Lecture 24

Szemerédi's Regularity Lemma

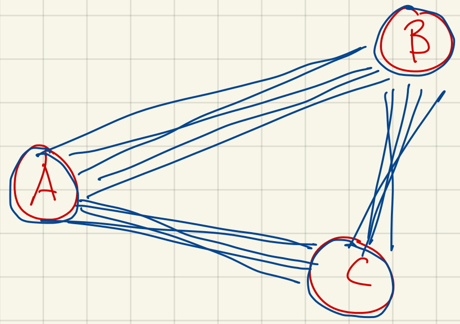
Testing dense graph properties via SRL:

Δ-freeness

Graphs with "random" properties:

Example question:

How many triangles in a random tripartite graph?



density η

$\forall$ $u \in A$, $v \in B$, $w \in C$:

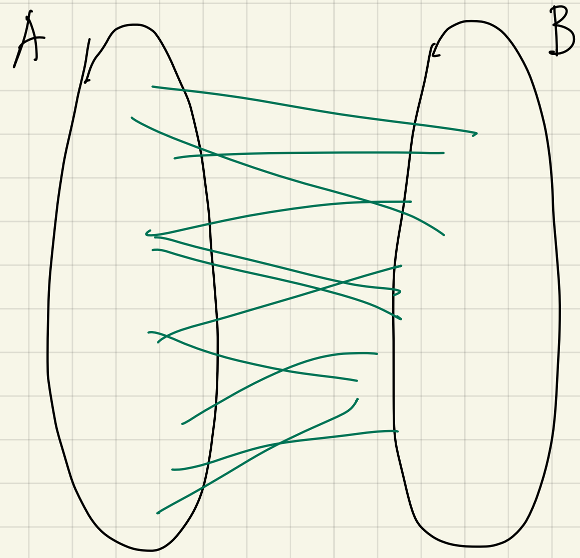
$$\Pr[u \sim v \sim w] = \eta^3 \qquad \sigma_{u,v,w} = \begin{cases} 1 & \text{if } u \sim v \sim w \\ 0 & \text{o.w.} \end{cases}$$

$$E[\sigma_{u,v,w}] = \eta^3$$

$$E[\#\text{ triangles}] = E\Big[\sum_{\substack{u \in A \\ v \in B \\ w \in C}} \sigma_{u,v,w}\Big] = \eta^3 \cdot |A| \cdot |B| \cdot |C|$$

Can we make weaker assumptions + still get reasonable bounds?

Density & Regularity of set pairs:



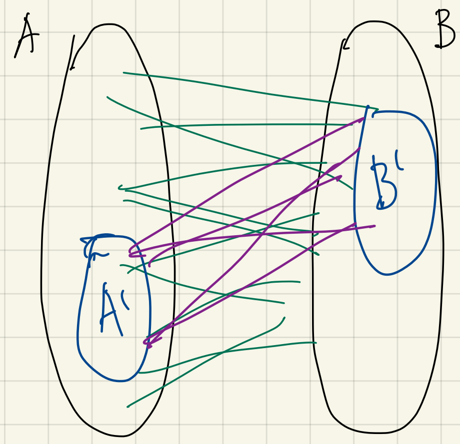
__def.__ For $A, B \leq V$ s.t.

(1) $A \cap B = \emptyset$

(2) $|A|, |B| > 1$

Let $e(A,B)$ = # edges between $A + B$

+ density $d(A,B) = \dfrac{e(A,B)}{|A|\cdot|B|}$



Say A, B is γ-regular if $\forall\, A' \leq A, B' \leq B$

s.t. $|A'| \geq \gamma |A|$

$|B'| \geq \gamma |B|$

$$|d(A',B') - d(A,B)| < \gamma$$

behaves like a "random graph"

(parameter γ used in 2 ways to "conserve" on parameters)

Lemma

density

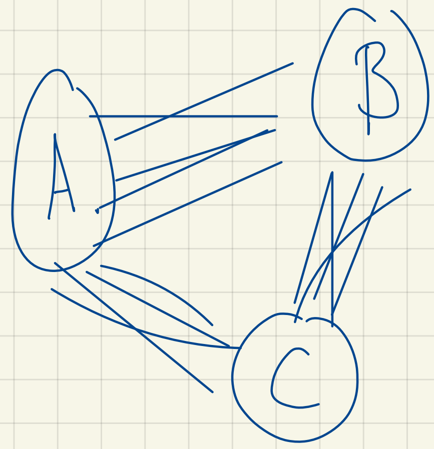
$\forall \eta > 0$ $\quad \exists \gamma = \frac{1}{2}\eta \equiv \gamma^{\Delta}(\eta)$

regularity parameter, depends only on η

$\delta = (1-\eta)\frac{\eta^3}{8} \geq \frac{\eta^3}{16} \equiv \delta^{\Delta}(\eta)$

if $\eta < 1/2$

triangles, depends only on η



s.t. if A, B, C disjoint subsets of V s.t. each pair is γ-regular with density $> \eta$ then G contains $\geq \delta \cdot |A| \cdot |B| \cdot |C|$ distinct Δ's with node in each of A, B, C.

Wow! differ only by factor of 16

Compare to random graphs: $\geq \eta^3 |A| \cdot |B| \cdot |C|$

Proof:

$A^* \leftarrow$ nodes in A with $\geq |\eta - \gamma| \cdot |B|$ nbrs in B

$\geq |\eta - \gamma| \cdot |C|$ nbrs in C

{ if this were a random graph, you would expect to have $\geq \eta |B|$ nbrs, $\geq \eta |C|$ $\pm$ some error. So these are the nodes that "look random"

Claim $|A^*| \geq (1 - 2\gamma) |A|$

Why? (Pf of claim)

$A' \leftarrow$ "bad" nodes w.r.t. B $\;(< |\eta - \gamma| \cdot |B|$ nbrs in $B)$

$A'' \leftarrow$ " " " C $\;($ " " " " $C)$

then $|A'| \leq \gamma |A|$ $\quad (\& \; |A''| \leq \gamma |A|)$

why? consider pair A', B.

$$d(A', B) < \frac{|A'| \cdot |\eta - \gamma| \cdot |B|}{|A'| \cdot |B|} = \eta - \gamma$$

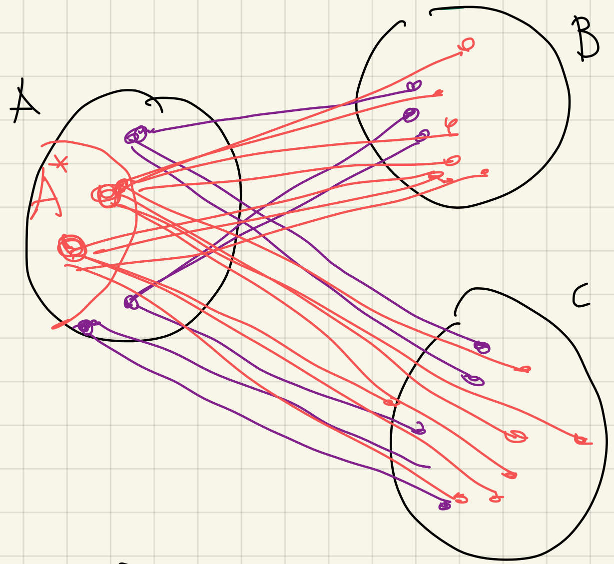
def of A'

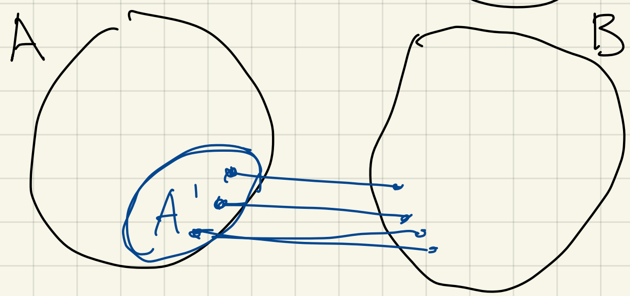
but $d(A, B) > \eta$

so $|d(A', B) - d(A, B)| > \gamma$

+ $|B| \geq \gamma |B|$

so if $|A'| \geq \gamma |A|$ then (A, B) is not γ-regular



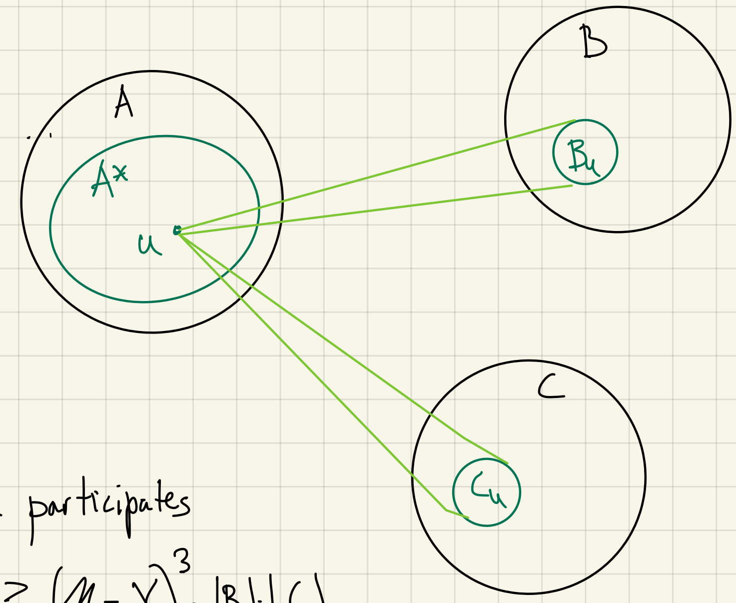


Let $A^* = A \setminus (A' \cup A'')$ then $|A^*| \geq |A| - |A'| - |A''| \geq |A| - 2\gamma |A| = (1 - 2\gamma) \cdot |A|$

Using claim:

For each $u \in A^*$: define $B_u \equiv$ nbrs of u in B
$C_u \equiv$ nbrs of u in C

both pretty big by def of A^*



since $\gamma < \frac{\eta}{2}$ $(\eta - \gamma > \gamma)$, $|B_u| \geq (\eta-\gamma)|B| \geq \gamma |B|$

$|C_u| \geq (\eta - \gamma)|C| \geq \gamma |C|$

Find lots of distinct Δ's

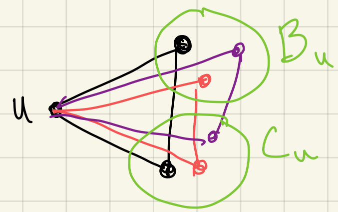
edges between $B_u + C_u$ $\Rightarrow$ lower bnd on # distinct Δ's in which u participates

$$d(B,C) \geq \eta \Rightarrow d(B_u, C_u) \geq \eta - \gamma \Rightarrow e(B_u, C_u) \geq (\eta - \gamma)|B_u||C_u| \geq (\eta-\gamma)^3 \cdot |B||C|$$

B_u, C_u big enough & (B,C) is γ regular

so total # Δ's $\geq \underbrace{(1-2\gamma)|A|}_{\text{this is where claim gets used}} \cdot (\eta-\gamma)^3 |B||C| \geq (1-\eta)(\eta/2)^3 |A||B||C|$ ∎

choose $\gamma < \eta/2$



Do interesting graphs have regularity properties?

Yes in some sense, all graphs do "can be approximated as small collection of random graphs"

Szemerédi's Regularity Lemma

would like it to say:

"one can equipartition nodes of V into $V_1 \dots V_k$ (for const k) s.t.

Sometimes need $k > m$ for some m ($k=1$, $k=n$ trivial) useful in applications

all pairs (V_i, V_j) are ε-regular"

only most $\leq \varepsilon \binom{k}{2}$ are not

Szemerédi's Regularity Lemma: (especially useful version)

$\forall\, m, \varepsilon > 0$ $\quad \exists\, T = T(m,\varepsilon)$ s.t. given $G = (V,E)$ s.t. $|V| > T$

& $\mathcal{A}$ an equipartition of V into sets

then exists equipartition $\mathcal{B}$ into k sets which refines $\mathcal{A}$

independent of n

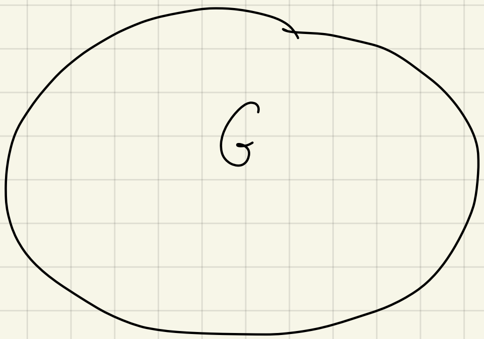
s.t. $m \le k \le T$

& $\le \varepsilon\binom{k}{2}$ set pairs <u>not</u> ε-regular

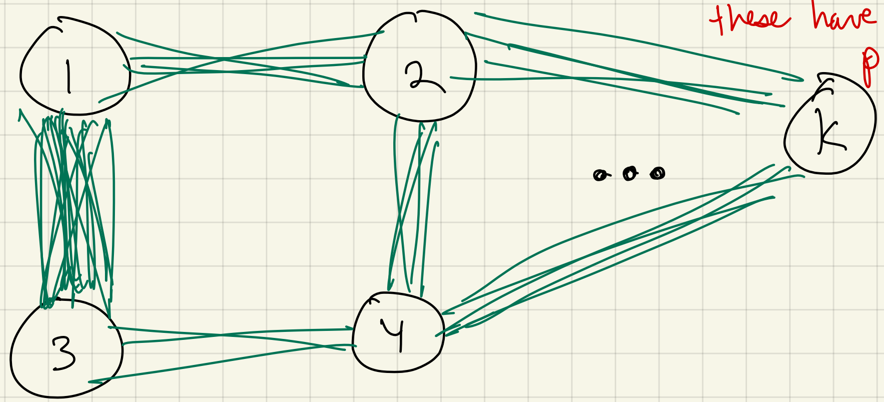
const # partitions s.t. each pair behaves like random graph

$\underbrace{\text{random graph}}_{\text{these have nice properties!}}$

Note: T does <u>not</u> depend on $|V|$



$\Rightarrow$



Why was SRL first studied?

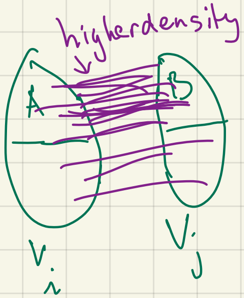
to prove conjecture of Erdös + Turán: sequences of ints have long arithmetic progressions

Very rough idea of proof:

"expectation of $d^2(V_i, V_j)$" $\longrightarrow$ $\displaystyle \operatorname{ind}(V_1 \dots V_k) = \frac{1}{k^2} \sum_{i=1}^{k} \sum_{j=i+1}^{k} d^2(V_i, V_j) \le \frac{1}{2}$ "variance of d"

same densities

note: $E[d(V_i, V_j)] = \dfrac{|E|}{|V|^2}$



issue: what if split V_i for many V_j?
$\Rightarrow$ split into exponential subsets

if a partition violates, can refine s.t. $\operatorname{ind}(V_1', \dots, V_{k'}')$ grows significantly (i.e. by $\approx \varepsilon^c$) so in less than $\frac{1}{\varepsilon^c}$ refinements, have good partition

note, if refine, Cauchy Schwartz $\Rightarrow$ ind can't decrease

How big is k? u.b. Tower of size $\frac{1}{\varepsilon^c}$

l.b. " " " $\frac{1}{\varepsilon^{c'}}$

$\left.2^{2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}}\right\} \frac{1}{\varepsilon^c}$

An application of the SRL:

Given G in adj matrix form

Is it Δ-free?

desired behavior: if G is Δ-free, output PASS

if G $\underbrace{\varepsilon\text{-far}}_{\text{must delete } \geq \varepsilon n^2 \text{ edges}}$ from Δ-free output FAIL

1-sided error

Algorithm: Do $O(\delta^{-1})$ times:

Pick $v_1, v_2, v_3 \in_r V$

if Δ reject & halt

Accept

fctn of ε only

Thm $\forall \varepsilon$, $\exists \delta$ s.t. $\forall G$ s.t. $|V| = n$

& s.t. G is ε-far from Δ-free,

then G has $\geq \delta \binom{n}{3}$ distinct Δ's

Corr Algorithm has desired behavior

Why?
- if Δ-free: we never reject
- if ε-far from Δ-free:

  $\geq \delta \binom{n}{3}$ Δ's

  $\Rightarrow$ each loop passes with prob $\leq 1 - \delta$

  $$\Pr[\text{don't find } \Delta] \leq (1-\delta)^{c/\delta}$$
  $$\leq e^{-c} < 1/3$$

  for proper choice of c

  $\Rightarrow$ reject with prob $\geq 2/3$

Thm $\forall \varepsilon, \exists \delta$ s.t. $\forall G$ s.t. $|V| = n$

& s.t. G is ε-far from Δ-free,

then G has $\geq \delta \binom{n}{3}$ distinct Δ's

Proof

Use regularity to get equipartition $\{V_1 \cdots V_k\}$ s.t.

$$\frac{5}{\varepsilon} \leq k \leq T\left(\frac{5}{\varepsilon}, \varepsilon'\right)$$ $\Leftarrow$ need $\geq \frac{5}{\varepsilon}$ sets in partition so that no set has $\geq \frac{\varepsilon}{5}$ fraction of nodes

equivalent: $\frac{\varepsilon n}{5} \geq \frac{n}{k} \geq \frac{n}{T(\frac{5}{\varepsilon}, \varepsilon')}$

how? start with arbitrary equipartition into $5/\varepsilon$ sets $\leftarrow$ this is why we need ability to refine any partition

for $\varepsilon' \equiv \min\left\{\frac{\varepsilon}{5}, \gamma^{\Delta}\left(\frac{\varepsilon}{5}\right)\right\}$

s.t. $\leq \varepsilon' \binom{k}{2}$ pairs not ε'-regular assume $\frac{n}{K}$ is integer

$G' =$ take G and

1) delete edges internal to any V_i

(if # nodes per partition small, few internal edges)

how many? $\leq \frac{n}{K} \cdot n \leq \frac{\varepsilon n^2}{5}$

deg w/in V_i

sum over all nodes

2) delete edges between ε'-non regular pairs

how many?

$$\leq \underbrace{\varepsilon' \binom{K}{2}}_{\text{\# non regular pairs}} \cdot \underbrace{\left(\frac{n}{K}\right)^2}_{\text{max \# edges per pair}} \leq \frac{\varepsilon}{5} \cdot \frac{K^2}{2} \cdot \frac{n^2}{K^2} \leq \frac{\varepsilon}{10} n^2$$

since $|V_i| \approx |V_j| = \frac{n}{K}(+1)$

$$\frac{\varepsilon n}{5} \geq \frac{n}{K} \geq \frac{n}{T\left(\frac{5}{\varepsilon}, \varepsilon'\right)}$$

$$d(A,B) = \frac{e(A,B)}{|A| \cdot |B|}$$

A, B is γ-regular if $\forall A' \subseteq A, B' \subseteq B$

s.t. $|A'| \geq \gamma |A|$

$|B'| \geq \gamma |B|$

$|d(A', B') - d(A,B)| < \gamma$

$\gamma^{\Delta}(\eta) = \frac{1}{2}\eta$

$\delta^{\Delta}(\eta) = (1-\eta)\frac{\eta^3}{8} \geq \frac{\eta^3}{16}$

$\varepsilon' = \min\{\frac{\varepsilon}{5}, \gamma^{\Delta}(\frac{\varepsilon}{5})\}$

$\Rightarrow \leq \varepsilon' \binom{K}{2}$ pairs not ε'-regular

3) delete edges between $\underbrace{\text{low density}}_{\text{use } < \varepsilon/5 \text{ as cutoff}}$ pairs

how many?

$$\leq \sum_{\substack{\text{low}\\\text{density}}} \left(\frac{\varepsilon}{5}\right)\left(\frac{n}{K}\right)^2$$

note $\sum \left(\frac{n}{K}\right)^2 \leq \binom{n}{2}$

$$\leq \frac{\varepsilon}{5}\binom{n}{2} \approx \frac{\varepsilon n^2}{10}$$

$$\frac{\varepsilon n}{5} \geq \frac{n}{K} \geq \frac{n}{T\left(\frac{5}{\varepsilon}, \varepsilon'\right)}$$

$$d(A,B) = \frac{e(A,B)}{|A|\cdot|B|}$$

A,B is γ-regular if $\forall A' \subseteq A, B' \subseteq B$
s.t. $|A'| \geq \gamma |A|$
$|B'| \geq \gamma |B|$

$$|d(A',B') - d(A,B)| < \gamma$$

$\varepsilon' = \min\{\frac{\varepsilon}{5}, \gamma^*(\frac{\varepsilon}{5})\}$

$\forall$ $\leq \varepsilon' \binom{K}{2}$ pairs not ε'-regular

Total deleted edges: $\leq \frac{\varepsilon n^2}{5} + \frac{\varepsilon n^2}{10} + \frac{\varepsilon n^2}{10} < \varepsilon n^2$

But G is ε-far from Δ-free (must delete $\geq \varepsilon n^2$ edges to remove all Δ's)

so G' must still have a triangle!!!

Δ in G' must connect:

1) nodes in 3 distinct $V_i\ V_j\ V_k$

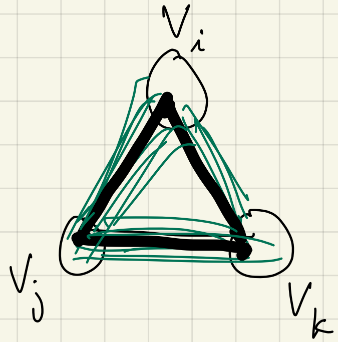
since no edges internal to partition in G'

2) regular pairs

since nonregular pair edges deleted in G'

3) high density pairs

since removed low density pairs in G'



$\therefore \exists\ i, j, k$ distinct s.t. $x \in V_i,\ y \in V_j,\ z \in V_k$

V_i, V_j, V_k all $\geq \frac{\varepsilon}{5}$ density pairs ($\frac{\varepsilon}{5} = 4\eta$)

$+ \geq \underbrace{\gamma^{\Delta}\left(\frac{\varepsilon}{5}\right)}_{\geq \frac{\eta}{2} \geq \frac{\varepsilon}{10}}$-regular

Δ-counting lemma $\Rightarrow$

$$\geq \delta^{\Delta}\left(\frac{\varepsilon}{5}\right) |V_i| \cdot |V_j| \cdot |V_k| \quad \text{triangles in } G'$$

where $\delta^{\Delta} = (1-\eta)\frac{\eta^3}{8}$

$$\geq \frac{1}{2} \cdot \frac{\varepsilon^3}{8000} = \frac{\varepsilon^3}{16000}$$

$$\geq \frac{\delta^{\Delta}\left(\frac{\varepsilon}{5}\right) n^3}{\left(T(5/\varepsilon, \varepsilon')\right)^3} \quad \Delta's$$

$$\geq \delta' \cdot \binom{n}{3} \quad \Delta's \text{ in } G' \text{ (and thus in } G)$$

for $\delta' = 6\,\delta^{\Delta}\left(\frac{\varepsilon}{5}\right)\left(T\left(\frac{5}{\varepsilon}, \varepsilon'\right)\right)^{-3}$

$\blacksquare$

This is a powerful technique!

- similar lemma to Δ counting holds for all const sized subgraphs
- almost "as is" can use same method to test all

"1st order" graph properties:

$$\exists\, u_1, u_2 u_3 \cdots u_K \quad \forall\, v_1 \cdots v_\ell \quad \underbrace{R(u_1 \cdots u_K\, v_1 \cdots v_\ell)}_{R \text{ defined via } \wedge, \vee, \neg \text{ + neighbors}}$$

nodes

queries to adj matrix

i.e. $\forall\, u_1, u_2, u_3 \quad \neg\, (\underbrace{u_1 \sim u_2,\ u_2 \sim u_3,\ u_3 \sim u_1}_{\text{triangle}})$

more generally,
H-freeness for all const sized H