Lecture 10

Testing dense graph properties via SRL:

$\Delta$-freeness

Begin lower bound
Density & Regularity of set pairs:

def. 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 \( d(A, B) = \frac{e(A, B)}{|A||B|} \)

Say \( A, B \) is \( \delta \)-regular if \( \forall A', B' \subseteq A, B \)
s.t. \( |A'| = \delta |A| \)
\( |B'| = \delta |B| \)

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

behave like "random graph"
Lemma

∀η > 0

∃ δ = \frac{1}{2} \eta = \eta^A(\eta)

# triangles depends only on n

\delta = (1 - \eta)^2 \frac{n^3}{8} = \frac{n^3}{16} = \delta^A(\eta)

if \eta < \frac{1}{2}

Sit. if A, B, C disjoint subsets of V st. each pair

is δ-regular with density > η

then G contains

\geq \delta \cdot |A| \cdot |B| \cdot |C| \geq \frac{n^3}{16} \cdot |A| \cdot |B| \cdot |C|

with node in each of A, B, C.

Compare for random tripartite graphs: \eta^3 \cdot |A| \cdot |B| \cdot |C|

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

A, B is δ-regular if ∀ A' ≤ A, B' ≤ B

s.t. |A'| ≥ η |A|

|B'| ≥ η |B|

| d(A',B') - d(A,B) | < δ
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 $V_1,...,V_k$ (for const $k$) s.t.

\[
\forall \text{ pairs } (V_i, V_j) \text{ are } \epsilon \text{- regular}
\]

will get "most"

\[
\leq \epsilon(k^2)
\]

are not regular.

Sometimes need $k > m$ for some $m$

$k=1$ & $k=n$ trivial.
Szemerédi's Regularity Lemma: (especially useful version)

\[ A \in \mathbb{R}, \epsilon > 0 \quad \exists \quad T = T(m, \epsilon) \quad \text{s.t.} \quad \forall G = (V,E) \quad \text{s.t.} \quad |V| > T \]

\[ + A \text{ an equipartition of } V \quad \text{into } m \text{-sets} \]

\[ \text{then exists equipartition } B \text{ into } K \text{-sets which refines } A \]

\[ m \leq K \leq T \]

\[ + \leq 3 \left( \frac{k}{2} \right) \text{ set pairs not } \delta \text{-regular} \]

Note: \( T \) does not depend on \( |V| \)

\[ \Rightarrow \]

\[ \bullet \quad \Rightarrow \]

\[ \Rightarrow \]

\[ \Rightarrow \]

\[ \Rightarrow \]

\[ \Rightarrow \]
An application of the SRL:

Given $G$ in adjacency matrix form

Is it $\Delta$-free?

desired behavior: if $G$ is $\Delta$-free, output PASS
if $G$ $\varepsilon$-far from $\Delta$-free output FAIL, with prob $\geq 3/4$

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

Algorithm:

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

Pick $V_1, V_2, V_3 \in V$
if $\Delta$ reject and halt

Accept

Just because you need to delete $\geq \varepsilon n^2$ edges, do we know that there are a lot of $\Delta$'s???
Theorem \( \forall \varepsilon, \exists \delta \) s.t. \( \forall G \) s.t. \( |V| = n \)

\( \delta \) s.t. \( G \) is \( \varepsilon \)-far from \( \Delta \)-free,
then \( G \) has \( \geq \delta \left( \frac{n}{3} \right) \) distinct \( \Delta \)'s

Corollary: Algorithm has desired behavior

Why? 
1. if \( \Delta \)-free: we never reject
2. if \( \varepsilon \)-far from \( \Delta \)-free:
   \[ \geq \delta \left( \frac{n}{3} \right) \] \( \Delta \)'s

\Rightarrow \text{each loop passes with prob } \geq 1 - \gamma

\[ \Pr[ \text{don't find } \Delta ] \leq (1 - \gamma) \]

\[ \leq e^{-c} < \frac{1}{4} \]

\Rightarrow reject with prob \( \geq \frac{3}{16} \)

\text{for proper choice of } c
**Theorem**

\[ \forall \varepsilon, \exists G \text{ s.t. } |V| = n \]

\[ \forall \varepsilon, \exists G \text{ s.t. } |V| = n \]

\[ \forall \varepsilon, \exists G \text{ s.t. } |V| = n \]

\[ \forall \varepsilon, \exists G \text{ s.t. } |V| = n \]

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

**Proof**

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

\[ \frac{s}{8} \leq k \leq 7\left(\frac{5}{\varepsilon}, \varepsilon'\right) \]

Equivalent: size of partitions

\[ \frac{3^n}{5} \geq \frac{n}{k} \geq \frac{n}{7(\varepsilon, \varepsilon')} \]

How? Start with arbitrary equipartition \( A \) into \( 5/\varepsilon \) sets

For \( \varepsilon' \equiv \min \left\{ \frac{3}{8}, \varepsilon' \right\} \)

\[ \forall \varepsilon', \exists'(\frac{k}{2}) \text{ pairs not } \varepsilon'\text{-regular} \]

\[ \forall \varepsilon', \exists'(\frac{k}{2}) \text{ pairs not } \varepsilon'\text{-regular} \]

\[ \forall \varepsilon', \exists'(\frac{k}{2}) \text{ pairs not } \varepsilon'\text{-regular} \]
Assume $\frac{n}{K}$ is integer

$G' = \text{take } G$ and

1) delete edges internal to any $V_i$ (if nodes per partition small, few internal edges)

how many? \[ \leq \frac{n}{K}, n \leq \frac{n^2}{K} \leq \frac{\varepsilon n^2}{5} \]

2) delete edges between $\varepsilon'$-non-regular pairs

how many? \[ \leq \varepsilon' \binom{K}{2} \left( \frac{n^2}{K} \right) \leq \varepsilon' \frac{K^2}{2} \frac{n^2}{K} \leq \frac{\varepsilon n^2}{10} \]

$A, B$ is $\varepsilon$-regular if $A' \leq A, B' \leq B$

s.t. $|A'| \geq |A|, |B'| \geq |B|$

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

$\frac{\delta(n)}{\bar{\delta}(\eta)} \leq \frac{\frac{1}{2} \eta}{\bar{\delta}(\eta)}$

$|\delta(n)| \leq (1 - \eta) \frac{\eta^2}{8} \leq \frac{\eta^3}{16}$

$\varepsilon' = \min \varepsilon' \left( \frac{n}{K}, \frac{\varepsilon}{5} \right)$

$\delta \leq \varepsilon' \left( \frac{K}{5} \right)$ pairs not $\varepsilon'$-regular
3) delete edges between low density pairs

how many?

\[ \leq \sum_{\text{low density pairs}} \left( \frac{\epsilon}{5} \right) \left( \frac{n^2}{K} \right) \]

\[ \leq \frac{\epsilon}{5} \binom{n}{2} \approx \frac{\epsilon n^2}{10} \]

Total deleted edges:

\[ \leq \frac{3n^2}{5} + \frac{3n^2}{10} + \frac{3n^2}{10} < 3n^2 \]

But \( G \) is \( \epsilon \)-far from \( \Delta \)-free (must delete \( \geq \epsilon n^2 \) edges to remove all \( \Delta \)'s)

so \( G' \) must still have a triangle!
\[ \Delta \text{ in } G' \text{ must connect:} \]

1) nodes in 3 distinct \( V_i, V_j, V_k \) since deleted all internal edges.

2) regular pairs since deleted all edges between irregular pairs.

3) high density pairs since deleted all edges between low density pairs.

\[ \exists \ ij, jk \text{ distinct s.t. } x \in V_i, y \in V_j, z \in V_k \]

\[ V_i, V_j, V_k \text{ all } \geq \frac{\varepsilon^5 n}{3} \text{ density pairs} \]

\[ \Rightarrow \exists A \left( \frac{\varepsilon}{5} \right) - \text{regular} \]

\[ \Rightarrow \frac{M}{2} \geq \frac{\varepsilon}{10} \]
$$\Delta$$-counting lemma \(\Rightarrow\)

\[ n \geq \delta^A \left( \frac{\epsilon}{8} \right) \frac{1}{V_{\xi, V_{\lambda}, V_k}} \quad \text{triangles in } G', \text{ where } \delta^A = (1 - \eta) \frac{n^3}{8} \geq 2 (\epsilon^3) \]

\[ n \geq \delta^1 \left( \frac{\epsilon}{3} \right) \Delta^1 \text{ in } G', \text{ thus in } G \]

For \( \delta^1 = \delta^A \left( \frac{\epsilon}{8} \right) \frac{1}{1 \left( \frac{\epsilon}{8}, \epsilon' \right)} \)

\[ G \Rightarrow G' \]

\(G\) not-regular

\(G'\) too sparse
This is a powerful technique!

- Similar lemma to \( \Delta \)-counting holds for all constant sized subgraphs.

- Almost "as is" can use same method to test all "1st order" graph properties:

\[
\exists u_1, u_2, u_3 \ldots u_K \quad \forall V_1 \ldots V_K \quad R(u_1 \ldots u_K, V_1 \ldots V_K) \quad \text{queries to adj matrix}
\]

\[
R \text{ defined via } \forall V, \text{1-4 neighbors}
\]

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

More generally, triangle freeness for all constant sized \( H \)
For dense graphs, testable properties

• 1-sided error const time $\approx$ hereditary graph properties (closed under vertex removal: chordal, perfect, interval)

  difficulty = infinite set of forbidden subgraphs

• 2-sided error const time $\approx$ any property that can be reduced to testing if satisfies one of finite # of Szemeredi partitions

Maybe the reason that the dependence on $\varepsilon$ is so bad is that this technique is too "general purpose"?
Maybe specific properties (e.g., $\Delta$-freeness) have better testers?
An intriguing characterization of bipartite graphs:

For graphs in adjacency matrix model:

- If $H$ bipartite, poly($\frac{1}{\varepsilon}$) is enough
- If $H$ not bipartite, no poly($\frac{1}{\varepsilon}$) suffices

we will prove for $H = \Delta$
is a terrible dependence
on \( \epsilon \) required?
is there a better algorithm?
even just for testing \( \Delta \)-freeness?

Lower bounds for testing

\[ \Delta \text{-freeness}: \]

No, superpoly dependence on \( \epsilon \) required!

i.e., \( \geq \left( \frac{c}{\epsilon} \right)^{c \log \left( \frac{c}{\epsilon} \right)} \) for some const \( c \)
Main tool #1: Additive number theory lemma

\[ \text{Lemma} \quad \forall m, \exists X \leq M = \{1, 2, \ldots, m^2\} \text{ of size } \geq m^{1+\frac{1}{625}} \]

with no nontrivial soln to

\[ X_1 + X_2 = 2X_3 \]

no three evenly spaced points \[ X_3 = \frac{X_1 + X_2}{2} \]

will use to construct graphs which are
(1) far from \( \Delta \)-free
(2) any algorithm needs lots (in terms of) queries to find \( \Delta \)
Lemma. \( \forall m, \exists X < M = \{1, 2, \ldots, m^2\} \) of size \( \geq \frac{m}{e^{10 \sqrt{\log m}}} \)

with no nontrivial soln to

\[ X_1 + X_2 = 2X_3 \]

Examples:

Bad \( X \): \( \{1, 2, 3, 3\} \)
\( \{5, 9, 13, 3\} \)

Good \( X \): \( \{1, 2, 4, 5, \times, \times, \times, 10, \ldots, 3\} \leftarrow \text{how big?} \)
\( \{1, 2, 4, 8, 16, 32, \ldots, 3\} \leftarrow \text{only size} \log m \)
Proof. Let $d$ be integer

$$k \left\lfloor \frac{\log m}{\log d} \right\rfloor - 1$$

Define $X_B = \{ x = \sum_{i=0}^{k} x_i d^i \mid x_i < \frac{d}{2} \text{ for } 0 \leq i \leq k \text{ and } \sum_{i=0}^{k} x_i^2 = \frac{B}{d^2} \}$

View $x \in M$ in base $d$ representation $X = (x_0 x_1 \ldots x_k)$

Claim. $X_B \leq M$

Pick $B$ s.t. $|X_B|$ maximized:

how big can $B$ be?

how small can $\sum |X_B|$ be?

So $\exists B$ s.t. $|X_B| \geq \ldots$

Lemma. $\forall m$, $\exists X \leq M = \{ x_1, \ldots, m^2 \}$ of size $\geq \frac{m}{e^{0.5 \log m}}$

with no nontrivial soln to $x_1 + x_2 = 2x_3$