Lecture 8:

Testing Δ-freeness in dense graphs
Testing "Triangle Freeness" for Dense Graphs

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

Claim (will prove in homework)
If there is a properly testing algorithm for $\Delta$-freeness 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?

Let's take a detour:

How many triangles in a random tripartite graph?

\[
\begin{align*}
A & \in A, \ V \in B, \ W \in C:\ \\
\Pr\left[\text{uvw}\right] & = \eta^3 \\
E & \left[6_{uvw}\right] = \eta^3 \\
E & \left[\#\text{triangles}\right] = E\left[\sum_{\text{\tiny\#\text{triangles}}} 6_{uvw}\right] = \eta^3 \cdot 1A\|1B\|1C
\end{align*}
\]
One possibility:

Density & Regularity of set pairs:

- **def.** For \( A, B \subseteq V \) let:
  1. \( A \cap B = \emptyset \)
  2. \( |A|, |B| \geq 1 \)

Let \( e(A, B) \) = # edges between \( A \cup B \)

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

Say \( A, B \) is \( \delta \)-regular if \( \forall A' \subseteq A, B' \subseteq B \)

\[
|d(A', B') - d(A, B)| \leq \delta
\]

Lemma [Komlos-Simonovitz]:

\( \delta > 0 \)

\( \exists \gamma \) (regularity parameter depends only on \( \gamma \))

\( \delta \) (number of triangles, depends only on \( \gamma \))

\( \frac{1}{2} \gamma^3 \leq \delta \gamma \)

\( \frac{1}{16} \gamma^3 \leq \delta \gamma \)

\( \delta \leq \frac{1}{16} \)

\( \gamma \leq \sqrt{2} \)

St. if \( A, B, C \) disjoint subsets of \( V \) each pair is \( \delta \)-regular with density \( > \gamma \)

then \( G \) contains \( \geq 8 |A| |B| |C| \) distinct \( \Delta \)'s with vertex from each of \( A, B, C \)
Proof: (simplification of \cite{ok})

\[ A^* \subseteq \text{nodes in } A \text{ with } \geq (n-k) |B| \text{ nbrs in } B \]
\[ \quad \geq (n-k) |C| \text{ nbrs in } C \]

Claim: \[ |A^*| \geq (1-2\gamma) |A| \]

Proof of Claim:

Let \( A' \subseteq "\text{bad" nodes of } A \text{ wrt. } B \) (i.e., \( \leq (n-k) |B| \text{ nbrs in } B \))

Let \( A'' \subseteq "\text{" nodes of } C \) (i.e., \( \geq |C| \text{ nbrs in } C \))

Then \[ |A'| \leq \gamma |A| \]
\[ |A''| \leq \gamma |A| \]

Why? otherwise consider pair \( A', B \)

\[ d(A', B) \leq \frac{|A'| (n-k) |B|}{|A'| |B|} = n-k \]

but \( d(A', B) \geq n \)

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

Contradicts \( \gamma \)-regularity!

\( \text{Similarly for } A'' \)

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 $u \in A^*$:

\[ d \left( B_u, C_u \right) \geq \eta \]

$\Rightarrow d \left( B_u, C_u \right) \geq \eta - \delta \quad (\text{since } |B_u|, |C_u| \text{ big enough + } B, C \eta \text{-regular})$

$\Rightarrow e \left( B_u, C_u \right) \geq (\eta - \delta) |B_u||C_u|$

$\geq (\eta - \delta)^3 |B||C|$ gives lb. on # triangles with $u$ as a vertex

$\Rightarrow \text{total # } A^1_s \geq (1 - 2\delta)|A| \cdot (\eta - \delta)^3 |B||C|$

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

Choosing $\delta = \eta / 2$
Do any interesting graphs have regularity properties?

in some sense, all graphs do!

i.e., every graph (in some sense) can be approximated by random graphs.

Szemerédi's Regularity Lemma

would like it to say:

"one can equipartition the nodes $V$ into $V_1 \ldots V_k$

(for some constant $K$) s.t. all pairs $(V_i, V_j)$ are $\varepsilon$-regular"

"only most

i.e. $\leq \varepsilon(K)$

don't have to be regular"

more useful version:

Lemma

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

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

do a an equipartition of $V$ into sets

then $\exists$ equipartition $\beta$ into $k$ sets which

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

$s \leq \varepsilon(K)$ set pairs not $\varepsilon$-regular
"Picture:

\[ \text{G} \rightarrow \text{SRL} \rightarrow \text{K} \]

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 Erdős & Turán:

sequences of integers must always contain long arithmetic progressions.
An application of the SRK:

Property testing:

Given $G$, adjacency matrix format

Desired Behavior:

- if $G$ is $\delta$-free, output PASS
- if $G$ is $\varepsilon$-far from $\delta$-free, \( \Pr[\text{output}=\text{PASS}] \geq 3/4 \)
  
  must delete $\geq \varepsilon n^2$ edges
  
  to make it $\delta$-free

How much time does this require?

- trivial $O(n^3)$, $O(n^w)$, ... $O(1)$?

Algorithm:

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

  - Pick $V_1, V_2, V_3$

  - if $\Delta$ reject & halt

Accept
**Theorem**

\[ \forall \epsilon, \delta \in \mathbb{R} \quad \exists \text{ } G \text{ s.t. } |V| = n \]

\[ \text{st. } G \text{ is } \epsilon \text{-far from } \Delta \text{-free } \]

then \( G \) has \( \geq \delta(n^3) \) distinct \( \Delta \)'s

**Corollary**

Algorithm has desired behavior

ie. if \( \Delta \)-free, accepts with prob \( 1 \)

if \( \epsilon \)-far, \( \geq \delta(n^3) \) \( \Delta \)'s

\[ \Pr \left[ \text{don't find } \Delta \text{ in } \frac{\epsilon}{8} \text{ loops} \right] \leq (1 - \frac{\epsilon}{8})^{\frac{\epsilon}{8}} \]

\[ \leq e^{-\epsilon} \leq \frac{1}{2} \quad \text{for big enough } c \]

**Proof of Theorem**

Use regularity to get equipartition \( \exists V_1, \ldots, V_k \)

\[ \text{st. } \frac{5}{\epsilon} \leq k \leq \frac{1}{T(5\epsilon^{-1}, \epsilon')} \]

equivalently:

\[ \frac{En}{5} \geq \frac{n}{k} \geq \frac{n}{T(5\epsilon^{-1}, \epsilon')} \]

(nodes per partition)

(do this by starting with arbitrary equipartition into \( 5/\epsilon \) sets as \( \Delta \))

for \( \epsilon' \equiv \min \left\{ \frac{\epsilon}{5}, \sqrt[3]{\frac{\epsilon}{5}} \right\} \)

\[ \text{st. } \leq \epsilon'(\frac{k}{2}) \text{ pairs not } \epsilon'-\text{regular} \]
Need: # of partitions fairly large at # edges
inside a partition not too big

slight check: Assume \( n/K \) is integer

\[ G' = \text{take } G \text{ and } \]

1) delete edges of \( G \) internal to any \( V_i \):

how many?

\[ \leq \frac{n}{K} \cdot n \leq \frac{\varepsilon n^2}{5} \]

\([\text{choice of } K]\)

\(\text{deg } \cup \text{in } V_i \text{ sum over all } n \text{ nodes}
\)

since \( |V_i| \leq \frac{n}{K} \)

2) delete edges between \( \varepsilon' \)-nonregular pairs:

how many?

\[ \leq \varepsilon' (\frac{n}{k})^2 \leq \frac{\varepsilon'}{5} \cdot \frac{n^2}{k^2} \leq \frac{\varepsilon'}{10} n^2 \]

\(\text{# non-regular pairs max # edges per pair here we use: equipartition } \Rightarrow |V_i| = \frac{n}{k} \)

3) delete edges between low density pairs

how many?

\[ \leq \frac{\varepsilon}{5} \left( \frac{n}{k} \right)^2 \]

\(\text{low } \leq \frac{\varepsilon}{5} \)

\[ \leq \frac{\varepsilon}{5} \left( \frac{n}{k} \right)^2 \approx \frac{\varepsilon n^2}{10} \]

So total deleted edges from \( G \leq \varepsilon n^2 \)

\(\text{so chart is not so bad} \)
But $G$ was $\epsilon$-far from $\Delta$-free, so $G'$ must still have a $\Delta$!!!

Furthermore, by the way we constructed $G'$, we know a lot about the $\Delta$: $\Delta$'s abc $\in V_i, V_j, V_k$

1) it must be that $ij, jk, ik$ distinct
   since removed all edges within partitions

2) $(ij), (jk), (ik)$ are regular pairs
   since removed non-regular pairs

3) $(ij), (jk), (ik)$ are high density pairs
   since removed low density pairs

\[ \forall \delta_{ij, jk, ik} \text{ distinct st. } a \in V_i, b \in V_j, c \in V_k \]

\[ V_i, V_j, V_k \text{ all } \geq \frac{\epsilon^3}{\delta_{ij, jk, ik}} \text{-density pairs} \]

\[ \Rightarrow \frac{\epsilon}{2} \geq \frac{\epsilon}{10} \]

\[ \Delta \text{-counting Lemma } \Rightarrow \]

\[ \geq \delta_{ij, jk, ik}^{\Delta} \text{ triangles in } G' \]

where $\delta_{ij, jk, ik}^{\Delta} = (1-\gamma) \frac{\eta^3}{8}$

\[ \geq \delta_{ij, jk, ik}^{\Delta} \frac{\eta^3}{(T(\frac{\eta}{2}, \epsilon^3))} \Delta^{1/2} \]

\[ \geq \delta_{ij, jk, ik} \Delta \text{'s in } G' \text{ thus in } G \]

for $\delta' = 6 \delta_{ij, jk, ik}^{\Delta} (T(\frac{\eta}{2}, \epsilon^3))^{-3}$
Extensions

* Komlos-Simonovits holds for all constant sized subgraphs

* Almost "as is" can use method to test all 1st order graph properties

\[ \exists u_1, u_2, \ldots, u_k \forall v_1, v_2 \quad \exists (u_1, u_2, v_1, v_2) \]

defined by \( V, A, T \) neighbors

I.e., \( \forall u_1, u_2, u_3 \quad \exists (u_1, u_2, u_3) \)

encodes

\( \forall (u_1, u_2, u_3, u_4, u_5) \)

- Freeness for constant size \( H \)

\[ \square \quad \text{vs.} \quad \square \quad \square \quad \square \quad \square \]

- 1-sided constant time \( \approx \) hereditary graph props [Alan 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 constant time \( \approx \) regular partition is hardest testing problem

property testable if and only if can reduce to testing [Alan-Fisher-Neeman]

if satisfies one of finitely many Szemeredi partitions.

see also work by [Borgs, Chayes, Lovasz, Sos, Szego, Veszegy]

[Goldreich, Golovin, Ron] +...