

(Lecture 6: Global Correlation Rounding)

Gaussian Rounding : (p.d. degree 2)

G.R + Squared triangle inequalities (p.d. of deg 4)

today: one-way to use higher degree SOS.

Running application: Max-cut problem on certain restricted class of graphs

Recalling basic Spectral Graph Theory

G : d -regular graph on $[n]$ vertices

Normalized Adj Matrix : $A(i,j) = \begin{cases} 1/d & \text{if } \{i,j\} \in E \\ 0 & \text{o/w} \end{cases}$

$\equiv$ Random Walk matrix of G

Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigs of A

Obsⁿ 1: $\lambda_1 = 1$ (all 1s vector is the corr eig-vector)

Obsⁿ 2: $\lambda_2 = 1$ iff G is disconnected.

G is disconnected: C_1 & C_2 are conn components

1, vector c_2

$$v: \langle v, \mathbb{1} \rangle = 0.$$

if $Av = v$ then $v_i = v_j$ if $i \rightsquigarrow j$
the only way $v \neq c \cdot \mathbb{1}$ is if G is disconnected.

G is expander : $\Phi_G \geq \delta$

$$\lambda_2 \leq 1 - O(\delta^2) \quad \swarrow$$

G is a random d -reg graph:

$$\lambda_2(A) = \frac{2\sqrt{d-1}}{d} + o(1)$$

Def (Threshold Rank)

Let G be a d -reg graph. $\lambda_1 \gg \dots \gg \lambda_n$

The p -threshold rank of G $\text{rank}_p(G)$

$$= \# \text{ eigs of } G \geq p.$$

Ex: Suppose G is an expander graph.

$$p: 1 - o(1)$$

$$\text{What is } \text{rank}_p(G) = 1$$

is a linear combination of

Bounded threshold rank $\equiv$ "generate threshold expanders"

Running example / main Goal.

Theorem: Let G be a d -reg graph on $[n]$.

$$\text{Let } f_G = \frac{1}{4} \sum_{\{i,j\} \in E} (x_i - x_j)^2$$

Let $r = \text{rank}_{\delta^2}(G) / \delta^4$. Then given a d -reg $r+2$ p.d. μ , can find a $\text{dist}^n x'$

$$\text{s.t. } \mathbb{E} f_G(x') \geq \underbrace{\mathbb{E}_{\mu} f_G} - O(\delta) \cdot m$$

" $n^{O(r)}$ time algo to compute additive approx."

$\equiv n^{O(r)}$ time algo to compute $(1 - O(\delta))$ multiplicative approx. to max-cut.

Global Correlation Rounding

We'll first study some basic props of higher degree pseudo-distⁿs.

set $S \subseteq [n]^r$

Def "Marginal of a p.d. on a set $S \subseteq [n]$ of degree r ."

Let μ be a p.d. on $\{\pm 1\}^n$. For any

$S \subseteq [n]$, $\mu|_S : \{\pm 1\}^S \rightarrow \mathbb{R}$ is the

following f^n : $\mu|_S(y) = \sum_{x: x|_S = y} \mu(x)$.

Remark: $\forall U \subseteq S$, $\mathbb{E}_{\mu|_S} X_U = \mathbb{E}_{\mu} X_U$.

Thus, $\forall f$ of degree r , $\mu|_S$ only depends on variables in S ,

$$\mathbb{E}_{\mu|_S} f = \mathbb{E}_{\mu} f.$$

Prop: Let μ be a p.d. on $\{\pm 1\}^n$ of degree r .

Let $S \subseteq [n]$ of size at most $r/2$ ($|S| \leq r/2$).

Then $\mu|_S$ is an actual prob. distⁿ

Proof: μ is a p.d. of degree r

$\Downarrow$
 $\mu|_S$ is also a p.d. of degree r

[Note all expectations of $f: \{\pm 1\}^S \rightarrow \mathbb{R}$]

Since $|S| \leq \frac{r}{2}$: using that $z, 2r$ deg p.d. are the same)
on $\xi \in \{1\}^S$ are actual distⁿ $\rightarrow$ D.

Def / Notation / Name: $\forall |S| \leq \frac{r}{2}$, $\mu|_S$ the
 "local distⁿ" of μ on S .

Def (Reweighting):

Let μ be a p.d. of deg r . Suppose p is an
 SoS polynomial of degree $r' \leq r$.
 (Note: $\tilde{\mathbb{E}}_{\mu} p \approx 0$). Assume $\tilde{\mathbb{E}}_{\mu} p > 0$.

Then, μ' is said to be reweighting of μ by
 p if

$$\mu'(x) = \frac{\mu(x) \cdot p(x)}{\tilde{\mathbb{E}}_{\mu} p}$$

Remark:

$$\tilde{\mathbb{E}}_{\mu'} f = \sum_x \mu'(x) \cdot f(x)$$

If μ' has to be
 a p.d.

$$\sum_x \mu'(x) = 1$$

$$\sum_x \mu(x) \cdot p(x) \cdot z = 1$$

$$z \cdot \tilde{\mathbb{E}}_{\mu} p = 1$$

$$\Rightarrow z = \frac{1}{\tilde{\mathbb{E}}_{\mu} p}$$

$$\mu = \sum_x \underbrace{\mu(x) \cdot (p(x) - f(x))}_{\tilde{E}_{\mu} \cdot p}$$

$$= \frac{\tilde{E}_{\mu} \cdot p \cdot f}{\tilde{E}_{\mu} \cdot p}$$

Prop: Suppose μ is a deg r p.d. on $\{\pm 1\}^n$.
 Let p be an sos polynomial of deg $\leq r'$.
 & $\tilde{E}_{\mu} p > 0$. Then the reweighting $\mu' = \frac{\mu(x) \cdot p(x)}{\tilde{E}_{\mu} p}$
 is also a pseudo-distⁿ of deg $\leq r - r'$.

Proof: 1) $\tilde{E}_{\mu'} 1 = 1$?

2) $\tilde{E}_{\mu'} f^2 \geq 0 \quad \forall f: \deg(f) \leq \frac{r-r'}{2}$.

$$\tilde{E}_{\mu'} 1 = \frac{\tilde{E}_{\mu} \cdot p \cdot 1}{\tilde{E}_{\mu} p} = 1 \quad \checkmark$$

$$\tilde{E}_{\mu'} f^2 = \frac{\tilde{E}_{\mu} \cdot p \cdot f^2}{\tilde{E}_{\mu} p} \geq 0$$

μ'
 $\tilde{E}_{\mu, p}$
 Claim: $p \cdot f^2$ is sos of $\deg \leq r$

D

Special, interesting case

Conditioning

Let μ be a p.d. of $\deg r$.

Let $|S| \leq r'/2, S \subseteq [n]$

Let $\alpha \in \{\pm 1\}^S$.

Then $\mu|_{X_S=\alpha} =$ reweighting of μ by the polynomial f_α^2 ✓
 where $f_\alpha(x) = \begin{cases} 1 & \text{if } x_S = \alpha \\ 0 & \text{o/w} \end{cases}$

$\deg(f_\alpha) \leq |S|$

Goal: $\mathbb{E}_{\mu} f_\alpha^2 > 0$ on the event $X_S = \alpha$

↓ assuming
 $\mathbb{E}_{\mu} f_\alpha^2 > 0$

Note: $\mu|_{X_S=\alpha}$ is a p.d. of $\deg \leq r - r'$

random-distⁿs?

how to round high deg pseudo

here's a simple "high-deg" rounding.

Def (Independent Rounding)

Let μ be a p.d. of $\deg = r$.

Let x' be the following actual distⁿ.

Set $x'_i = \begin{cases} 1 & \text{w.p. } \frac{1}{2} + \frac{1}{2} \tilde{\mathbb{E}} x_i \\ -1 & \text{o/w} \end{cases}$ independently for $1 \leq i \leq n$.

$$\mathbb{E} x'_i = \tilde{\mathbb{E}}_{\mu} x_i$$

"round each variable independently"

When is it good?

Prop: Let x' be indep rounding of a p.d.

μ of $\deg \geq 2$. Then.

$$\mathbb{E} f_G(x') = \tilde{\mathbb{E}}_{\mu} f_G(x)$$

$$+ \sqrt{\sum_{i,j} \tilde{\text{Cov}}(x_i, x_j)}$$

Def (Pseudo-covariance):

$$\tilde{E}_{\mu} (x_i - \tilde{E}_{\mu} x_i)(x_j - \tilde{E}_{\mu} x_j) \stackrel{\text{def}}{=} \widetilde{\text{Cov}}(x_i, x_j)$$

Proof: $E(x_i' - x_j')^2$

$$= E x_i'^2 + E x_j'^2 - 2 \cdot E x_i' \cdot x_j'$$

$$= 1 + 1 - 2 \cdot (E x_i')(E x_j')$$

$$\tilde{E}_{\mu} (x_i - x_j)^2 = 2 - 2 \cdot \tilde{E}_{\mu} x_i x_j$$

$$\frac{1}{4} E(x_i' - x_j')^2 = \frac{1}{4} E(x_i' - x_j')^2 + \frac{1}{2} \underbrace{\left(\tilde{E}_{\mu} x_i x_j - E x_i' \cdot E x_j' \right)}$$

$$= \frac{1}{4} \cdot \tilde{E}_{\mu} (x_i - x_j)^2 + \frac{1}{2} \cdot \left(\tilde{E}_{\mu} x_i x_j - \tilde{E}_{\mu} x_i \tilde{E}_{\mu} x_j \right)$$

$$= \frac{1}{4} \cdot \tilde{E}_{\mu} (x_i - x_j)^2 + \frac{1}{2} \widetilde{\text{Cov}}(x_i, x_j)$$

done by linearity.

Main insight: Start from a high deg p.d. μ .
and produce a low average covariance p.d.
by conditioning μ . while ensuring that μ_1

$$\tilde{\mathbb{E}}_{\mu} f_G \approx \mathbb{E}_{\mu_1} f_G$$

Then can independently rand on μ_1 .

$$\text{Low average covariance} = \underbrace{\mathbb{E}_{\{i,j\} \in E} |\tilde{\text{Cov}}(X_i, X_j)|}_{\leq \delta}$$

$$\begin{aligned} \text{Then, } \mathbb{E} f_G(X') &\approx \mathbb{E}_{\mu_1} f_G - \delta \cdot m \\ &\approx \tilde{\mathbb{E}}_{\mu} f_G - \delta \cdot m \end{aligned}$$

$$\mu \rightarrow \mu_2$$

$$\mu_2(x) = \frac{\mu(x) + \mu(-x)}{2}$$

Under Symmetrization, 2nd moments
are preserved, covariances are not

not Covariance

$$\tilde{\mathbb{E}}_{\mu} f_G(X') \approx \mathbb{E}_{\mu_1} f_G(X') = \mathbb{E}_{\mu_1} \cdot XX^T$$

$$\mathbb{E}_{u_2} X = 0, \quad \mathbb{E}_{\mu_2} X X' = \underline{\underline{\mu}}$$

Def (Local Correlation of a pseudodistⁿ)

Let μ be a p.d. on $\{\pm 1\}^n$.

Let G be any graph on $[n]$, $G([n], E)$.

Then, the local correlation of μ under G

$$= \mathbb{E}_{\{i,j\} \in E} |\widetilde{\text{Cov}}(X_i, X_j)|$$

Summary of: If Indep rounding fails to give additive δ approx. then

$$\mathbb{E}_{\{i,j\} \in E} |\widetilde{\text{Cov}}(X_i, X_j)| \geq 2.5\delta$$

Can't expect to find p.d. of low local corr for arbitrary G . Instead we'll define a "global Correlation" & relate G.C. to L.C. for "nice"

Graphs

Basic Information Theory

CSPs with alphabet of size 10

$[q]^n \equiv$ Space of assignments to alphabet of size 10

X : be a r.v. on $[q]$.

$$H(X) \stackrel{\text{def}}{=} - \sum_{i \in [q]} \text{prob}(X=i) \cdot \log(\text{prob}(X=i))$$

$$I(X; Y) = \sum_{i, j \in [q]} \text{Pr}(X=i, Y=j) \log \frac{\text{Pr}(X=i, Y=j)}{\text{Pr}(X=i) \cdot \text{Pr}(Y=j)}$$

(mutual information)

$$H(X|Y) = \sum_{i \in [q]} \text{Pr}(Y=i) H(X|Y=i)$$

(Conditional entropy)

Note: If X & Y are indep. then $H(X|Y) = H(X)$.

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$

Note: $I(X; Y) = 0$ when X & Y are indep.

"measure of how far from indep are X & Y "

Fact: Let X, Y be $[q]$ -valued r.v.s.

Then $|\text{Cov}(X, Y)| \lesssim O_q(\sqrt{\mathbb{I}(X, Y)})$.

Def (Global Correlation) Let μ be a p.d. of $\deg \geq 4$.

$$GC(\mu) \stackrel{\text{def}}{=} \mathbb{E}_{1 \leq i, j \leq n} \underbrace{\mathbb{I}(X_i, X_j)}_{\substack{\text{well-defined}}}.$$

Lemma: Let μ be a p.d. on $\{ \pm 1 \}^n$ of degree

$\frac{2}{\eta} + 2$. Then there's a reweighting μ' of $\deg \geq 2$

with degree $\frac{2}{\eta}$ soS polynomial. s.t. averaging

$$GC(\mu') \leq \eta, \quad \underbrace{\mathbb{E}_{\text{random choices}} \tilde{\mathbb{E}}_{\mu'} f_G}_{\text{random choices}} \approx \underbrace{\mathbb{E}_{\mu} f_G}_{\text{random choices}} = 0$$

Proof: We'll build a sequence of reweightings

$\mu_0 = \mu, \mu_1, \dots, \mu_{t-1} \rightarrow \mu_t$. each μ_t will be a "degree 2" reweighting of μ_{t-1} . & one of them will satisfy the conclusion

1. Sample a uniformly random $1/n$ -type of variables from $[n]$: $i_1, \dots, i_{1/n}$.

For $1 \leq t \leq 1/n$ @ sample $x_{i_t} \sim d_t$ from local distⁿ on i_t under μ_{t-1} .

(b) Set $\mu_t = \mu_{t-1} \cdot \frac{\mathbb{1}(X_{i_t} = d_t)}{\mathbb{E}_{\mu_{t-1}}[\mathbb{1}(X_{i_t} = d_t)]}$ "Condition on $x_{i_t} = d_t$ "

(c) Stop if $GC(\mu_t) < \eta$.

Proof that one of μ_i s works:

Idea: Potential Function: $\mathbb{E}_i H_{\mu_t}(x_i)$

Assume for $t \leq 1/n$, $\mathbb{E}_{\mu_t} \mathbb{I}(X_i; X_j) \geq \eta$.

High G.C.: $1 \leq i, j \leq n$

"typical" i :

$$\mathbb{E}_j \mathbb{I}_{\mu_t}(X_i; X_j) \geq \eta.$$

What happens when I condition on $X_i = d_i$?

$$H(X_j) = H(X_j | X_i)$$

$$H_{\mu_t}(X_j) - H_{\mu_{t-1}}(X_j)$$

$$\mathbb{E} (H_{\mu_{t-1}}(X_j) - H_{\mu_t}(X_j))$$

$$\stackrel{j}{=} H_{\mu_{t-1}}(X_j) - H_{\mu_{t-1}}(X_j | X_i)$$

$$= \mathbb{E} \mathbb{I}_{\substack{j \\ \mu_{t-1}}} (X_i; X_j) \geq \eta$$

$\Rightarrow$ Potential has fallen by $\geq \eta$

OTOM. $\mathbb{E}_i H_{\mu_t}(X_i) \leq 1$

Thus if all of $\mu_1, \dots, \mu_{1/\eta}$ fail then
Contradiction.

This analysis is in expectation over the

$i_1, \dots, i_t$ & $d_1, \dots, d_t$

By averaging:

$\Rightarrow$ assignment to $i_1, \dots, i_{1/\eta}$ that

→ if an assignment

works.

→ n^n time algorithm to "brute force"

Search for such a set of vertices & their assignment

What's the connection betⁿ Local & Global Correlations? No conn. for arbitrary graphs.

will prove relationship when $\text{rank}(G)$ is small.

Connection Lemma:

Let G be a graph on $[n]$.

Let μ be a p.d. on $\{\pm 1\}^n$.

Then, If $L: C_G(\mu) \leq \beta$

then $\# \mu \leq \beta^2$, $G \cdot C(\mu) \leq \left(\frac{\beta^2 - \rho}{\text{rank}_\rho(G)} \right)^2$

Cor: $G \cdot C(\mu) \geq \left(\frac{\beta^2}{2 \cdot \text{rank}_{\beta/2}(G)} \right)^2$ ✓

Cor: If $G \cdot C(\mu) \leq \eta$ ✓

$$\Rightarrow \beta \leq 4 \cdot \text{rank}_{\frac{\beta}{2}}(G) \cdot \delta.$$

For our application, choose $\beta = \delta$.

$$\& \quad \eta = \frac{\delta^4}{4 \text{rank}_{\delta/2}(G)^2}.$$

This lemma roughly tells us that starting with $\frac{16 \text{rank}_{\delta/2}(G)^2}{\delta^4} + 2$ deg p.d. is enough to get rounding additive error $\leq \delta m$.

Let G be a graph on (n) .

Technical Lemma: Let M be a $n \times n$ psd matrix

$$\text{s.t. } \text{tr}(M) \leq n, \quad |M(i,j)| \leq 1 \quad \forall i,j.$$

Then $\mathbb{E}_{\{i,j\} \in E(G)} M(i,j) \approx \beta$ (local correlation)

$$\Rightarrow \mathbb{E}_{1 \leq i,j \leq n} M(i,j)^2 \approx \left(\frac{\beta - \rho}{\text{rank}_{\rho}(G)} \right)^2.$$

Proof of Conn. Lemma using
Technical lemma

Assume $L.C. G(\mu) \approx \delta$ - otherwise indep rounding works.

$$\begin{aligned} \overline{X_i} &\stackrel{\text{def}}{=} X_i - \tilde{\mathbb{E}} X_i \\ \overline{X_j} &\stackrel{\text{def}}{=} X_j - \tilde{\mathbb{E}} X_j \end{aligned}$$

$$\Rightarrow \mathbb{E}_{\{i,j\} \in E} \left| \mathbb{E}_{\mu} \bar{x}_i \cdot \bar{x}_j \right| \geq \delta$$

Jensen's: $\mathbb{E}_{\{i,j\} \in E} \left(\mathbb{E}_{\mu} \bar{x}_i \cdot \bar{x}_j \right)^2 \geq \delta^2 \cdot \frac{*}{2}$

$$\Rightarrow \left(\mathbb{E}_{\{i,j\} \in E} \left| \mathbb{E}_{\mu} \bar{x}_i \cdot \bar{x}_j \right| \right)^2$$

Let $M(i,j) = \left(\mathbb{E}_{\mu} \bar{x}_i \cdot \bar{x}_j \right)^2$.

Then $M \geq 0$.

* : $\mathbb{E}_{\{i,j\} \in E} M(i,j) \geq \delta^2$

Apply technical lemma

$$\Rightarrow \mathbb{E}_{1 \leq i,j \leq n} M(i,j)^2 \leq \left(\frac{\delta^2 - \rho}{\text{rank}_p(G)} \right)^2$$

Proof of $M \geq 0$: $\tilde{E}_{\mu} \cdot \bar{X} \cdot \bar{X}^T$

$$M = \tilde{\text{Cov}} \circ \tilde{\text{Cov}}$$

↓
Hadamard product

Note: $M(i,j)^2 \leq |M(i,j)| \quad \forall i,j$

Thus: $\mathbb{E}_{1 \leq i,j \leq n} M(i,j) \leq \left(\frac{\delta^2 - \rho}{\text{rank}_p(G)} \right)^2$

$$\Rightarrow \mathbb{E}_{i,j} \text{Cov}(X_i, X_j)^2 \geq \left(\frac{\delta^2 - p}{\text{rank}_p(G)} \right)^2$$

Info theory fact:

$$\text{Cov}(X_i, X_j)^2 \leq O(\mathbb{I}(X_i; X_j))$$

$$** \Rightarrow \mathbb{E}_{i,j} \mathbb{I}(X_i; X_j) \geq \Omega\left(\frac{\delta^2 - p}{\text{rank}_p(G)}\right)^2$$

$$G \cdot C(\mu) = \Omega\left(\frac{\delta^2 - p}{\text{rank}_p(G)}\right)^2$$

Choose $\delta = \sqrt{p}/4$ in our application

Proof of the technical lemma

$$\mathbb{E}_{\{i,j\} \in E} M(i,j) = \beta \quad , \quad \mathbb{E}_{1 \leq i,j \leq n} M(i,j)^2 \geq \left(\frac{\beta - p}{\text{rank}_p(G)} \right)^2$$

$$\beta = \mathbb{E}_{\{i,j\} \in E} M(i,j)$$

Facts:

$$A = \sum_i \lambda_i \tilde{v}_i \tilde{v}_i^T$$

$$= \frac{1}{n} \sum_i A_{ij} \cdot M(i,j)$$

v_i are orthonormal

$$\|v_i\|_2 = 1$$

$$2. \sum_i v_i^T M v_i = \text{tr}(M)$$

$$= \frac{1}{n} \sum_i \lambda_i \cdot v_i^T M \cdot v_i$$

$$\leq \frac{1}{n} \sum_{i: \lambda_i \geq \rho} \lambda_i \cdot v_i^T M v_i + \rho \cdot \frac{1}{n} \sum_{i=1}^n v_i^T M v_i$$

(M is psd
 $v_i^T M v_i \geq 0$)

$$\leq \frac{1}{n} \sum_{i: \lambda_i \geq \rho} \lambda_i \cdot v_i^T M v_i + \rho \cdot 1$$

Rearranging:

$$\beta - \rho \leq \frac{1}{n} \sum_{i: \lambda_i \geq \rho} \lambda_i \cdot v_i^T M v_i$$

at most $r = \text{rank}_p(G)$ terms

$$\exists v_i: \lambda_i \cdot v_i^T M v_i \geq \left(\frac{\beta - \rho}{r} \right) \cdot n$$

$$r \stackrel{\text{def}}{=} \text{rank}_p(G)$$

$\Downarrow$

(n, n)

$$v_i^T M v_i \geq \left(\frac{\beta - \rho}{\gamma} \right) \cdot n.$$

$$\max_{v: \|v\|_2=1} v^T M v \stackrel{\text{def}}{=} \|M\|_2 \geq \left(\frac{\beta - \rho}{\gamma} \right) \cdot n$$

$$\sum_{i,j} M(i,j)^2 \stackrel{\text{def}}{=} \|M\|_F^2 \geq \left(\frac{\beta - \rho}{\gamma} \right)^2 \cdot n$$

$$= \sum_i \lambda_i^2(M)$$

$$\Rightarrow \sum_{i,j} M(i,j)^2 \geq \left(\frac{\beta - \rho}{\gamma} \right)^2 \cdot n^2$$

$$\Rightarrow \frac{1}{n^2} \cdot \sum_{i,j} M(i,j)^2 \geq \left(\frac{\beta - \rho}{\gamma} \right)^2$$

Have we designed an algo for Max-Cut on low threshold rank graphs?

Roadmap

Start from μ . If $LC_G(\mu) < \delta$, done.

If not $GC(\mu) \geq \frac{\delta^4}{16 \text{rank}_{\frac{\delta}{2}}(G)^2}$



$GC(\mu') < \eta$.

for $\eta = \frac{\delta^4}{16 \text{rank}_{\frac{\delta}{2}}(G)^2}$

$$\downarrow$$

$$L.C.G(\mu') < \delta$$

$$\downarrow$$

Independently Round

Diderot: ① Constrained p.d.

② Arora-Barak-Steurer algorithm

More Applications in HW 3

Need:

$$\mathbb{E}_{\mu'} f_G \geq \mathbb{E}_{\mu} f_G - \delta$$

1. manage by adding constraint $\{f_G = \text{opt}\}$
2. or by averaging argument in GC lowering lemma