15-859GG: Lecture 2

Sum-of-Squares Algorithm, Application to Max-Cut

<u>Recall</u>: $f: \{-1,1\}^n \to \mathbb{R}$, multilinear, deg d.

Suppose f has a deg d certificate of non-neg.

$\equiv$ $f = \sum g_i^2$, g_i has deg $\leq d/2$

How can we compute it?

Suppose f ~~has~~ doesn't have a deg d cert,

<u>then</u> $\exists \mu$ s.t. $\tilde{\mathbb{E}}_\mu \cdot f < 0$.

Can we compute it?

Last time we showed: f is <u>deg d SoS</u>.

① $f = \sum_{i=1}^{r} g_i^2$, $r \leq n^{d/2}$

② ~~coeffs of g_i rationals with bit complexity poly(size(f)).~~ this is not clear

<u>OPEN Question</u>: f has rational coeff, has deg d SoS cert $\to$ is there always a deg d certificate where the constituent polynomials themselves have

rational coeff

$$f = \sum_{i=1}^{r} g_i^2$$

<u>Sturmfels</u>, Scheiderer '2013:

$\exists$ a deg d SoS polynomial with rational <u>coeffs</u> but there's no SoS representation with rational coeffs.

$$f(x) = x_0^4 + x_0 x_1^3 + x_1^4 - 3x_0^2 x_1 x_2 - 4 x_0 x_1^2 x_2 + 2 x_0^2 x_2^2 + x_0 x_2^3 + x_1 x_2^3 + x_2^4$$

What was wrong with our "proof"?

$$f(x) = \langle (1,x)^{\otimes d/2}, A\,(1,x)^{\otimes d/2} \rangle$$

$$tr(A) = f_\phi$$

$$\|A\|_F^2 \le tr(A)^2 \le f_\phi^2$$

Today we'll prove

<u>Lemma</u>: If f has deg d SoS cert then can find a deg d SoS cert for $f+\varepsilon$ in time

poly$(n^d \log 1/\varepsilon)$:

$\text{poly}(n, \log \frac{1}{\varepsilon})$

Cor: f has a deg d cert $\Rightarrow$ $f+\varepsilon$ has a deg d cert with $r \le n^{d/2}$ polynomials & size(g_i) $\le$ poly(n^d, $\log \frac{1}{\varepsilon}$, size($f$))

Proof: f has a deg d cert

$\Rightarrow \exists\ A \succcurlyeq 0$, $\underline{f(x)} = \langle (1,x)^{\otimes d/2}, \underbrace{A\,(1,x)^{\otimes d/2}\rangle}_{g_A}$

$\rightarrow$ we'll try to find the matrix A.

$$A \succcurlyeq 0$$

$$\forall\, x \in \{-1,1\}^n : g_A(x) = f(x)$$

$$\equiv \underline{\text{multilinear}(g_A) = \text{multilinear}(f)}$$

$$\underset{\substack{\subseteq [n] \\ |U| \le d}}{\forall U}: \quad \hat{f}_U = \sum_{\substack{\alpha, \beta \\ \text{odd}(\alpha+\beta) = U}} A_{\alpha,\beta} \quad \Big\}\ \underset{\text{constraints}}{\le (n+1)^d}$$

We've reduced the problem to finding a psd matrix satisfying linear equations. $\equiv$ SDP

Question: Can we decide feasibility of this constraint system?

... Not known. even if this is in NP.

Answer: No

Know: is in PSPACE.

We're going to approximately solve this.

Grötschel-Lovász-Schrijver '1981

$K \subseteq \mathbb{R}^N$

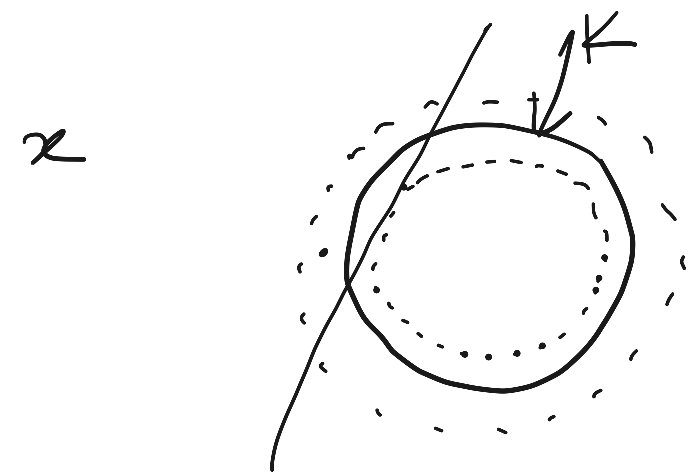
Weak Separation Oracle

a sep^n oracle for K takes input a rational vector $x \in \mathbb{R}^N$ and $\varepsilon > 0$.

either correctly asserts that $x \in K + \underbrace{B(0,\varepsilon)}_{\equiv \{x \mid \|x\|_2 \le \varepsilon\}}$

or returns an "almost separating hyperplane" in the form of $y \neq 0 \in \mathbb{R}^N$, s.t.

$$\langle y, x\rangle > \langle y, z\rangle - \varepsilon \cdot \|y\|_2 \quad \forall z \in K.$$



Thm: (GLS'81)

K: convex body: closed, convex, bounded set

$\exists R \geq r \geq 0$: $B(0,r) \subseteq K \subseteq B(0,R)$

Assume that we've a poly time weak sep^n oracle for K.

Then given any $v \in \mathbb{R}^n$ (rational).

Can compute a rational vector $x \in \mathbb{R}^n$ s.t.

① $x \in K$

② $\langle v, x \rangle \geq \langle v, z \rangle - \varepsilon \quad \forall z \in K.$

running time: $\text{poly}\left(\log \frac{R}{r} + \log \frac{1}{\varepsilon} + N\right)$

Let's apply this to our problem.

$$S = \{ A \mid A \succcurlyeq 0, \langle C_i, A \rangle = b_i \ \forall 1 \leq i \leq m, \ \|A\|_F^2 \leq f_\phi^2 \}$$

find a point inside

Convex ✓

bounded? ✓

closed ✓

$B(p, r) \not\subseteq S$

relax: $\{A \mid A \succcurlyeq 0, \langle C_i, A\rangle \in [b_i - \varepsilon, b_i + \varepsilon] \ \forall\, 1 \le i \le m\}$

$p = A + \delta \cdot \mathbb{I}$ for some fixed but tiny $\delta > 0$ works.

$A + \delta\mathbb{I} + B$, $\|B\|_F < \delta$

$A + \delta\mathbb{I} + B \succcurlyeq 0$

going back.

$$\hat{f}_u = \sum_{\substack{\alpha, \beta \\ \mathrm{odd}(\alpha+\beta) = u}} A_{\alpha, \beta} \pm \varepsilon$$

Can find some f' such that f' has a deg d SOS certificate

& $|\hat{f}_u - \hat{f}'_u| \lesssim \varepsilon$

KNOW: $f(x) = f'(x) + (f - f')(x)$

$\sum |\hat{f} - \hat{f}'| \le \varepsilon \cdot (n+1)^d$

<u>Note:</u> $\sum_{|U|\le d} |f_U - f'_U| \le \varepsilon \cdots \longrightarrow L$

<u>Then</u> we know $L + f - f'$ has a deg d certificate

f' has deg d cert

$L + f - f'$ has deg d cert

$\Downarrow$

$L + f$ has a deg d cert

$\downarrow$

$\varepsilon \cdot (n+1)^d + f$ has a deg d certificate

$\square$

Suppose f does not have a deg d cert.

$\Rightarrow \exists \mu$: $\tilde{\mathbb{E}}_\mu f < 0$

<u>degreed.</u>

<u>Thm:</u> $\exists$ a poly$(n^d, \log 1/\varepsilon, \text{Size}(f))$ algorithm to compute a pseudo-distn μ such that

$\tilde{\mathbb{E}}_\mu f < \varepsilon$.

Proof: $\tilde{\mathbb{E}}_\mu x_S$, $|S| \leq d$.

$\tilde{\mathbb{E}}_\mu 1 = 1$, $\underbrace{\tilde{\mathbb{E}}_\mu (1,x)^{\otimes d/2} \cdot (1,x)^{\otimes d/2}}$ $\succeq 0$

$\tilde{\mathbb{E}} f = \underline{\sum \hat{f}_S \cdot \tilde{\mathbb{E}}_\mu x_S} < 0$

Algorithms for computing SoS cert / pseudo-dist^n } SoS algorithm

---

Max-Cut

G: $G(V,E)$, $|V| = n$

Goal: find $S \subseteq V$, $|E(S,\bar{S})| = |\{\{i,j\} \mid i \in S, j \in \bar{S}\}|$

is maximized.

Max-Cut is NP-Hard!

Min-Cut has a poly time algo.

<u>Approx. algorithms:</u>

<u>Erdős'</u> algorithm: return a random cut

<u>approx. factor</u> $= \frac{1}{2}$

Question: Is there an LP-based algo that gets $0.5+\varepsilon$ ($\varepsilon>0$) approx?

Chan-Lee-Raghavendra-Steurer '13
K-Meka-Raghavendra '17
$\Big\}$ $\nexists\ 2^{n^\delta}$ size LP that gets $0.5+f(\delta)$ approx.

1993: Delorme-Poljak suggested an SDP relaxation

1994: Goemans-Williamson $\to$ 0.878 approx. for max-cut! $\leftarrow$ Our Goal today

Theorem: 1

Let $G = G(V,E)$. $opt(G) =$ max-cut(G).

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

for $x \in \{-1,1\}^n$. $\max_{x\in\{-1,1\}^n} f_G(x) \equiv$ max-cut(G).

Then, $\forall G$.

$$\frac{opt(G)}{0.878} - f_G(x) \text{ has a degree 2}$$

$opt(G)$

SoS certificate $\equiv f_G(x) \leq \frac{\text{opt}(G)}{0.878}$

<u>Theorem</u>: 2 Let μ be a degree 2 pseudo-distribution on $\{-1,1\}^n$. <u>Then</u>, there is an 'actual dist^n μ' such that

$$\mathbb{E}_{\mu'} f_G(x) \geq 0.878\ \tilde{\mathbb{E}}_{\mu} f_G(x).$$

<u>Proof of Thm 2 $\rightarrow$ Thm 1</u>:

Suppose $\frac{\text{opt}(G)}{0.878} - f_G(x)$ is not SoS_2

<u>then</u>, $\exists\ \mu$, deg 2, s.t. $\tilde{\mathbb{E}}_{\mu}\left(\frac{\text{opt}(G)}{0.878} - f_G\right) < 0$

$$\tilde{\mathbb{E}}_{\mu} f_G > \frac{\text{opt}(G)}{0.878}$$

$$\Rightarrow \underline{\exists \mu'} \quad \mathbb{E}_{\mu'} f_G \geq 0.878 \cdot \tilde{\mathbb{E}}_{\mu} f_G > \text{opt}(G)$$

$$\Rightarrow \underline{\mathbb{E}_{\mu'} f_G > \text{opt}(G)} \rightarrow \underline{\text{Contradiction}}.$$

<u>Theorem 2</u> $\equiv$ "rounding"

$$\boxed{\underline{\text{pseudo-dist}^n} \longrightarrow \text{dist}^n}$$

μ, $\tilde{\mathbb{E}}_{\mu} f_G(x) = \mathrm{opt}_{SOS_2}$

find μ' $\downarrow$

$\mathbb{E}_{\mu'} f_G(x) \geq \underline{0.878}\ \mathrm{opt}_{SOS_2}$

$\underline{\mu \longrightarrow \mu' \text{ efficient}}$ $\Rightarrow$ give us an algorithm to approximate max cut

---

Proof of theorem 2

"Idealized": $\mu \leftarrow$ p.d. over $\{-1,1\}^n$

$\downarrow$

$\mu' \leftarrow$ actual distn.

so that $\mathbb{E}_{\mu'} (1,x)^{\otimes 2} = \tilde{\mathbb{E}}_{\mu} (1,x)^{\otimes 2}$

not possible

Generalized Moment Problem

in 1D $\longrightarrow$ Lasserre

Lemma: (Gaussian Sampling).

for any deg 2 p.d. μ, $\exists$ an actual distn

... over $\mathbb{R}^n$ with same 1st & 2nd moments.

Proof: $\tilde{\mathbb{E}}_\mu x$

$\tilde{\mathbb{E}}_\mu x x^T \succcurlyeq 0$, $\underbrace{\tilde{\mathbb{E}}_\mu (x - \tilde{\mathbb{E}}_\mu x)(x - \tilde{\mathbb{E}}_\mu x)^T \succcurlyeq 0}$

$\tilde{\mathbb{E}}_\mu 1 = 1$

$\tilde{\mathbb{E}}_\mu (1,x)(1,x)^T \succcurlyeq 0$

We know, For any vector m, any psd matrix Σ $\exists$ a gaussian dist^n with mean m, & covariance Σ.

μ pseudo-dist^n

$\downarrow$

μ' on $\{-1,1\}^n$

Proof: $\mu \longrightarrow g$ $\quad \tilde{\mathbb{E}}_\mu (1,x)(1,x)^T = \mathbb{E}(1,g)(1,g)^T$

Issue: g does not have entries in $\{\pm 1\}$.

Suppose μ was an actual dist^n

$x \sim \mu$, $\begin{matrix} x \\ \text{or } -x \end{matrix}$ uniformly. (Symmetrization)

$\to$ mean $= 0$.

$\tilde{\mathbb{E}}\, xx^T \longrightarrow \tilde{\mathbb{E}}_\mu f_G(x)$?

Look at the p.d. with mean 0

& 2^{nd} moments $= \tilde{\mathbb{E}}_\mu xx^T$

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

$$= \frac{1}{4} \sum x_i^2 + x_j^2 - 2 \cdot x_i x_j$$

$$= \frac{1}{4} \sum (2 - 2 x_i x_j)$$

$$\tilde{\mathbb{E}}_\mu f_G(x) = \frac{1}{4} \cdot \sum (2 - 2 \cdot \tilde{\mathbb{E}}_\mu x_i x_j)$$

WLOG: can assume that $\tilde{\mathbb{E}}_\mu x = 0$

1) take $g \sim N(0, \tilde{\mathbb{E}}_\mu xx^T)$

2) $\hat{x}_i = \mathrm{Sign}(g_i)$, $\hat{x} \in \{-1,1\}^n$

$\Big\}$ algorithmic

Call μ': dist^n on $\hat{x}$.

Claim! 1 $\mathbb{E}_{\mu'} f_G(x) \geq 0.878 \cdot \tilde{\mathbb{E}}_\mu f_G(x)$.

$\left[\tilde{\mathbb{E}}_\mu (x_i - x_j)^2\right] 1/4$

$\frac{1}{4} \cdot \left[\mathbb{E}_{\mu'} (\hat{x}_i - \hat{x}_j)^2\right]$

$\quad \shortparallel \; \Pr[\hat{x}_i \neq \hat{x}_j]$

Claim 2: (Sheppard's Lemma)

$$\frac{\Pr[\text{sign}(g_i) \neq \text{sign}(g_j)]}{\mathbb{E}[(g_i-g_j)^2]} \geq \frac{2\cdot\arccos(\rho)}{\pi(1-\rho)} \quad \text{where } \rho = \tilde{\mathbb{E}}_\mu x_i x_j$$

Claim 2 $\Rightarrow$ Thm 2:

$$\min_{\rho\in[-1,1]} \frac{2\arccos(\rho)}{\pi(1-\rho)} \geq 0.878\ldots \; (\alpha_{GW})$$

$$\Rightarrow \frac{1}{4}\sum_{\{i,j\}\in E} \mathbb{E}_{\mu'}(\hat{x}_i-\hat{x}_j)^2 \geq \alpha_{GW}\frac{1}{4}\sum_{\{i,j\}\in E}\tilde{\mathbb{E}}_{\mu}(x_i-x_j)^2$$

$\square$

(minimized at $\rho = -0.69$)

Pf of Claim 2:

g_i & g_j:
$$\mathbb{E}\, g_i g_j = \tilde{\mathbb{E}}_\mu x_i x_j = \rho$$
$$\mathbb{E}\, g_i^2 = \tilde{\mathbb{E}}_\mu x_i^2 = 1$$

gaussian distn is completely specified by mean & covariance.

1. Let v, w be 2 vectors in $\mathbb{R}^2$ s.t.

$\|v\|_2 = \|w\|_2 = 1$, $\langle v, w \rangle = \rho$.

2. take $h \sim N(0, \mathbb{I}_2)$ (2D, std gaussian)

$\mathbb{E} h_i^2 = \mathbb{E} h_2^2 = 1$

$\mathbb{E} h_i h_j = 0$.

3. $\hat{g}_i = \langle h, v \rangle$

$\hat{g}_j = \langle h, w \rangle$

<u>Then</u>, $\mathbb{E}\, \hat{g}_i^2 = \mathbb{E} \langle h, v \rangle^2 = \|v\|_2^2 = 1$

$\mathbb{E}\, \hat{g}_j^2 = 1$

$\mathbb{E}\, \hat{g}_i \cdot \hat{g}_j = \rho$ (direct computation).

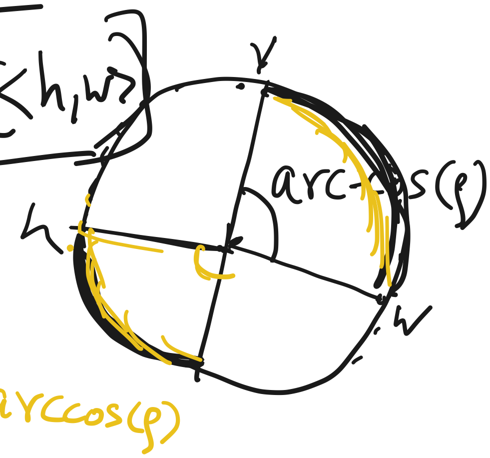
<u>What are we interested in?</u>

$$\Pr[\, \text{sign}(g_i) \neq \text{sign}(g_j) \,]$$

$$= \underline{\Pr[\, \text{sign}(\hat{g}_i) \neq \text{sign}(\hat{g}_j) \,]}$$

$$= \Pr\{\, \underline{\text{sign}(\langle h, v \rangle)} \neq \underline{\text{sign}(\langle h, w \rangle)} \,\}$$

$\langle v, w \rangle = \rho$



$$\frac{2 \cdot \text{arc-cos}(\rho)}{2\pi}$$

$$= \text{arc-cos}(\rho)$$

$$\frac{1}{4}\mathbb{E}_\mu (x_i - x_j)^2 = \mathbb{E}[(g_i - g_j)^2] = \mathbb{E}[(\hat{g}_i - \hat{g}_j)^2]$$

$$= \frac{1}{4}(2 - 2\rho) = \frac{1}{2} - \frac{\rho}{2}$$

$$\underline{\text{ratio}} = \frac{2 \arccos(\rho)}{\pi \cdot (1-\rho)}$$

<u>Corollary</u>: Suppose $\tilde{\mathbb{E}}_\mu f_G(x) \geq (1-\delta) \cdot |E|$.

<u>Then</u>. $\mathbb{E}_{\mu'} f_G(x) \geq (1 - O(\sqrt{\delta})) \cdot |E|$.

$\Rightarrow$ approx. ratio of the algo > 0.5

<u>Can we do better?</u>

<u>With deg 2 SoS?</u> <u>No</u>.

deg 4 SoS? $\rightarrow$ OPEN

$\vdots$

deg log(n) SoS? OPEN

<u>Unique Games Conjecture</u>:

<u>Khot-Kindler-Mossel-O'donnell '04</u>:

UGC $\Rightarrow$ $(\alpha_{GW} + \varepsilon)$-approx. Max-Cut is NP-Hard

$\forall \varepsilon > 0.$

Is UGC true? don't know.

2018: Dinur - Khot - Kindler - Minzer - Safra

2-to-2 Games Conjecture (a variant of UG is NP-Hard)

Fine-grained Question

Suppose $opt(G) \geq (1-\delta)\cdot |E|.$

"round" $\mathbb{E}_{\mu'} f_G(x) \geq \underbrace{(1 - O(\sqrt{\delta}))}\cdot |E|.$

→ is this rounding improvable?

Gaussian rounding is NOT optimal.

RPR^2 roundings

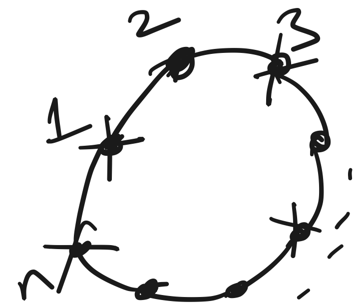
Integrality Gaps?

Suppose you want a degree 2 SoS certificate of non-neg for $\frac{opt(G)}{c} - f_G(x)$

What's the largest c for which a certificate exist?

→ We'll "prove" that $c = 0.878$ is optimal.

→ take $G = C_n$ for odd n



Theorem: C_n : cycle on n vertices
n odd.

(Note: $\text{max-cut}(C_n) = \text{opt}(C_n) = \frac{n-1}{} = (1-\frac{1}{n}) \cdot |E|$)

There's a pseudo-dist^n of deg 2 μ s.t.

$$\tilde{\mathbb{E}}_\mu f_{C_n}(x) \geq \left(1 - O\left(\frac{1}{n^2}\right)\right) \cdot |E|$$

Choose $n = 1/\delta$. then $\text{opt}(C_n) = (1-\delta)|E|$

$\cdot$ $\text{opt}_{sos_2}(C_n) \geq (1 - O(\delta^2))|E|$

⇒ Corollary for small δ is tight up to constant factors

Cycle = "discretized" 2-dim sphere

⋮

"discretized" high-dim sphere

(Feige-Schechtman '02)

<u>Proof (Sketch)</u>:

<u>Idea</u>: $f_G(x) = \frac{1}{4} \sum_{\{i,j\}\in E} (x_i - x_j)^2$

L_G: Laplacian matrix of G (regular) : $\underbrace{\deg \cdot \mathbb{I} - A}$ $\succeq 0$

<u>Claim</u>: $x^T L_G x = \sum_{\{i,j\}\in E} (x_i - x_j)^2 \geq 0$

<u>Goal</u>: Max-Cut $= \max_{x\in\{-1,1\}^n} x^T L_G x$

<u>related</u>: $\max_{x\in \|x\|_2=\sqrt{n}} x^T L_G x$

$= n \cdot \|L_G\|_2$

largest singular value of L_G

<u>How do</u> we construct a pseudo-distⁿ μ of deg 2 for the cycle $\tilde{\mathbb{E}}_\mu f_G(x) \geq 1 - O(1/n^2)$?

$f_G(x) = \underbrace{x^T L_G x}$

Choose a dist^n on x that are in the "largest eigenspace" of L_G.

$\tilde{\mathbb{E}}_\mu (1,x)(1,x)^T \succcurlyeq 0$, $\quad \tilde{\mathbb{E}}_\mu x_i^2 = 1$

$\rightarrow$ μ is a deg 2 p.d. over $\{-1,1\}^n$

$\Leftrightarrow$ $\tilde{\mathbb{E}}(1,x)(1,x)^T \succcurlyeq 0$, $\tilde{\mathbb{E}}_\mu = 1$

& $\tilde{\mathbb{E}}_\mu x_i^2 = 1 \ \forall i$

$$\underline{\text{Only}} \quad \underline{\tilde{\mathbb{E}}_\mu x x^T} = n\begin{pmatrix} 1 & & ? & & \\ & 1 & & & \\ & & 1 & & \\ ? & & & 1 & \\ & & & & 1 \end{pmatrix} = M$$

$\tilde{\mathbb{E}}_\mu f_G(x) \succcurlyeq 1 - O(1/n^2)$.

$\underline{\text{Idea:}}$ (1) L_G top eigenvalue: $\underline{1 - O(1/n^2)}$

$\downarrow$

top eigspace is 2-dim, Y_1, Y_2

$M = \tilde{\mathbb{E}}_\mu x x^T = \underline{Y_1} Y_1^T + \underline{Y_2} Y_2^T$

$$= \begin{pmatrix} 1 & & ? & \\ & 1 & & \\ & & 1 & \\ ? & & & 1 \end{pmatrix}$$

Idea: $\nexists$ boolean coordinate eigenvector of the top eig-space of L_G

but $\exists$ two vectors unit such that $v_1 v_1^T + v_2 v_2^T$ has diagonal $\vec{1}$.

Cycles: C_n.

eigenvectors: x_k , y_k

$$x_k(u) = \cos\left(\frac{2\pi k u}{n}\right)$$

$$y_k(u) = \sin\left(\frac{2\pi k u}{n}\right).$$

$0 \leq k \leq \frac{n-1}{2}$ (remember n is odd)

$$x_k \cdot x_k^T + y_k \cdot y_k^T = \sin^2\theta + \cos^2(\theta) = 1$$

$$\theta = \frac{2\pi k u}{n}$$

$\{i, j\} \rightarrow$ random edge

$\rho_{\ldots} \approx -0.69$

$\mathbb{E}_{\mu} x_i x_j \quad \sim \quad)_{GW} \qquad)_{GW}$

deg $\log(n)$ p.d.

Question: Is there a $2^{\sqrt{n}}$ time algo to get $0.878 + 0.0000001$ approx. for max-cut?

U.G. has a $2^{n^{\delta}}$ time approx. algo.

$G(n,p)$: $\approx \frac{1}{2} + \frac{p}{2}$