Testing "Triangle Freeness" for Dense Graphs

def. G is Δ-free if $\nexists\ x,y,z$ s.t. $A(x,y)=A(y,z)=A(x,z)=1$

Claim (will prove in homework)

If there is a property testing algorithm for Δ-free-ness then there is an algorithm that works as follows:

pick random x,y,z

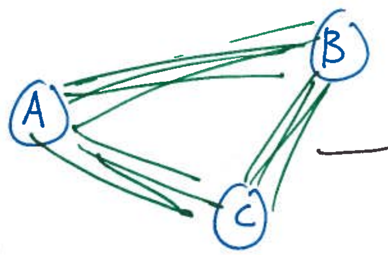
test if $A(x,y)=A(y,z)=A(x,z)=1$

But the question remains... how many times do you need to repeat the test?

Lets take a detour:

which weaker assumptions give similar bounds?

How many triangles in a random tripartite graph?



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

$$\Pr[u\sim v\sim w]=\eta^3$$

$$\sigma_{u,v,w}=\begin{cases}1 & \text{if } u\sim v\sim w\\ 0 & \text{o.w.}\end{cases}$$

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

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

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 between A & B

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

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

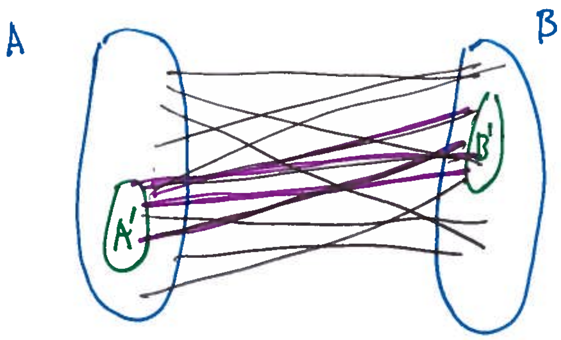
s.t. $|A'| \geq \gamma |A|$

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

behaves like a "random" graph

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

these two parameters don't have to be the same, they are here just to reduce # of parameters



lose only factor of 16!

Lemma [Komlos Simonovitz]

(density) $\forall \eta > 0$ $\exists$ γ (regularity parameter, depends only on η) $= \frac{1}{2}\eta \equiv \gamma^{\delta}(\eta)$

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

if $\eta < 1/2$

s.t. if A, B, C disjoint subsets of V & each pair is γ-regular with density $> \eta$

then G contains $\geq \delta\, |A||B||C|$ distinct Δ's with vertex from each of A, B, C

Proof (simplification of [Alon Fischer Krivelevich Szegedy])

$A^* \leftarrow$ 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 (" " " $|C|$ " " C)

then $|A'| \leq \gamma |A|$
& $|A''| \leq \gamma |A|$

Why? otherwise consider pair A', B ← trivially size $\geq \gamma|B|$ since $\gamma < 1$

size $> \gamma|A|$ by assumption for contradiction

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

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

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

Contradicts γ-regularity!

(similarly for A'')

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

so $|A^*| \geq |A| - |A'| - |A''|$
$\geq |A| - 2\gamma|A|$
$= (1-2\gamma)|A|$ ∎ (End of proof of claim)

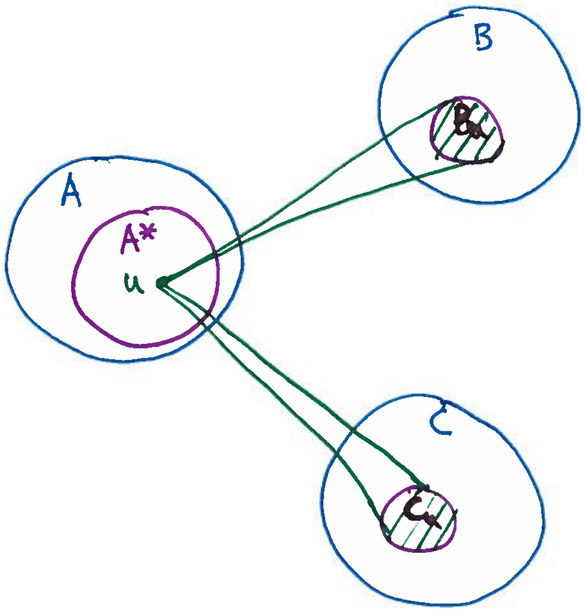
Finishing proof of lemma:

for each $u \in A^*$:

assumption on γ choice

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

$C_u \equiv$ " " " " C $\leftarrow$ so $|C_u| \geq (\eta-\gamma)|C| \geq \gamma|C|$



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

<u>Note:</u> # 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)

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

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

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

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

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.

Szemerédi's Regularity Lemma

K is const ≥ 1
may need $K > m$ for given m ($K=1, K=n$ trivial)

would like it to say:

"one can equipartition the nodes V into $V_1 \dots V_K$
(for some constant K) s.t. all pairs (V_i, V_j) are ε-regular"

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

more useful version:

Really Really huge!! a tower of 2's of size poly($1/\varepsilon$)

Lemma

huge constant, does not depend on $|V|$

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

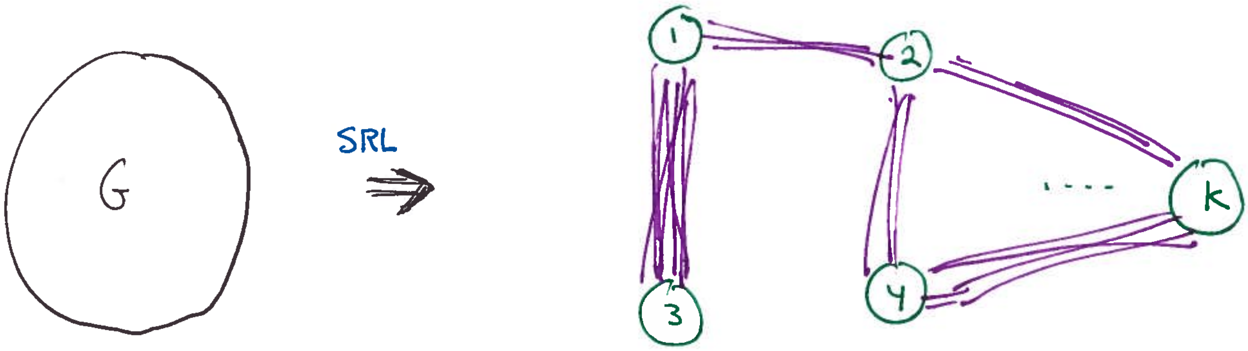
given $G = (V, E)$ s.t. $|V| > T$

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

then $\exists$ equipartition $\mathcal{B}$ into k sets which refine $\mathcal{A}$ + s.t. $m \leq k \leq T$

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

"Picture":



Why is this good?

- partition big graph into "constant" # partitions
  s.t. 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 integers must always contain long
arithmetic progressions

An application of the SRL:

Property testing!

Given G, adjacency matrix format

Desired Behavior if G is Δ-free, output PASS

if G is $\underbrace{\varepsilon\text{-far}}_{\substack{\text{must delete} \geq \varepsilon n^2 \text{ edges} \\ \text{to make it } \Delta\text{-free}}}$ from Δ-free, $\Pr[\text{output FAIL}] \geq 3/4$

How much time does this require?

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

Algorithm

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

← note this theorem is specific to notion of ε-far from Δ-free defined above "Adj matrix model"

Corollary Algorithm has desired behavior
ie. if Δ-free, accepts with prob 1 ✓

if ε-far, $\geq \delta \binom{n}{3}$ Δ's

$$\Pr[\text{don't find } \Delta \text{ in } \tfrac{c}{\delta} \text{ loops}] \leq (1-\delta)^{c/\delta}$$
$$\leq e^{-c} < 1/4$$
for big enough c

Proof of Thm

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

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

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

↖ # nodes per partition

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

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

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

slight cheat → Assume n/k is integer

Clean up G!

$G' \equiv$ take G and

1) delete edges of G internal to any V_i

how many?

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

2) delete edges between ε'-nonregular pairs ← note $\varepsilon' = \min\left(\frac{\varepsilon}{5}, \gamma^{\Delta}\left(\frac{\varepsilon}{5}\right)\right)$

how many?

$$\le \underbrace{\varepsilon' \binom{k}{2}}_{\#\text{ non-regular pairs}} \underbrace{\left(\frac{n}{k}\right)^2}_{\substack{\text{max \# edges per pair} \\ \text{here we use: equipartition} \Rightarrow |V_i| = \frac{n}{k}}} \le \frac{\varepsilon}{5} \cdot \frac{k^2}{2} \cdot \frac{n^2}{k^2} \le \frac{\varepsilon}{10} n^2$$

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

how many?

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

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

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

So total deleted edges from $G < \varepsilon n^2$ ← so cheat is not so bad

But, G was ε-far from Δ-free,

so G' must still have a Δ !!!

Furthermore, by the way we constructed G', we know a lot about the Δ: $\forall$ Δ's $a,b,c \in V_i, V_j, V_k$

1) it must be that i,j,k distinct

since removed all edges within partitions

2) (i,j) (j,k) (i,k) are regular pairs

since removed non-regular pairs

3) (i,j) (j,k) (i,k) are high density pairs

since removed low density pairs

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

$V_i V_j V_k$ all $\geq \overset{\text{“}\eta\text{”}}{\frac{\varepsilon}{5}}$-density pairs

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

Δ-counting Lemma $\Rightarrow$

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

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

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

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

$\geq \delta'\binom{n}{3}$ Δ's in G' & thus in G

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

Extensions

- Komlos- Simonovits holds for all const sized subgraphs

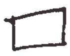
- almost "as is" can use method to test all 1st order graph properties

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

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

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

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 Fisher Newman Shapira]
  if satisfies one of finitely many Szemerédi partitions.

  see also work by [Borgs Chayes Lovasz Sos Szegedy Vesztergombi]