Lecture 4

- Last time:

  estim. avg degree with $(2+\varepsilon)$-approx using deg queries

- Today: estim avg deg with deg queries & neighb. queries

  on $(v, i) \to$ i^{th} neighb of v

$$\bar{d} = \frac{\sum d_v}{n}$$

$$\beta = \frac{\varepsilon}{c}; \quad c > 0$$

$$B_i = \{ v \mid (1+\beta)^{i-1} < \deg v \leq (1+\beta)^i \}$$

Alg

Sample $S = \sqrt{n}(\text{polylog } n)\text{poly}\frac{1}{\varepsilon}$

Let $S_i = B_i \cap S$

Estimate $\frac{|B_i|}{n} : \approx \frac{|S_i|}{|S|}$

for all i

if $|S_i| \geq \sqrt{\frac{\varepsilon}{n}} \frac{|S|}{ct} = \text{poly}(\log n, \frac{1}{\varepsilon})$

take $p_i = \frac{|S_i|}{|S|}$ $\quad \mathbb{E}\left[\frac{|S_i|}{|S|}\right] = \frac{|B_i|}{n}$

→ ow. $p_i = 0$ ← ignore small buckets

Output $\sum p_i (1+\beta)^{(i-1)}$ $\quad 1+\beta \approx 1+\frac{\varepsilon}{c}$ multpl. approx

Recall $\sum_i \frac{|B_i|}{n}(1+\beta)^{i-1} \leq \bar{d} \leq \sum_i \frac{|B_i|}{n}(1+\beta)^{(i)}$

Showed: if $|B_i| > \frac{2\sqrt{\varepsilon n}}{ct}$ twice as big as threshold for S_i to be large

$\mathbb{E}[S_i] \geq (2)\sqrt{\frac{\varepsilon}{n}} \cdot \frac{1}{ct}$

So whp Alg will label S_i as large. 2T

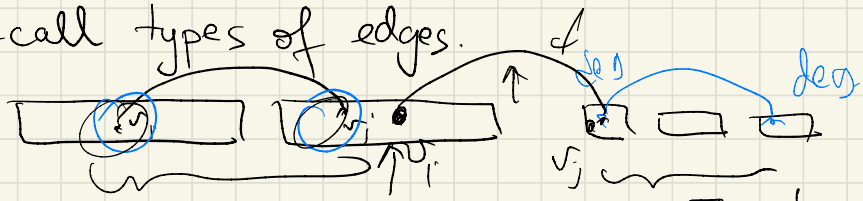
- Showed that if $B_i < \frac{2\sqrt{\varepsilon n}}{ct}$ (B_i is small)

$$\bigcup_{B_i \text{ is small}} B_i \leq \frac{2\sqrt{\varepsilon n}}{c} = O(\sqrt{\varepsilon n})$$

So # edgs bet B_i, B_j that are both small $< \varepsilon n$ $\Rightarrow$

If ignored w. ncur an $(1+\varepsilon)$ multiplic error in estimating $\overline{d}$. (So these are safe to ignore)

Recall types of edges.



① Eds $v_i - v_j$ bet big Buckets

alg considers all degs in big buckets so edge considered both in deg v_i & deg v_j, which is the correct way to comp $\sum \deg v_i$.

② Eds $v_i - v_j$ bet vts in small buckets

Alg ignores these, but only εn such edges anyway

so $(1+\varepsilon)$ mult. loss

③ Eds $v_i - v_j$ with $v_i \in$ Small

$v_j \in$ Big bucket.

These eds are counted once in $\sum \deg v_i$,

i.e. only when taking into acct $\deg v_i$.

Should be counted twice, so we're

underestimating by a fact ≤ 2.

$$\sum (1+\beta)^{i-1} \frac{|B_i|}{n} \leq \bar{d} \leq \sum (1+\beta)^{i} \frac{|B_i|}{n} \leftarrow$$

vs. ↙

$$\frac{(1-\gamma)}{2} \cdot \frac{1}{2} \leq \sum \frac{|S_i|}{|S|} \cdot (1+\beta)^{i-1} + 0 \leq \frac{(1+\beta)}{(1-\gamma)} \bar{d}$$

$\gamma \approx \varepsilon$.

$\frac{|B_i|}{n}(1+\gamma)$ for big buckets.

Hence: Alg gives a $(2+\varepsilon)$-approx to $\bar{d}$ in $O(\sqrt{n}\,\text{poly}\log n\,\text{poly}(\frac{1}{\varepsilon}))$ time with deg queries.

This is tight: $\exists$ graphs that require $\Omega(\sqrt{n})$ deg. queries.

Q: Can we get $(1+\varepsilon)$ mult. approx?

Yes: New query: random neighbor query(v)

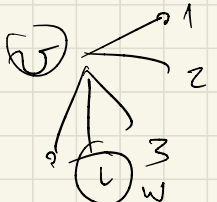
- given v, return rand neighb of v:

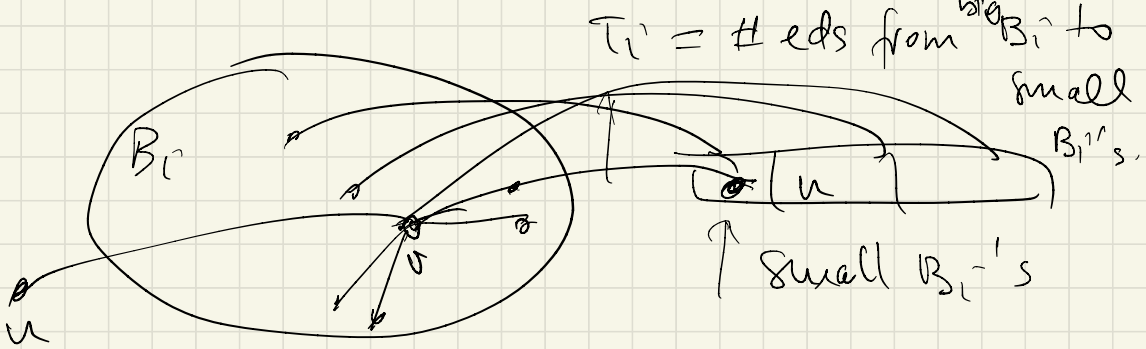
  ie. make deg query to v $= \deg(v)$

  pick random $i \in [1 \dots \deg(v)]$.

  neighbor query (v, i)

new type of query.





repeat $s = O(1/\delta)$

pick rand $v \in B_i$

make rand neighb query. ✓

$\Rightarrow$ negh. u

$\rightarrow$ $a_j = \begin{cases} 1 \\ 0 \end{cases}$ if $(v, u) \in$ Big-Small

Output $\frac{\sum a_j}{s} = \tilde{a}$ ←

$$\mathbb{E}[a_j] = \frac{\#\text{ eds from } B_i \text{ to small buckets}}{\#\text{eds from } B_i \text{ to } V}$$

$$= \frac{T_i}{\#\text{eds } B_i \to V}$$

↑

$$(1+\beta)^{i-1}|B_i| \le \#\text{eds } B_i \to v \le (1+\beta)^{i}|B_i|$$

$$\underbrace{\mathbb{E}[a_j](1+\beta)^{i-1}|B_i|}_{} \le T_i \le \mathbb{E}[a_j](1+\beta)^{i}|B_i|$$

want estimate. for big B_i

know:

whp. (Chernoff) $(1-\gamma)|B_i| \le \frac{|S_i|}{S} n \le (1+\gamma)|B_i|$

$\forall i$ big.

whp

$$(1-\gamma)\mathbb{E}[a_j] \le \tilde{a} \le (1+\gamma)\mathbb{E}[a_j].$$

Can output $T_i \approx \tilde{a}(1+\beta)^{i-1}\frac{|S_i|}{S}$

$(1 \pm \gamma)$

We proved $(1+\varepsilon)$ - approx deg queries & rand neighb. queries.

Testing Δ-freeness in dense graphs

Δ-freeness: G doesn't have any Δ's.

dense $\Theta(n^2)$

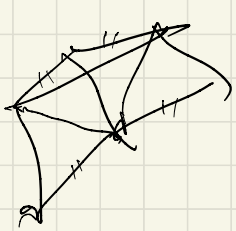
We'll show

Thm: Δ-freeness is testable with ct. # of queries

Test: on ε, G on n vts & $\Theta(n^2)$ eds.

- acc if G is Δ-free.
- rej wp $\geq 2/3$ if G is ε-far from being Δ-free.

ε-far: need to remove $\Omega(\varepsilon n^2)$ eds in order to get to a Δ-free graph.



Basic test.

Sample 3 rand vert. $v_1 v_2 v_3$

if the form a Δ then rej.

ow acc.

Completeness: if G is Δ-free. then $Pr[acc] = 1$

Soundness: if G is ε-far from Δ-free.

want $Pr[\text{Basic rejects}] = \delta(\varepsilon, n)$

Note that run Basic test for $\frac{2}{\delta}$ times

then $Pr[\text{rej}] > \frac{2}{3}$. $\leftarrow (1 - (1-p)^{2/p})$

$\rightarrow Pr[\text{Basic test reject}] = \frac{\#\text{ triangles in } G}{\text{total} \rightarrow \binom{n}{3}}$

$\#$ u, v, w vts.

want $= \delta(\varepsilon, n) \rightarrow O(n)$

$\rightarrow f(\varepsilon) \rightarrow$ ot !!!

$1/\delta(\varepsilon, n) = 2^{2^{2^{\cdot^{\cdot^{2}}}}}$ ↕ $\mathrm{poly}(1/\varepsilon)$

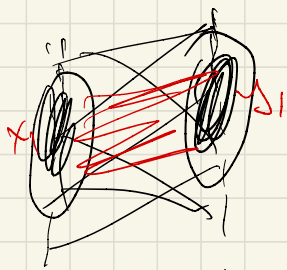
⇕

$\Theta\left(\frac{1}{\varepsilon}\right)^2$

comes from SRL.

Szemeredi Reg Lemma

$$d(V_1, V_2) = \frac{e(V_1, V_2)}{|V_1|\,|V_2|}$$ ← density



ε-regularity

$G = (V_1 \cup V_2, E)$

G is ε-regular if $\forall\, X \subseteq V_1$ $|X| > \varepsilon |V_1|$

$\forall\, Y \subseteq V_2$ $|Y| > \varepsilon |V_2|$

Then $\left| d(V_1, V_2) - d(X, Y) \right| < \varepsilon$

- homogeniety
- expansion prop. (pseudorand.)

Equipartition: $G = (V, E)$

partition of V into $V_1 \dots V_k$

s.t. $\big||V_i| - |V_j|\big| \le 1$

$G[S_1, S_2]$ the subgraph induced by $S_1 \& S_2$

Szemeredi's Reg. Lemma

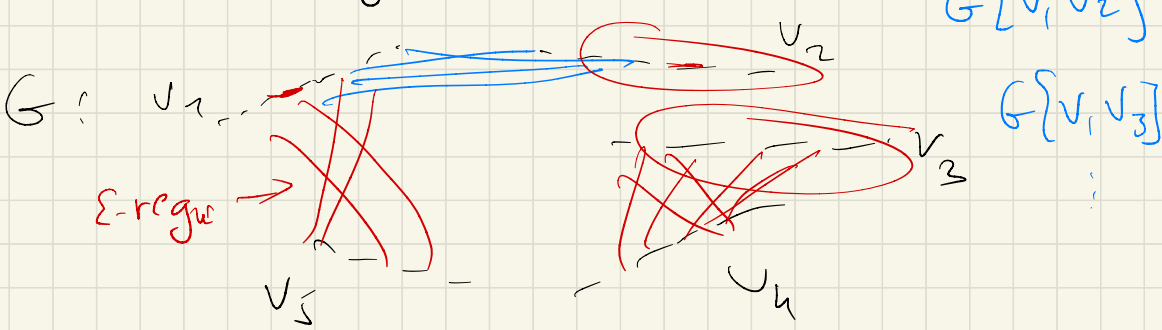
$\forall \varepsilon \quad \exists M = M(\varepsilon)$ s.t. $\forall G$ on $\ge M$ vts

there is an equip of V into $V_1 \dots V_k$

$$\frac{1}{\varepsilon} \le k \le M(\varepsilon) \text{ s.t.}$$

At least $(1-\varepsilon)\binom{k}{2}$ of $G[V_i, V_j]$

are ε-regular.



Existential: $M(\varepsilon) = 2^{2^{\cdot^{\cdot^{2}}}}$ ↑ $\text{poly}(1/\varepsilon)$

For, Lovasz $\Theta(1/\varepsilon^2)$.

Algorithmic: Alon et al: can find a partit. in $O(n^{2.3})$ ↑ $n^{W \leftarrow \text{exp}}$ of mat.

Next time:

- Δ-Removal Lemma (via Komlós Simonovits Lemma)
- Szem. reg. lemma (the proof)