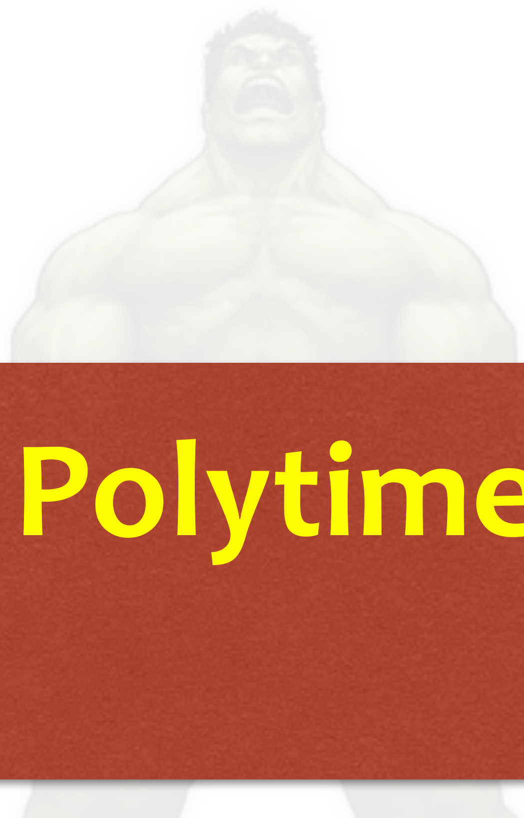


SOS

Planted Clique

SoS vs Planted Clique?

[Meka-Wigderson'13]



Conjecture [Meka-Wigderson'13]

Polytime SoS cannot find $\tilde{o}(\sqrt{n})$-size planted cliques.

SOS

Planted Clique

$\tilde{o}$ ignores $n^{o(1)}$ factors.

SoS vs Planted Clique?

[Meka-Wigderson'13]

Conjecture [Meka-Wigderson'13]

Const. *d*-SoS cannot find $\tilde{o}(\sqrt{n})$-size planted cliques.

SOS

Planted Clique

$\tilde{o}$ ignores $n^{o(1)}$ factors.

SoS vs Planted Clique?

[Meka-Potechin-Wigderson'15]	d	$\tilde{O}(n^{\frac{1}{d}})$
[Deshpande-Montanari'15]	4	$\tilde{O}(n^{\frac{1}{3}})$
[Hopkins-K-Potechin'15] [Raghavendra-Schramm'15]	4	$\tilde{O}(\sqrt{n})$

$\tilde{O}$ ignores $n^{o(1)}$ factors.

SoS vs Planted Clique?

[Meka-Potechin-Wigderson'15]	d	$\tilde{O}(n^{\frac{1}{d}})$
[Deshpande-Montanari'15]	4	$\tilde{O}(n^{\frac{1}{3}})$
[Hopkins-K-Potechin'15] [Raghavendra-Schramm'15]	4	$\tilde{O}(\sqrt{n})$
[Barak-Hopkins-Kelner-K-Moitra-Potechin'16]	d	$\tilde{O}(\sqrt{n})$

Proving MW Conj. for all d!

SoS vs Planted Clique?

[Meka-Potechin-Wigderson'15]	d	$\tilde{O}(n^{\frac{1}{d}})$
[Deshpande-Montanari'1		
[Hopkins-K-Potechin' **[Raghavendra-Schramm'**		
[Barak-Hopkins-Kelner-K-Moitra-Potechin'16]	**d**	$\tilde{O}(\sqrt{n})$

"Bayesian" view of SoS Algorithm.

Proving MW Conj. for all d!

Talk Plan

1. Sum of Squares Proofs and Pseudodistributions

2. Meka-Wigderson Pseudodistribution

3. MW Pseudodistribution is not PSD

4. The Bayesian Pseudodistribution

Talk Plan

1. Sum of Squares Proofs and Pseudodistributions

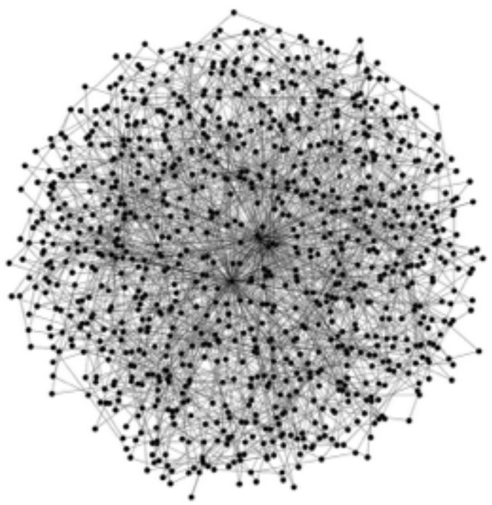
2. Meka-Wigderson Pseudodistribution

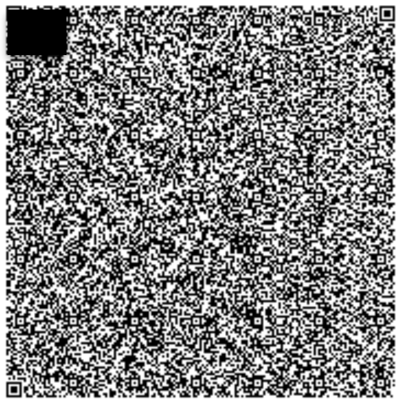
3. MW Pseudodistribution is not PSD

4. The Bayesian Pseudodistribution

Planted Clique

[Jerrum'92, Kucera'95]

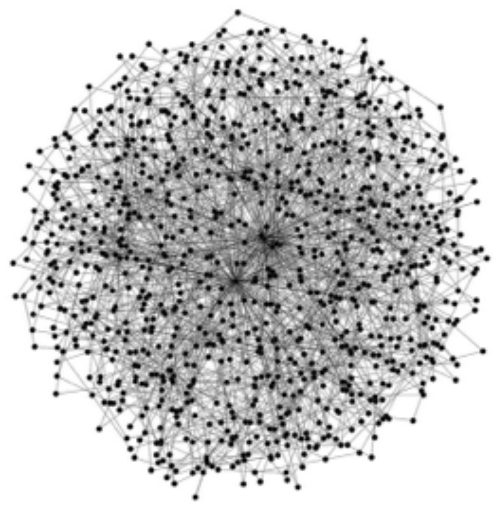


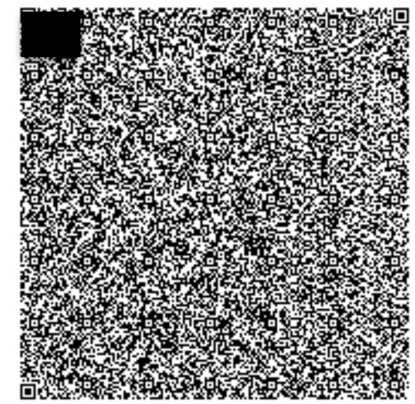


Is there an ω clique in G?

Planted Clique

[Jerrum'92, Kucera'95]





x indicates a clique

x indicates a set of size ω

x is a vector of 0-1 indicators.

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

Is there an ω clique in G?

$$\sum x_i = \omega$$

$$x_i^2 = x_i \quad \forall\, i$$

Sum of Squares Algorithm

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

μ : **"a prob. dist. on ω cliques."**

μ :"knowledge" of large cliques in G.

Sum of Squares Algorithm

μ :a prob. dist. on ω-cliques.

$\mathbb{E} = \mathbb{E}_\mu$, q: any polynomial.

μ satisfies constraints.

$$\mathbb{E}[q(x)\left(\sum x_i - \omega\right)] = 0$$

$$\mathbb{E}[q(x)x_ix_j] = 0 \qquad i \not\sim j \text{ in G}$$

$$\mathbb{E}[q(x)\left(\sum x_i^2 - x_i\right)] = 0 \quad \forall i$$

Sum of Squares Algorithm

μ :a prob. dist. on ω-cliques.

$\mathbb{E} = \mathbb{E}_\mu$, q: any polynomial.

μ satisfies constraints.

$$\mathbb{E}[q(x)\left(\sum x_i - \omega\right)] = 0$$

$$\mathbb{E}[q(x)x_i x_j] = 0 \qquad i \not\sim j \text{ in G}$$

$$\mathbb{E}[q(x)\left(\sum x_i^2 - x_i\right)] = 0 \quad \forall i$$

μ is positive semidefinite. $\mathbb{E}[q^2] \geq 0$

Sum of Squares Algorithm

μ :a prob. dist. on ω-cliques.

$\mathbb{E} = \mathbb{E}_\mu$, q: any polynomial.

μ satisfies constraints.

$$\mathbb{E}[q(x)\left(\sum x_i - \omega\right)] = 0$$

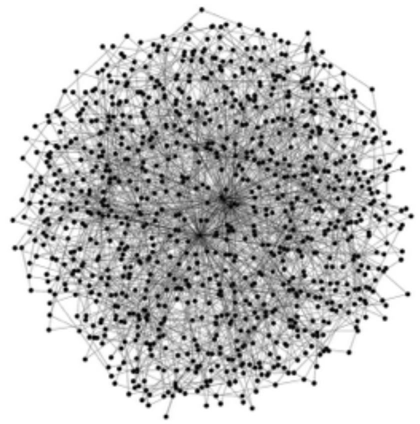
$$\mathbb{E}[q(x)x_ix_j] = 0 \qquad i \not\sim j \text{ in G}$$

$$\mathbb{E}[q(x)\left(\sum x_i^2 - x_i\right)] = 0 \quad \forall i$$

μ is positive semidefinite. $\mathbb{E}[q^2] \geq 0$

μ is normalized. $\mathbb{E}[1] = 1$

Sum of Squares Algorithm



μ :a prob. dist. on ω-cliques.

$\mathbb{E} = \mathbb{E}_\mu$, q: any polynomial.

$\mathbb{E}$ is linear.

Moments $\mathbb{E}[\Pi_{i \in S} x_i]$

μ satisfies constraints.

$$\mathbb{E}[q(x)\left(\sum x_i - \omega\right)] = 0$$

$$\mathbb{E}[q(x)x_i x_j] = 0 \qquad i \not\sim j \text{ in G}$$

$$\mathbb{E}[q(x)\left(\sum x_i^2 - x_i\right)] = 0 \quad \forall i$$

μ is positive semidefinite. $\mathbb{E}[q^2] \geq 0$

μ is normalized. $\mathbb{E}[1] = 1$

Sum of Squares Algorithm

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

Sum of Squares Algorithm

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mu}$: "a ***pseudo*** **dist. on** ω **cliques."**

$\tilde{\mu}$:"knowledge" of large cliques in G.

Sum of Squares Algorithm

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mu}$: **"a *pseudo* dist. on ω cliques."**

*of a **"computationally bounded"** agent*

$\tilde{\mu}$:"knowledge" of large cliques in G.

Sum of Squares Algorithm

$\tilde{\mu}$:a prob. dist. on ω-cliques.

$$\tilde{\mathbb{E}} = \tilde{\mathbb{E}}_{\tilde{\mu}}$$

$\tilde{\mathbb{E}}$ is linear.

PseudoMoments $\tilde{\mathbb{E}}[\Pi_{i\in S} x_i]$

$\tilde{\mu}$ satisfies constraints.

$$\tilde{\mathbb{E}}[q(x)\left(\sum x_i - \omega\right)] = 0$$

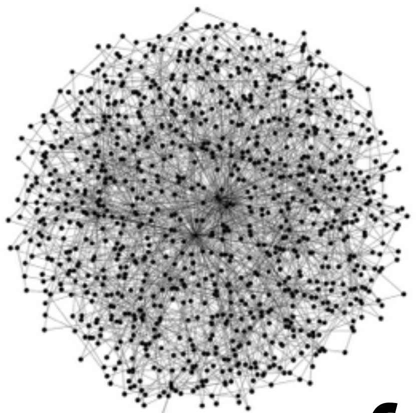
$$\tilde{\mathbb{E}}[q(x)x_i x_j] = 0 \quad i \not\sim j \text{ in G}$$

$$\tilde{\mathbb{E}}[q(x)\left(x_i^2 - x_i\right)] = 0 \quad \forall i$$

$\tilde{\mu}$ is positive semidefinite. $\tilde{\mathbb{E}}[q(x)^2] \geq 0$

$\tilde{\mu}$ is normalized. $\tilde{\mathbb{E}}[1] = 1$

Sum of Squares Algorithm



$\tilde{\mu}$:a prob. dist. on ω-cliques.

$$\tilde{\mathbb{E}} = \tilde{\mathbb{E}}_{\tilde{\mu}}$$

Total degree of argument of $\tilde{\mathbb{E}}$ ***is at most d!***

$\tilde{\mathbb{E}}$ is linear. **PseudoMoments** $\tilde{\mathbb{E}}[\Pi_{i \in S} x_i]$

$\tilde{\mu}$ satisfies constraints.

$$\tilde{\mathbb{E}}[q(x)\left(\sum x_i - \omega\right)] = 0$$

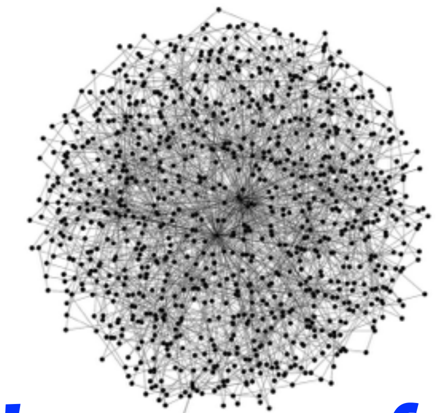
$$\tilde{\mathbb{E}}[q(x)x_i x_j] = 0 \quad i \not\sim j \text{ in G}$$

$$\tilde{\mathbb{E}}[q(x)\left(x_i^2 - x_i\right)] = 0 \quad \forall i$$

$\tilde{\mu}$ is positive semidefinite. $\tilde{\mathbb{E}}[q(x)^2] \geq 0$

$\tilde{\mu}$ is normalized. $\tilde{\mathbb{E}}[1] = 1$

Sum of Squares Algorithm



$\tilde{\mu}$:a prob. dist. on ω-cliques.

$$\tilde{\mathbb{E}} = \tilde{\mathbb{E}}_{\tilde{\mu}}$$

Total degree of argument of $\tilde{\mathbb{E}}$ is at most d!

$\tilde{\mathbb{E}}$ is linear. **PseudoMoments** $\tilde{\mathbb{E}}[\Pi_{i \in S} x_i]$

Computed using SDP on $n^{O(d)}$ variables and constraints!

$= 0$

$\tilde{\mu}$ sa

n G

$$\tilde{\mathbb{E}}[q(x)\left(x_i^2 - x_i\right)] = 0 \quad \forall i$$

$\tilde{\mu}$ is positive semidefinite. $\tilde{\mathbb{E}}[q(x)^2] \geq 0$

$\tilde{\mu}$ is normalized. $\tilde{\mathbb{E}}[1] = 1$

Sum of Squares Proofs

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

Duality

No $\tilde{\mathbb{E}}$ *satisfying constraints*

Proof of unsatisfiability in a specific form.

Sum of Squares Proofs

Is this set of poly-equations feasible?

$$x_i x_j = 0 \;\; \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \;\; \forall\, i$$

$$\sum_i x_i = \omega$$

Duality

No $\tilde{\mathbb{E}}$ *satisfying constraints*

SoS Refutation

There exist deg d polynomials f and g

1. f = 0 on solutions.
2. g is sum of squares of deg d/2 polynomials, and

-1 = f + g

Proving Lower Bounds

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

Goal

For $d = \Theta(1)$ and $\omega = n^{0.5 - o(1)}$, show that there is a pseudo-expectation satisfying all the constraints, w.h.p. when $G \sim G(n, \frac{1}{2})$.

SoS "thinks" there's a large clique in a random graph!

Talk Plan

1. Sum of Squares Proofs and Pseudodistributions

2. Meka-Wigderson Pseudodistribution

3. MW Pseudodistribution is not PSD

4. The Bayesian Pseudodistribution

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

"fake" planted clique

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$
$$x_i^2 = x_i \quad \forall\, i$$
$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

Every vertex is equally likely to be in C.

$$\tilde{\mathbb{E}}[x_i] \approx \left(\frac{\omega}{n}\right)$$

 sloppy calculations only.

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

Every vertex is equally likely to be in C.

$$\tilde{\mathbb{E}}[x_i] \approx \left(\frac{\omega}{n}\right)$$

Every edge is equally likely to be in C.

$$\tilde{\mathbb{E}}[x_i x_j] \approx \begin{cases} \left(\frac{\omega}{n}\right)^2 & \text{if } i \sim j \\ 0 & \text{otherwise.} \end{cases}$$

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

$$\tilde{\mathbb{E}}[x_S] \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if } S \text{ is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

$$\tilde{\mathbb{E}}[x_S] = \begin{cases} \Theta(1)\left(\frac{\omega}{n}\right)^{|S|}\left(1 + O(n^{-|S|})\right) & \text{if S is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

The MW Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

$$\tilde{\mathbb{E}}[x_S] \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if } S \text{ is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

Talk Plan

1. Sum of Squares Proofs and Pseudodistributions

2. Meka-Wigderson Pseudodistribution

3. MW Pseudodistribution is not PSD

4. The Bayesian Pseudodistribution

MW Moments are not PSD...

Theorem **[Kelner]**

There is a degree 2 polynomial p $= p_G$ such that

for $\omega \gg n^{\frac{1}{3}}$, w.h.p. over G $\tilde{\mathbb{E}}_{MW}[p(x)^2] < 0$

generalizes to higher d.

MW Moments are not PSD…

Theorem **[Kelner]**

There is a degree 2 polynomial $p = p_G$ such that for $\omega \gg n^{\frac{1}{3}}$, w.h.p. over G $\tilde{\mathbb{E}}_{MW}[p(x)^2] < 0$

1. MW moments are "*fishy*" w.r.t Bayesian reasoning

Intuition

MW Moments are not PSD…

Theorem **[Kelner]**

There is a degree 2 polynomial p $= p_G$ such that for $\omega \gg n^{\frac{1}{3}}$, w.h.p. over G $\tilde{\mathbb{E}}_{MW}[p(x)^2] < 0$

1. MW moments are "
   Bayesian reasoning

Intuition

2. $\exists q$ s.t. , deg(q) = 4,

$q(x) \geq \omega^5$ if x indicates an ω-clique,

$\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

MW Moments are not PSD…

Theorem [Kelner]

There is a degree 2 polynomial p $= p_G$ such that for $\omega \gg n^{\frac{1}{3}}$, w.h.p. over G $\tilde{\mathbb{E}}_{MW}[p(x)^2] < 0$

1. MW moments are "

Bayesian reasoning

Intuition

2. $\exists q$ s.t. , deg(q) = 4,

$q(x) \geq \omega^5$ if x indicates an ω-clique,

$\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$$\boxed{n\omega^2 \ll \omega^5 \quad \text{when} \quad \omega \gg n^{\frac{1}{3}}}$$

Going Bayesian

Intuition

$G \sim G(n, \frac{1}{2})$ and ***H:*** $(H, C) \sim G(n, \frac{1}{2}, \omega)$

"indistinguishable" for efficient algos.

$(H, C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

Going Bayesian

Intuition

$G \sim G(n, \frac{1}{2})$ and *H:* $(H, C) \sim G(n, \frac{1}{2}, \omega)$

"indistinguishable" for efficient algos.

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

$(H, C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

Going Bayesian

Intuition

$G \sim G(n, \frac{1}{2})$ and *H:* $(H, C) \sim G(n, \frac{1}{2}, \omega)$

"indistinguishable" for efficient algos.

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1 + 2\epsilon)$

$(H, C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

Going Bayesian

Intuition

$G \sim G(n, \frac{1}{2})$ and *H:* $(H, C) \sim G(n, \frac{1}{2}, \omega)$

"indistinguishable" for efficient algos.

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1 + 2\epsilon)$

2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1 + O(\frac{1}{n}))$

$(H, C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

MW: The Simple Moments

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_S]$ (pseudo) prob. of $S \subseteq C$

$$\tilde{\mathbb{E}}[x_S] \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if } S \text{ is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

$$\tilde{\mathbb{E}}[x_S] = \begin{cases} \Theta(1)\left(\frac{\omega}{n}\right)^{|S|}(1 + O(n^{-|S|})) & \text{if S is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

Going Bayesian

Intuition

$G \sim G(n, \frac{1}{2})$ and *H:* $(H,C) \sim G(n, \frac{1}{2}, \omega)$

"indistinguishable" for efficient algos.

Let $\omega = \varepsilon\sqrt{n}$, deg(1) =

large when ω,
approaches $\sqrt{n}$.

D

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1 + 2\epsilon)$

2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1 + O(\frac{1}{n}))$

$(H,C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

Going Bayesian

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1+2\epsilon)$

2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1+O(\frac{1}{n}))$

Proof Idea

$(H,C) \sim G(n,\frac{1}{2},\omega)$:random graph + planted ω-clique.

"prior" $\Pr[1 \in C] = \frac{\omega}{n}$ **deg(1) =** $\frac{n}{2} + \Delta$

"posterior" $\Pr[1 \in C \mid D]$ **?**

Going Bayesian

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1 + 2\epsilon)$
2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1 + O(\frac{1}{n}))$

Proof Idea

$(H, C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

"prior" $\Pr[1 \in C] = \frac{\omega}{n}$ **deg(1) =** $\frac{n}{2} + \Delta$

"posterior"

$\Pr[1 \in C \mid D]$ **?**

If $1 \notin C$, deg(1) $\sim \mathcal{N}(\frac{n}{2}, \frac{n}{4})$

If $1 \in C$, deg(1) $\sim \mathcal{N}(\frac{n}{2} + \omega, \frac{n}{4})$

"higher deg vertex more likely to be in the clique"

Going Bayesian

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1+2\epsilon)$

2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1+O(\frac{1}{n}))$

Proof Idea

$(H,C) \sim G(n, \frac{1}{2}, \omega)$:random graph + planted ω-clique.

"prior" $\Pr[1 \in C] = \frac{\omega}{n}$ **deg(1) =** $\frac{n}{2} + \Delta$

"posterior"

$$\Pr[1 \in C \mid D] = \frac{\omega}{n} \frac{\exp\left(\frac{2(\Delta-\omega)^2}{n}\right)}{\exp\left(\frac{2\Delta^2}{n}\right)} \approx \frac{\omega}{n} \exp\left(\frac{2\Delta\omega}{n}\right) \approx \frac{\omega}{n}(1+\frac{2\Delta\omega}{n})$$

Going Bayesian

Let $\omega = \varepsilon\sqrt{n}$, deg(1) = $\frac{n}{2} + \sqrt{n}$ **Event D**

Claim

1. $\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1+2\epsilon)$

2. $\mathbb{E}_G[\tilde{\mathbb{E}}[x_1] \mid D] \approx \frac{\omega}{n}(1+O(\frac{1}{n}))$

Proof Idea

$(H,C) \sim G(n,\frac{1}{2},\omega)$:random graph + planted ω-clique.

"prior" $\Pr[1 \in C] = \frac{\omega}{n}$ **deg(1) =** $\frac{n}{2} + \Delta$

$$\Pr[1 \in C \mid D] = \frac{\omega}{n}\frac{\exp\left(\frac{2(\Delta-\omega)^2}{n}\right)}{\exp\left(\frac{2\Delta^2}{n}\right)} \approx \frac{\omega}{n}\exp\left(\frac{2\Delta\omega}{n}\right) \approx \frac{\omega}{n}(1+\frac{2\Delta\omega}{n})$$

$$\mathbb{E}_{(H,C)}[x_1 \mid D] \approx \frac{\omega}{n}(1+2\epsilon)$$

Going Bayesian

MW moments are "*fishy*" w.r.t Bayesian reasoning.

Going Bayesian

MW moments are "*fishy*" w.r.t Bayesian reasoning.

Can

notice?

Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4$$

Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4$$

Proof Idea

x indicates an ω-clique

for any i: $x_i = 1$ $\left(\sum_{j \leq n} r_i(j) x_j \right)^4 \gtrapprox \omega^4$

There are ω such i.

Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

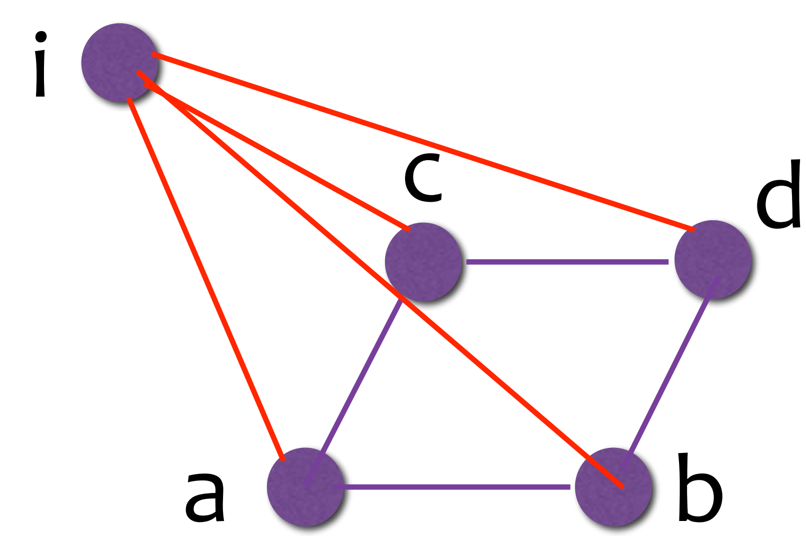
$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4 \qquad \tilde{\mathbb{E}}_{MW} \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if S is a clique.} \\ 0 & \text{otherwise.} \end{cases}$$

Proof Idea

$$\mathbb{E}_G \tilde{\mathbb{E}}[\left(\sum_{j \leq n} r_i(j) x_j \right)^4] = \sum_{a,b,c,d} \mathbb{E}_G \left[r_i(a) r_i(b) r_i(c) r_i(d) \tilde{\mathbb{E}}[x_a x_b x_c x_d] \right]$$

notice

$i \notin \{a, b, c, d\}$



Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. $G \sim G(n, \frac{1}{2})$.

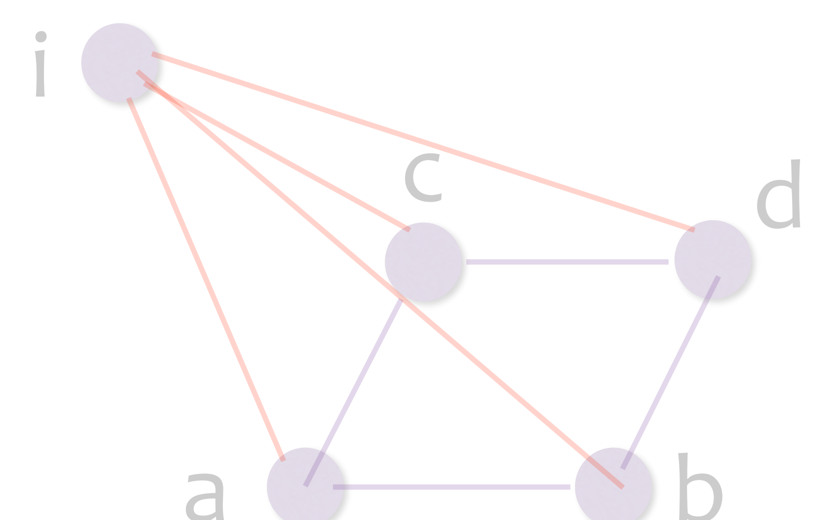
$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4 \qquad \tilde{\mathbb{E}}_{MW} \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if S is a clique.} \\ 0 & \text{otherwise.} \end{cases}$$

Proof Idea

$$\mathbb{E}_G \tilde{\mathbb{E}}[\left(\sum_{j \leq n} r_i(j) x_j \right)^4] = \sum_{a,b,c,d} \mathbb{E}_G \left[r_i(a) r_i(b) r_i(c) r_i(d) \tilde{\mathbb{E}}[x_a x_b x_c x_d] \right]$$

suppose

$i \notin \{a, b, c, d\}$



Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4 \qquad \tilde{\mathbb{E}}_{MW} \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if S is a clique.} \\ 0 & \text{otherwise.} \end{cases}$$

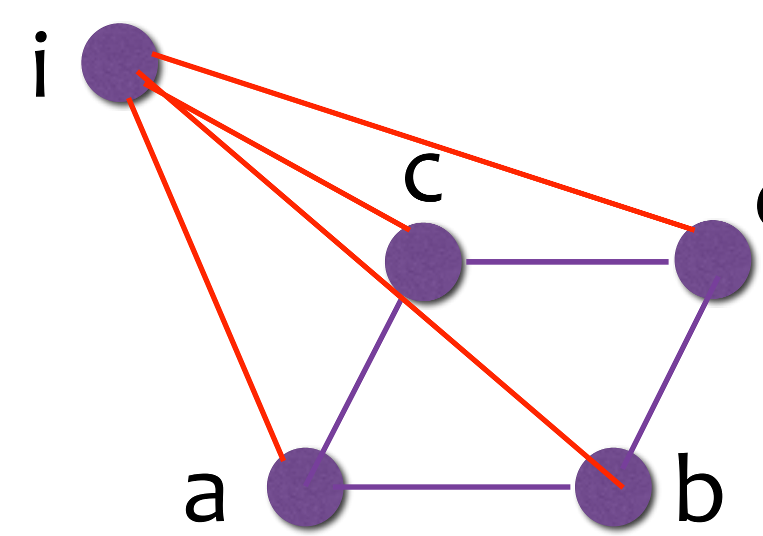
Proof Idea

$$\mathbb{E}_G \tilde{\mathbb{E}}[\left(\sum_{j \leq n} r_i(j) x_j \right)^4] = \sum_{a,b,c,d} \mathbb{E}_G \left[r_i(a) r_i(b) r_i(c) r_i(d) \tilde{\mathbb{E}}[x_a x_b x_c x_d] \right]$$

$$\approx \sum_{a,b,c,d} \mathbb{E}_G[r_i(a) r_i(b) r_i(c) r_i(d)] \cdot \tilde{\mathbb{E}}[x_a x_b x_c x_d]$$

Notice

$i \notin \{a, b, c, d\}$



Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases} \qquad q(x) = \sum_{i \leq n} \left(\sum_{j \leq n} r_i(j) x_j \right)^4 \qquad \tilde{\mathbb{E}}_{MW} \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if S is a clique.} \\ 0 & \text{otherwise.} \end{cases}$$

Proof Idea

$$\mathbb{E}_G \tilde{\mathbb{E}}[\left(\sum_{j \leq n} r_i(j) x_j \right)^4] = \sum_{a,b,c,d} \mathbb{E}_G \left[r_i(a) r_i(b) r_i(c) r_i(d) \tilde{\mathbb{E}}[x_a x_b x_c x_d] \right]$$

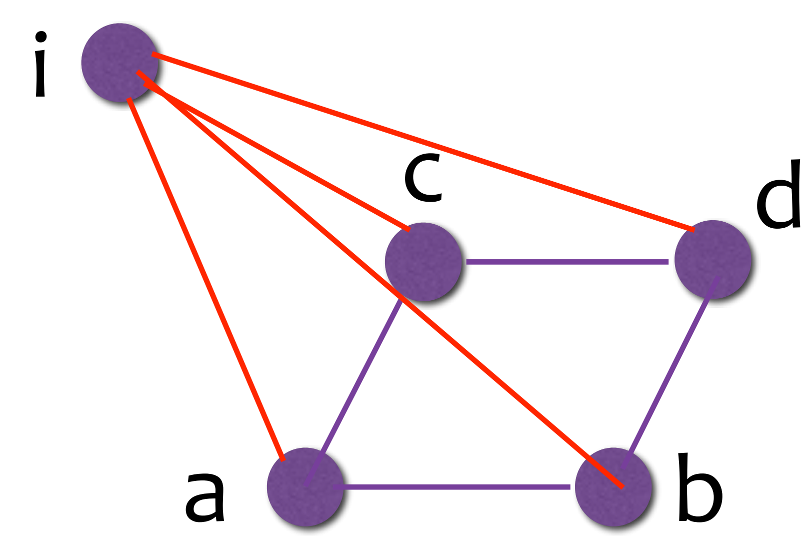
$$\approx \sum_{a,b,c,d} \mathbb{E}_G[r_i(a) r_i(b) r_i(c) r_i(d)] \cdot \tilde{\mathbb{E}}[x_a x_b x_c x_d]$$

$$\approx n^2 \cdot \frac{\omega^2}{n^2} = \omega^2$$

Notice

$i \notin \{a, b, c, d\}$

Use Markov.



Bayesian Reasoning

$\exists q$ s.t., deg(q) = 4, $q(x) \geq \omega^5$ if x indicates an ω-clique,

but $\tilde{\mathbb{E}}_{MW}[q(x)] \approx n\omega^2$ w.h.p. over $G \sim G(n, \frac{1}{2})$.

$\boxed{n\omega^2 \ll \omega^5 \quad \text{when} \quad \omega \gg n^{\frac{1}{3}}}$

Fixing MW at degree 4

Is this set of poly-equations feasible?

$$x_i x_j = 0 \;\; \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \;\; \forall\, i$$

$$\sum_i x_i = \omega$$

$$\tilde{\mathbb{E}}[x_S] \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if } S \text{ is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

Fixing MW at degree 4

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$$\tilde{\mathbb{E}}[x_S] \approx \begin{cases} \left(\frac{\omega}{n}\right)^{|S|} & \text{if } S \text{ is a clique} \\ 0 & \text{otherwise.} \end{cases}$$

$$\tilde{\mathbb{E}}[x_S] \approx \left(\frac{\omega}{n}\right)^{|S|} \left(1 + r_S \frac{\omega}{n}\right)$$

where $r_S = \sum_i \Pi_{j \in S} r_i(j)$

Fixing MW at degree 4

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$$\tilde{\mathbb{E}}[x_S] \approx \left(\frac{\omega}{n}\right)^{|S|} \left(1 + r_S \frac{\omega}{n}\right)$$

where $r_S = \sum_i \Pi_{j \in S} r_i(j)$

[HKP'15]
([RS'15])

q is the ***only*** "bad" polynomial for MW.

fixing q doesn't cause ***new*** problems.

Fixing MW at degree 4

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$$\tilde{\mathbb{E}}[x_S] \approx \left(\frac{\omega}{n}\right)^{|S|} \left(1 + r_S \frac{\omega}{n}\right)$$

where $r_S = \sum_i \Pi_{j \in S} r_i(j)$

[HKP'15]
([RS'15])

q is the ***only*** "bad" polynomial for MW.

fixing q doesn't cause ***new*** problems.

approximate eigen-split of MW Mom. matrix

norm of random matrices with limited dependence

Fixing MW at degree 4

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

$$\tilde{\mathbb{E}}[x_S] \approx \left(\frac{\omega}{n}\right)^{|S|} \left(1 + r_S \frac{\omega}{n}\right)$$

where $r_S = \sum_i \Pi_{j \in S} r_i(j)$

[HKP'15]
([RS'15])

q is the ***only*** "bad" polynomial for MW.

fixing q doesn't cause ***new*** problems.

doesn't even generalize to d = 6.

Talk Plan

1. Sum of Squares Proofs and Pseudodistributions

2. Meka-Wigderson Pseudodistribution

3. MW Pseudodistribution is not PSD

4. The Bayesian Pseudodistribution

Bayesian Reasoning

SoS seems to “reason” about deg(i).

Bayesian Reasoning

SoS seems to "reason" about deg(i).

1. top(G) :highest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \quad \tilde{\mathbb{E}}[x_{top(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)}]$$

$(H,C) \sim G(n,\frac{1}{2},\omega)$:random graph + planted ω-clique.

Bayesian Reasoning

SoS seems to "reason" about deg(i).

1. top(G) :highest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \quad \tilde{\mathbb{E}}[x_{top(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)}]$$

2. bot(G) :lowest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[x_{top(G)} x_{bot(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)} 1(C)_{bot(H)}]$$

$(H,C) \sim G(n,\frac{1}{2},\omega)$:random graph + planted ω-clique.

Bayesian Reasoning

SoS seems to "reason" about deg(i).

1. top(G) :highest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \quad \tilde{\mathbb{E}}[x_{top(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)}]$$

2. bot(G) :lowest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[x_{top(G)} x_{bot(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)} 1(C)_{bot(H)}]$$

3. top-edge(G) :edge contained in most triangles in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[x_{top-edge(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top-edge(H)}]$$

Bayesian Reasoning

SoS seems to "reason" about deg(i).

1. top(G) :highest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \quad \tilde{\mathbb{E}}[x_{top(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)}]$$

2. bot(G) :lowest degree vertex in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[x_{top(G)} x_{bot(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top(H)} 1(C)_{bot(H)}]$$

3. top-edge(G) :edge contained in most triangles in G

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[x_{top-edge(G)}] \approx \mathbb{E}_{(H,C)}[1(C)_{top-edge(H)}]$$

• • •

The Bayesian Commandment

Pseudoexpectation shall satisfy all simple Bayesian tests

The Bayesian Commandment

For every ***"simple"*** map $\{0,1\}^{\binom{[n]}{2}} \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[p_G(x)] \approx \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_H(1(C))]$$

$G = \{g_{\{i,j\}}\}_{i,j\in[n]}$ $\quad g_{\{i,j\}} = \begin{cases} +1 & \text{if } \{i,j\} \text{ is an edge in G.} \\ -1 & \text{otherwise.} \end{cases}$

p_G is a polynomial in x and $\{g_{\{i,j\}}\}_{i,j\in[n]}$

The Bayesian Commandment

For every ***"simple"*** map $\{0,1\}^{\binom{[n]}{2}} \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}\tilde{\mathbb{E}}[p_G(x)] \approx \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_H(1(C))]$$

$G = \{g_{\{i,j\}}\}_{i,j\in[n]}$ $\quad g_{\{i,j\}} = \begin{cases} +1 & \text{if } \{i,j\} \text{ is an edge in G.} \\ -1 & \text{otherwise.} \end{cases}$

p_G is a polynomial in x and $\{g_{\{i,j\}}\}_{i,j\in[n]}$

Solution Degree degree of p_G in x.

Instance Degree degree of p_G in $\{g_{\{i,j\}}\}_{i,j\in[n]}$.

The Bayesian Commandment

For every **"*simple*"** map $\{0,1\}^{\binom{[n]}{2}} \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}\tilde{\mathbb{E}}[p_G(x)] \approx \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_H(1(C))]$$

$$G = \{g_{\{i,j\}}\}_{i,j\in[n]} \qquad g_{\{i,j\}} = \begin{cases} +1 & \text{if } \{i,j\} \text{ is an edge in G.} \\ -1 & \text{otherwise.} \end{cases}$$

p_G is a polynomial in x and $\{g_{\{i,j\}}\}_{i,j\in[n]}$

Solution Degree degree of p_G in x.

Instance Degree degree of p_G in $\{g_{\{i,j\}}\}_{i,j\in[n]}$.

"*simple*" map ⟷ low ins. and sol. degree.

The Bayesian Commandment

For every ***low degree*** map $\{0,1\}^{\binom{[n]}{2}} \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}\tilde{\mathbb{E}}[p_G(x)] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_H(1(C))]$$

$$\tilde{\mathbb{E}}[x_S] : \{0,1\}^{\binom{n}{2}} \to \mathbb{R} \qquad \tilde{\mathbb{E}}[x_S] = \sum_{T\subseteq\binom{n}{2}} \widehat{\tilde{\mathbb{E}}[x_S]}(T)\cdot\chi_T$$

Observation

Bayesian Commandment

$\Downarrow$

Low Degree Fourier coefficients of $\tilde{\mathbb{E}}[x_S]$ determined!

The Bayesian Commandment

For every ***low degree*** map $\{0,1\}^{\binom{[n]}{2}} \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}\tilde{\mathbb{E}}[p_G(x)] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_H(1(C))]$$

$$\tilde{\mathbb{E}}[x_S] : \{0,1\}^{\binom{n}{2}} \to \mathbb{R} \qquad \tilde{\mathbb{E}}[x_S] = \sum_{T\subseteq\binom{n}{2}} \widehat{\tilde{\mathbb{E}}[x_S]}(T)\cdot\chi_T$$

Bayesian Commandment

Low Degree Fourier coefficients of $\tilde{\mathbb{E}}[x_S]$ determined!

Observation

No choice in the operator!

The Bayesian Pseudomoments

$$\tilde{\mathbb{E}}[x_S] = \sum_{T \subseteq \binom{[n]}{2} : |V(T)| \leq \tau} \widehat{\tilde{\mathbb{E}}[x_S]}(T) \chi_T$$

use a ***truncated*** Fourier polynomial as the pseudomoment.

Easy

$\tilde{\mathbb{E}}$ satisfies all constraints!

The Bayesian Pseudomoments

$$\tilde{\mathbb{E}}[x_S] = \sum_{T \subseteq \binom{[n]}{2} : |V(T)| \le \tau} \widehat{\tilde{\mathbb{E}}[x_S]}(T) \chi_T$$

use a ***truncated*** Fourier polynomial as the pseudomoment.

Easy

$\tilde{\mathbb{E}}$ satisfies all constraints!

Main Technical Lemma

For every q, deg(q) at most d/2

$\tilde{\mathbb{E}}[q(x)^2] \geq 0$ w.h.p. over $G \sim G(n, \frac{1}{2})$

The Bayesian Pseudomoments

Main Technical Lemma

For every q, deg(q) at most d/2

$\tilde{\mathbb{E}}[q(x)^2] \geq 0$ w.h.p. over $G \sim G(n, \frac{1}{2})$

$M(I, J) = \tilde{\mathbb{E}}[x_{I \cup J}]$

Moment Matrix

$M \approx M_{struct} + M_{random}$

M_{struct} eigenvectors are low-degree functions of graph

M_{random} eigenvectors are high-degree functions of graph

The Bayesian Pseudomoments

Main Technical Lemma

For every q, deg(q) at most d/2

$\tilde{\mathbb{E}}[q(x)^2] \geq 0$ w.h.p. over $G \sim G(n, \frac{1}{2})$

$M(I, J) = \tilde{\mathbb{E}}[x_{I \cup J}]$

Moment Matrix

$M \approx M_{struct} + M_{random}$

M_{struct} eigenvectors are low-degree functions of graph

fine by construction of the operator!

M_{random} eigenvectors are high-degree functions of graph

"random-like" -Wigner-like in appropriate subspaces

Main Theorem

For every d, Bayesian $\tilde{\mathbb{E}}$ satisfies all the constraints whenever $\omega \ll n^{0.5-\varepsilon}$.

$d = \Theta(1)$, $\epsilon = o_d(1)$

Conclusions

General procedure to construct SoS LB witness.

unify SoS LBs, prove LBs for planted problems, 2->4 norm…

Bayesian view of SoS

Applications to upper bounds?

Thank you for your attention.

Acknowledgment

Some illustrations based on Boaz's inputs.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall \ i \not\sim j \text{ in } G$$

$$x_i^2 = x_i \ \forall \ i$$

$$\sum_i x_i = \omega$$

$\tilde{\mathbb{E}}[x_i] \approx \left(\frac{\omega}{n}\right)$ Vertex i : deg(i) = $\frac{n}{2} + \Delta$, $\Delta \pm \Theta(\sqrt{n})$

If $i \notin C$, deg(i) $\sim \mathcal{N}(\frac{n}{2}, \frac{n}{4})$ **If $i \in C$,** deg(i) $\sim \mathcal{N}(\frac{n}{2} + \omega, \frac{n}{4})$

Given deg(i), posterior "probability" of i is in C?

$$\Pr[i \in C] \cdot \frac{\Pr[deg(i) = q \mid i \in C]}{\Pr[i \in C]\Pr[deg(i) = q \mid i \in C] + \Pr[i \notin C]\Pr[deg(i) = q \mid i \notin C]}$$

sloppy calculations only.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall \ i \not\sim j \text{ in } G$$
$$x_i^2 = x_i \ \forall \ i$$

$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

$$\mathbb{E}_{G \sim G(n, \frac{1}{2})} \tilde{\mathbb{E}}[\langle r_i, x \rangle^4] \approx \|x\|^4$$
$$\approx \omega^2$$

r_i is treated as a random vector.

⚠ sloppy calculations only.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall\, i \not\sim j \text{ in } G \qquad \sum_i x_i = \omega$$
$$x_i^2 = x_i \ \forall\, i$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}\tilde{\mathbb{E}}[\langle r_i, x\rangle^4] \approx \|x\|^4 \approx \omega^2 \qquad \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}\langle r_i, 1(C)\rangle^4] \approx \frac{\omega}{n}\omega^4 \approx \frac{\omega^5}{n}$$

sloppy calculations only.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall\, i \not\sim j \text{ in } G$$

$$x_i^2 = x_i \ \forall\, i$$

$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

prob i is in C

$$\mathbb{E}_{G \sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[\langle r_i, x\rangle^4] \approx \|x\|^4 \approx \omega^2$$

$$\mathbb{E}_{(H,C) \sim G(n,\frac{1}{2},\omega)} \langle r_i, 1(C)\rangle^4] \approx \frac{\omega}{n}\omega^4 \approx \frac{\omega^5}{n}$$

⚠ sloppy calculations only.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \;\forall\, i \not\sim j \text{ in } G$$
$$x_i^2 = x_i \;\forall\, i$$
$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

$$\mathbb{E}_{G \sim G(n,\frac{1}{2})} \tilde{\mathbb{E}}[\langle r_i, x\rangle^4] \approx \|x\|^4 \approx \omega^2$$

$$\mathbb{E}_{(H,C) \sim G(n,\frac{1}{2},\omega)} \langle r_i, 1(C)\rangle^4] \approx \frac{\omega}{n}\omega^4 \approx \frac{\omega^5}{n}$$

If $\omega \gg n^{\frac{1}{3}}$ pseudo expectation is too small.

⚠ sloppy calculations only.

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall \ i \not\sim j \text{ in } G$$

$$x_i^2 = x_i \ \forall \ i$$

$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

SoS captures "low-instance degree" Bayesian reasoning?

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall \ i \not\sim j \text{ in } G \qquad \sum_i x_i = \omega$$

$$x_i^2 = x_i \ \forall \ i$$

MW violates Bayesian reasoning.

"instance degree"of $\sum_S p_S x_S$ = deg of p_S edge indicators

SoS captures "low-instance degree" Bayesian reasoning?

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \ \forall\ i \not\sim j \text{ in } G \qquad \sum_i x_i = \omega$$
$$x_i^2 = x_i \ \forall\ i$$

MW violates Bayesian reasoning.

"instance degree"of $\sum_S p_S x_S$ = deg of p_S edge indicators

SoS captures "low-instance degree" Bayesian reasoning.

degree of vertices, number of triangles, etc.

Bayesian Reasoning

Is this set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

[Kelner]

for $\omega \gg n^{1/3}$ and some constant c,

w.h.p. $\tilde{\mathbb{E}}\left[\left(\langle r_i, x\rangle^2 - c x_i\right)^2\right] < 0$

Bayesian Reasoning

Is the set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

MW violates Bayesian reasoning.

$$r_i(j) = \begin{cases} +1 & \text{if } i \sim j \\ 0 & \text{if } j = i \\ -1 & \text{if } i \not\sim j \end{cases}$$

[Kelner]

for $\omega \gg n^{1/3}$ and some constant c,

w.h.p. $\tilde{\mathbb{E}}\left[\left(\langle r_i, x\rangle^2 - c x_i\right)^2\right] < 0$

MW Pseudomoments work for $\omega \ll n^{\frac{1}{\frac{d}{2}+1}}$.

[MPW'15]
[DM'15]
[HKP'15]

The Bayesian Pseudomoments

For every ***"simple"*** map $G \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[p_G(x)]] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_G(1(C))]$$

$$\tilde{\mathbb{E}}[x_S] : \{0,1\}^{\binom{n}{2}} \to \mathbb{R} \qquad \sum_{T\subseteq\binom{n}{2}} \widehat{\tilde{\mathbb{E}}[x_S]}(T) \cdot \chi_T$$

$$\widehat{\tilde{\mathbb{E}}[x_S]}(T) = \mathbb{E}_{G\sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[x_S]\chi_T(G)]$$

The Bayesian Pseudomoments

$$\tilde{\mathbb{E}}[x_S] : \{0,1\}^{\binom{n}{2}} \to \mathbb{R} \qquad \sum_{T \subseteq \binom{n}{2}} \widehat{\tilde{\mathbb{E}}[x_S]}(T) \cdot \chi_T$$

$$\widehat{\tilde{\mathbb{E}}[x_S]}(T) = \mathbb{E}_{G \sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[x_S]\chi_T(G)]$$

$$= \mathbb{E}_{(H,C) \sim G(n,\frac{1}{2},\omega)}[1(S \subseteq C)\chi_T(H)]$$

match “moments” (Fourier coeffs) with planted dist.

The Bayesian Pseudomoments

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[p_G(x)]] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_G(1(C))]$$

$$\widehat{\tilde{\mathbb{E}}[x_S]}(T) = \mathbb{E}_{G\sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[x_S]\chi_T(G)]$$

$$= \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[1(S\subseteq C)\chi_T(H)]$$

completely determined by the "planted" distribution!

The Bayesian Pseudomoments

Is the set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

build an operator that satisfies ***all*** Bayesian constraints of ***low-instance degree*****.**

"instance degree"of $\sum_S p_S x_S$ = deg of p_S edge indicators

The Bayesian Pseudomoments

Is the set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

build an operator that satisfies ***all*** Bayesian constraints of ***low-instance degree*****.**

For every **"simple"** map $G \to p_G(x)$

$$\mathbb{E}_{G\sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[p_G(x)]] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_G(1(C))]$$

The Bayesian Pseudomoments

Is the set of poly-equations feasible?

$x_i x_j = 0 \ \forall\ i \not\sim j$ in G

$x_i^2 = x_i \ \forall\ i$

$\sum_i x_i = \omega$

build an operator that satisfies

of ***low-instance degree***

No choice in the operator if want to satisfy all Bayesian constraints!

For every **"simple"** map $G \to p_G(x)$

$$\mathbb{E}_{G \sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[p_G(x)]] = \mathbb{E}_{(H,C) \sim G(n,\frac{1}{2},\omega)}[p_G(1(C))]$$

The Bayesian Pseudomoments

Is the set of poly-equations feasible?

$$x_i x_j = 0 \quad \forall\, i \not\sim j \text{ in G}$$

$$x_i^2 = x_i \quad \forall\, i$$

$$\sum_i x_i = \omega$$

build an operator that satisfies ***all*** Bayesian constraints of ***low-instance degree.***

For every **"simple"** map $G \to p_G(x)$

$$\mathbb{E}_{G \sim G(n,\frac{1}{2})}[\tilde{\mathbb{E}}[p_G(x)]] = \mathbb{E}_{(H,C)\sim G(n,\frac{1}{2},\omega)}[p_G(1(C))]$$

The Bayesian Pseudomoments

$$\tilde{\mathbb{E}}[x_S] : \{0,1\}^{\binom{n}{2}} \to \mathbb{R} \qquad \sum_{T \subseteq \binom{[n]}{2} : |V(T)| \leq \tau} \widehat{\tilde{\mathbb{E}}[x_S]}(T)\chi_T$$

use a ***truncated*** Fourier polynomial as the pseudomoment.