Lecture 9

Szemerédi's Regularity Lemma

Testing dense graph properties via SRL:

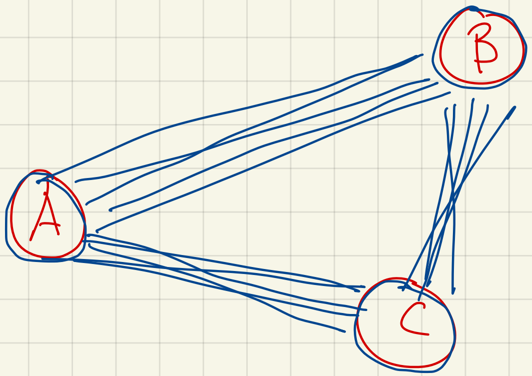
Δ-freeness

Graphs with "random" properties:

Example question:

How many triangles in a random tripartite graph?

no internal edges in A, B or C



density η

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

$$\sigma_{uvw} = \begin{cases} 1 & \text{if } uvw \text{ is } \triangle \\ 0 & \text{o.w} \end{cases}$$

$$\Pr[\sigma_{uvw} = 1] = \Pr[u \sim v \sim w \sim u] = \eta^3$$

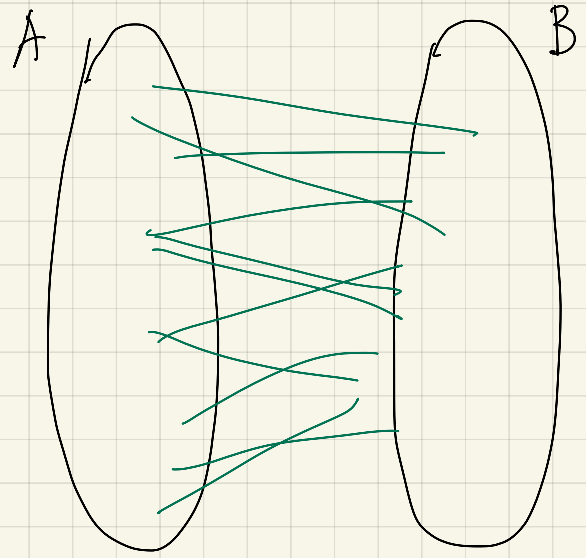
$$\| $$

$$E[\sigma_{uvw}] = \eta^3$$

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

Can we make weaker assumptions + still get reasonable bounds?

Density & Regularity of set pairs:



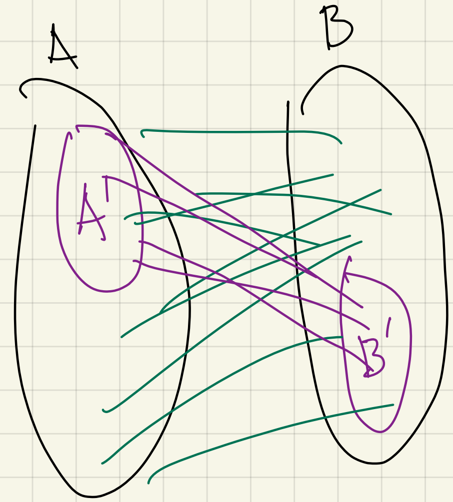
<u>def.</u> For $A, B \subseteq V$ s.t.

(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 $\forall\, A' \subseteq A, B' \subseteq B$

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

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

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

behaves like "random graph"

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

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

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

Lemma

density

regularity parameter, depends only on η

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

today assume $\eta < 1/2$

#triangles depends only on η

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

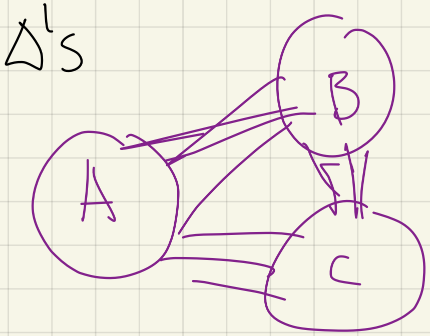
if $\eta < 1/2$

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

$\geq \eta^3/16 \cdot |A||B||C|$

with node in each of A, B, C.



Compare | for random case: $\eta^3 \cdot |A||B||C|$

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

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

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

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

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

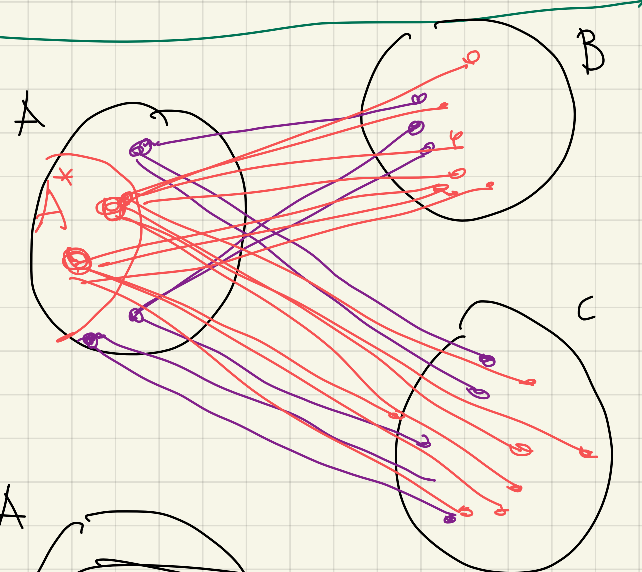
$$|d(A', B') - d(A,B)| < \gamma$$

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

"good nodes" $\qquad \geq |\eta - \gamma| \cdot |C|$ nbrs in C

not much less than expected

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



why? (Pf of claim)

$A' \leftarrow$ "bad" nodes wrt B $(< |\eta - \gamma||B|$ nbrs in $B)$

$A'' \leftarrow$ " " " C $(< |\eta - \gamma||C|$ " " $C)$

then $|A'| < \gamma |A|$ & $|A''| < \gamma \cdot |A|$

why? consider A', B

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

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

but if $|A'| \geq \gamma |A|$

then we have contradiction to (A, B) regular pair so $|A'| < \gamma |A|$

let $A^* = A \setminus (A' \cup A'')$ then $|A^*| \geq |A| - |A'| - |A''|$

$$\geq |A| - 2\gamma|A| = (1-2\gamma)|A|$$

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

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

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

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

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

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

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

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

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

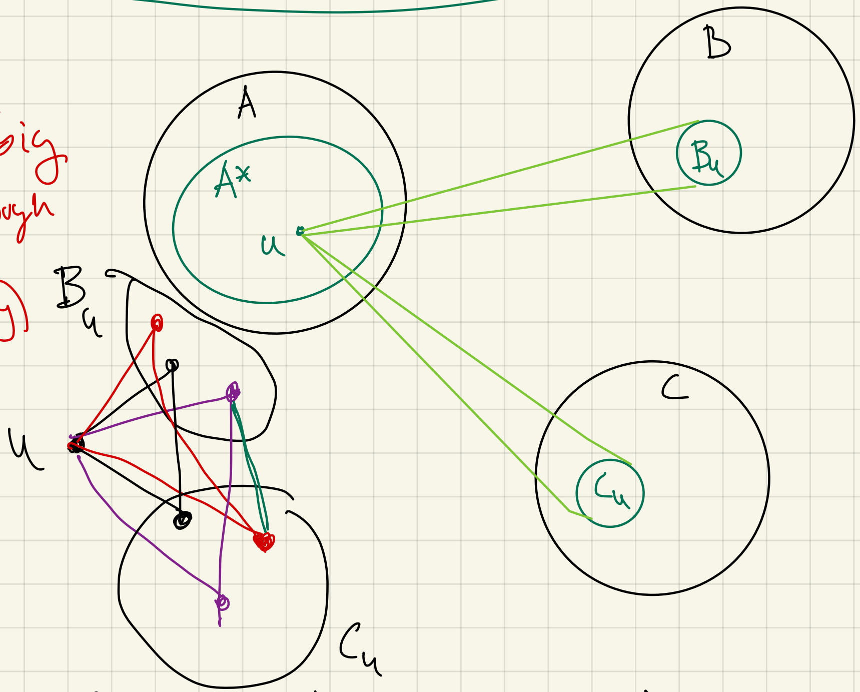
For each $u \in A^*$: define $B_u \equiv$ nbrs of u in B

$C_u \equiv$ nbrs of u in C

pretty big (big enough for regularity)

find lots of distinct Δ's using node u

edges between B_u + C_u give # distinct Δ's in which u participates



$$d(B, C) \geq \eta \Rightarrow d(B_u, C_u) \geq \eta - \gamma \Rightarrow$$

$B_u + C_u$ big enough + (B, C) is γ-regular

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

$$\geq (\eta - \gamma)^3 |B||C|$$

recall $|B_u| \geq |\eta - \gamma| \cdot |B|$

but will set $\gamma = \frac{1}{2}\eta$ so $|B_u|$ big

$$\text{total } \# \Delta\text{'s} \geq (1 - 2\gamma)|A| \cdot (\eta - \gamma)^3 |B||C|$$

$$\geq (1 - \eta)(\eta/2)^3 |A||B||C|$$

$\delta(n)$

Do interesting graphs have regularity properties?

Yes in some sense all graphs do

Can be approximated as small collection of random regular sets

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.

all pairs (V_i, V_j) are ε- regular"

will get only "most"

$\le \varepsilon \binom{k}{2}$ are not regular

to be useful

sometimes need $k \geq m$

for some m

$k=1$ & $k=n$ trivial

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 ← # is const ind of n $<< T$

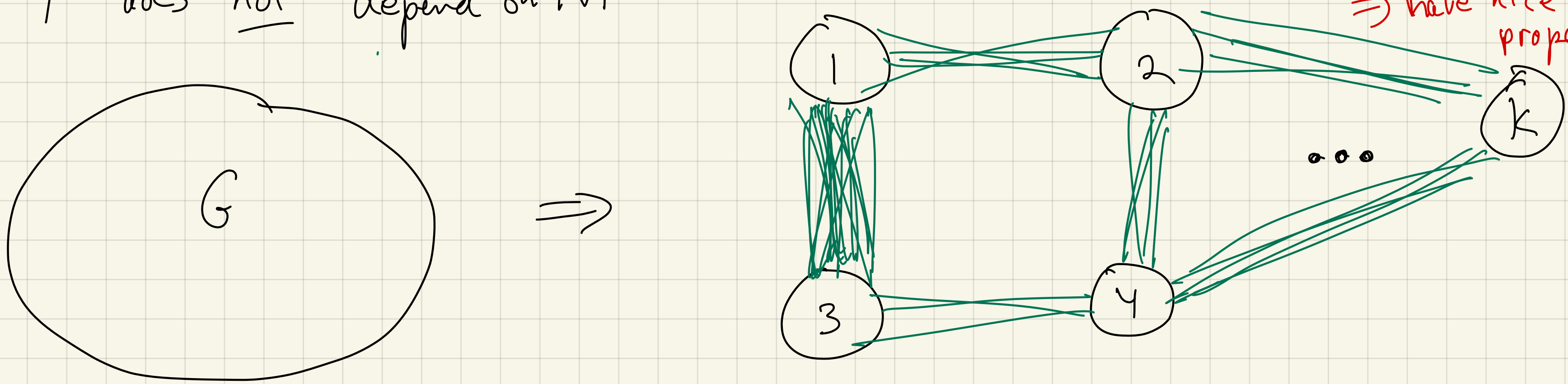
then exists equipartition $\mathcal{B}$ into k sets which refines $\mathcal{A}$

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

& $\leq \varepsilon \binom{k}{2}$ set pairs not ε-regular

const # partitions
each pairs behaves like random graph
⇒ have nice properties

Note: T does not depend on $|V|$



Why was SRL first studied?

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

Very rough idea of proof:

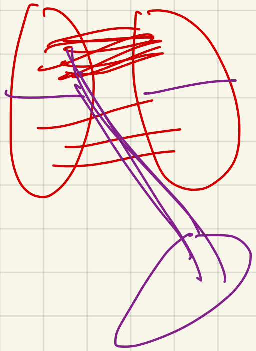
"expectation of $d^2(V_i, V_j)$"

same densities

$$I(V_1 \dots V_k) = \frac{1}{k^2} \sum_{i=1}^{k} \sum_{j=i+1}^{k} d^2(V_i, V_j) \leq \frac{1}{2}$$

"variance of d"

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



if a partition violates, can refine s.t. $I(V_1', \dots, V_k')$ grows significantly (i.e. by $\approx \varepsilon^c$)

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

so in less than $\frac{1}{\varepsilon^c}$ refinements have good partition

how big is k? each split can split into exponential subsets

upper bound $2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}$ $\frac{1}{\varepsilon^c}$ times tower of size $\frac{1}{\varepsilon^c}$

lower bound " " " $\frac{1}{\varepsilon^{c'}}$

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 $\underbrace{\varepsilon\text{-far}}_{}$ from Δ-free output FAIL

must delete $\geq \varepsilon n^2$ edges

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