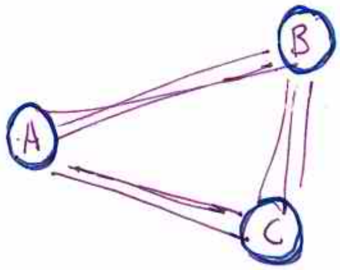
Lecture 10

- Szemerédi's Regularity Lemma

- Testing dense graph property of Δ-freeness

Graphs with "random" properties:

Example Question:
How many triangles in a random tripartite graph?



density η

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

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

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

$$E[\#\text{ triangles}] = E\Big[\sum_{\substack{u \in A \\ v \in B \\ w \in C}} \gamma_{u,v,w}\Big]$$

$$= \eta^3 |A||B||C| \quad \blacksquare$$

What weaker assumptions can we make to get similar bounds?

One possibility:
Density + regularity of set pairs:

def For $A, B \subseteq V$ s.t.

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

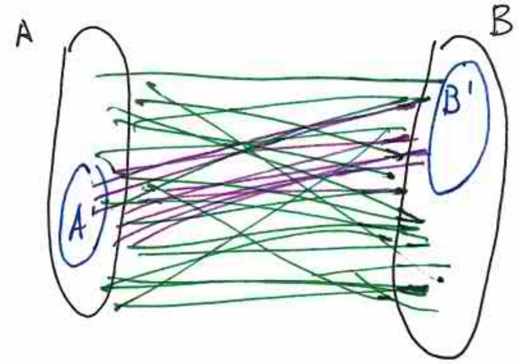
(2) $|A|, |B| > 1$

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

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

Say A,B is γ-regular if $\forall\ A' \subseteq A,\ B' \subseteq B$ s.t. $|A'| \geq \gamma |A|$, $|B'| \geq \gamma |B|$ } behaves like a "random" graph

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



So lose only factor of 16

Lemma [Komlos Simonovits] (density)

$\forall \eta > 0$ $\exists\ \gamma$ (regularity parameter, depends only on η) $= \frac{1}{2}\eta \;\; \left(= \gamma^{\Delta}(\eta)\right)$

δ (# triangles, depends only on η) $= (1-\eta)\dfrac{\eta^3}{8} \geq \dfrac{\eta^3}{16} = \delta^{\Delta}(\eta)$ (if $\eta < \frac{1}{2}$)

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 with vertex from each of A, B, C.

Proof ([Alon Fischer Krivelevich Szegedy] simplify).

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

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

Pf of Claim

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

then $|A'| \leq \gamma|A|$ (and $|A''| \leq \gamma|A|$)

why? otherwise: look at pair A', B ← size $|B| \geq \gamma|B|$

size $\geq \gamma|A|$ by assumption

by def of A'

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

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

$\rightarrow\leftarrow$ to γ-regularity

(simalarly for $|A''| \leq \gamma|A|$)

But $A^* = A \setminus (A' \cup A'')$

so $|A^*| \geq |A| - |A'| - |A''|$
$\geq |A| - 2\gamma|A| = (1-2\gamma)|A|$

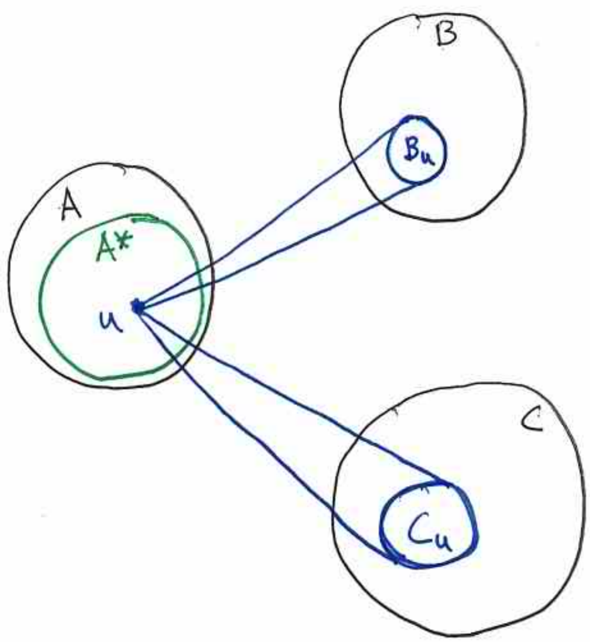
■ end-proof-of-claim

Finishing proof of Lemma:

For each $u \in A^*$:

def. $B_u \equiv$ nbrs of u in B so $|B_u| \geq (\eta - \gamma)|B| \geq \gamma |B|$

$C_u \equiv$ nbrs of u in C so $|C_u| \geq (\eta - \gamma)|C| \geq \gamma |C|$



Since γ chosen s.t. $\gamma < \frac{\eta}{2}$, we have $\eta - \gamma > \gamma$

Note: # edges between B_u + C_u $\Rightarrow$ lower bound on # distinct triangles with u as a vertex

$d(B, C) \geq \eta$

$\Rightarrow d(B_u, C_u) \geq \eta - \gamma$ (since $|B_u|, |C_u|$ big enough + B, C γ-regular)

$\Rightarrow e(B_u, C_u) \geq (\eta - \gamma)|B_u||C_u|$

$\geq (\eta - \gamma)^3 |B||C|$ ≡ l.b. on # triangles with u

$\Rightarrow$ total # Δ's $\geq (1 - 2\gamma)|A| \cdot (\eta - \gamma)^3 |B||C|$

$\geq (1 - \eta)(\eta/2)^3 \cdot |A| \cdot |B| \cdot |C| = (1 - \eta)\frac{\eta^3}{8} \cdot |A| \cdot |B| \cdot |C|$

Choosing $\gamma = \eta/2$

∎

Do any interesting graphs have regularity properties?

in some sense, <u>all</u> graphs do! i.e. every graph (in some sense) can be approximated by random graphs.

<u>Szemerédi's Regularity Lemma</u>

would like it to say:

"one can equipartition the nodes V into $V_1 \dots V_k$ (for some const k) s.t. $\underbrace{\text{all}}$ pairs (V_i, V_j) are ε-regular"

only most i.e. $\leq \varepsilon \binom{k}{2}$ don't have to be regular

k is const > 1 sometimes need $k > m$ for some m ($k=1$, $k=n$ trivial)

a more useful version:

<u>Lemma</u>

$\forall$ m, $\varepsilon > 0$ $\exists$ $T = T(m, \varepsilon)$ (huge constant) s.t.

given $G = (V, E)$ s.t. $|V| > T$

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

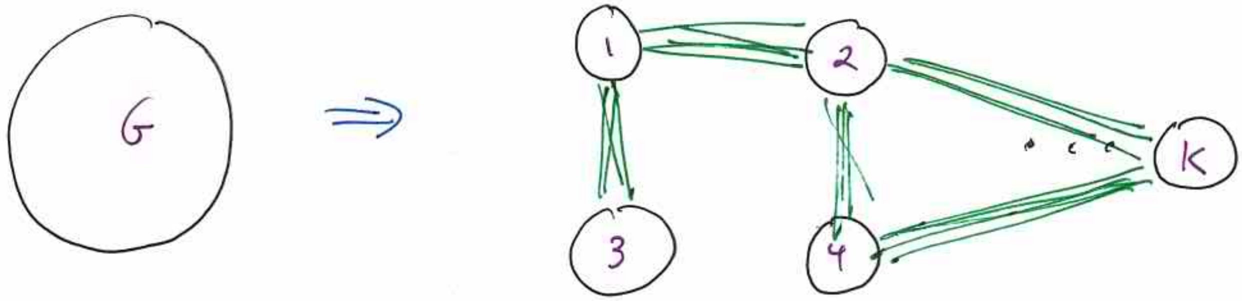
then $\exists$ equipartition β into k sets which refines $\mathcal{A}$,

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

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

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

"Picture":



Why is this good?

- partition big graph into "constant" # partitions st. each pair behaves like random bipartite graph

- random bipartite graphs have nice properties

Why was SRL first studied?

to prove conjecture of Erdös & Turán
sequences of ints must always contain long arithmetic progressions.

Very rough idea of a proof:

$$\text{ind}(V_1 \cdots V_k) = \frac{1}{k^2} \sum_{i=1}^{k} \sum_{j=i+1}^{k} d^2(V_i, V_j) \leq 1/2$$

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

How many classes (k')?

u.b.: tower of size $\frac{1}{\varepsilon}^c$: $\underbrace{2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}}_{\frac{1}{\varepsilon^c}}$

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

An application of the SRL:

Property testing of a graph: is it triangle-free?

Given: graph G, adjacency matrix format

Desired Behavior: if G is Δ-free, output PASS

if G is $\underbrace{\varepsilon\text{-far}}_{\text{must delete } \geq \varepsilon n^2 \text{ edges to make it } \Delta\text{-free}}$ from Δ-free then $\Pr[\text{output FAIL}] > 1/2$

1-sided error

How much time does this require?

trivial $O(n^3)$, $\underbrace{O(n^{\omega})}_{\text{matmult}}$, ..., $O(1)$?

Algorithm

← constant time!

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

Pick v_1, v_2, v_3

if Δ reject & halt

Accept

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

$\forall$ s.t. G is ε-far from Δ-free,
then G has $\geq \delta \binom{n}{3}$ distinct Δ's.

← note, this thm is specific to notion of ε-far from Δ-free defined above ("Adjacency matrix model")

Corollary Algorithm has desired behavior

Pf of Corr (Given Thm)

if Δ-free, accepts ✓

if ε-far,

$\geq \delta \binom{n}{3}$ Δ's

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 const c in "O" notation

Proof of Thm

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

s.t. $\frac{5}{\varepsilon} \leq k \leq T(5\varepsilon^{-1}, \varepsilon')$

\# nodes per partition

equivalently: $\frac{\varepsilon n}{5} \geq \frac{n}{k} \geq \frac{n}{T(5\varepsilon^{-1}, \varepsilon')}$

(do this by starting with arbitrary equipartition into $5/\varepsilon$ sets as $\mathcal{A}$)

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

Need: # of partitions fairly large s.t. # edges inside a partition not too big

Assume n/k is integer

$G' \equiv$ take G and

1) delete edges of G internal to any V_i

how many?

$$\le \underbrace{\frac{n}{k}}_{\text{deg w/in } V_i} \cdot \underbrace{n}_{\text{sum over all nodes}} \overset{\text{choice of } k}{\le} \frac{\varepsilon n^2}{5}$$

2) delete edges between ε'-nonregular pairs

how many?

$$\le \underbrace{\varepsilon' \binom{k}{2}}_{\text{\# non-regular pairs}} \underbrace{\left(\frac{n}{k}\right)^2}_{\text{max \# edges per pair - here we use equipartition} \Rightarrow \text{max size of } V_i \text{ is} \approx \frac{n}{k}(+1)?} \le \frac{\varepsilon}{5} \cdot \frac{k^2}{2} \cdot \frac{n^2}{k^2} \le \frac{\varepsilon}{10} \cdot n^2$$

3) delete edges between $\underbrace{\text{low density}}_{< \frac{\varepsilon}{5}}$ pairs

how many?

$$\le \sum_{\text{low density}} \frac{\varepsilon}{5} \left(\frac{n}{k}\right)^2$$

note $\sum \left(\frac{n}{k}\right)^2 \le \binom{n}{2}$

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

Total deleted edges from G: $< \varepsilon n^2$

But, G was ε-far from Δ-free

so G' must still have a Δ !!!

(the Δ must be 1) in 3 distinct $V_i\, V_j\, V_k$ since removed inter partition edges

2) regular pairs - since removed non-regular pairs

3) high density pairs - since removed low density pairs

by construction of 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 $\geq \underbrace{\frac{\varepsilon}{5}}_{\eta}$ density pairs

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

Δ-counting Lemma $\Rightarrow$

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

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

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

$\geq \dfrac{\delta^{\Delta}\left(\frac{\varepsilon}{5}\right)n^3}{\left(T\left(5\varepsilon^{-1}, \varepsilon'\right)\right)^3}$ Δ's

$\geq \delta'\binom{n}{3}$ Δ's in G'

(and thus in G)

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

$\blacksquare$

Extensions

- Komlos–Simonovits holds for all const sized subgraphs
- almost "as is" can use method to test

  all 1st-order graph properties

  $$\exists u_1 u_2 u_3 \ldots u_k \;\; \forall v_1 \ldots v_\ell \;\; R(\underbrace{u_1 \ldots u_k \, v_1 \ldots v_\ell}_{\text{defined by } \vee, \wedge, \neg \text{ nbrs}})$$

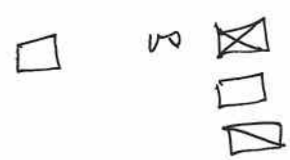
  i.e. $\forall u_1 u_2 u_3 \;\; R(\underbrace{u_1 u_2 u_3}_{\text{encodes}})$

  $\neg(u_1 \sim u_2,\ u_2 \sim u_3,\ u_1 \sim u_3)$

  H-freeness for const size H

  Induced vs forbidden

  not induced



- 1-sided const time ≈ hereditary graph props [Alon Shapira]

  closed under vertex removal (not necessarily edges)

  includes monotone graph props

  chordal
  perfect
  interval graph

  difficulty: infinite set of forbidden subgraphs also forbidden as induced.

- 2-sided const time ≈ regular partition is hardest testing problem

  property testable iff can reduce to testing [Alon Fischer Newman Shapira]

  if satisfies one of finitely many Szemerédi partitions.