LECTURE 16

Last time

• Testing triangle-freeness
• Testing other properties of dense graphs
• Behrend’s construction

Today

• Lower bound for testing triangle-freeness
• Canonical testers for the dense graph model

Project progress reports due today on Gradescope
Testing Triangle-Freeness

**Input:** parameters $\varepsilon, n$, access to undirected graph $G = (V, E)$ represented by $n \times n$ adjacency matrix.

**Goal:** Accept if $G$ has no triangles; reject w.p. $\geq \frac{2}{3}$ if $G$ is $\varepsilon$-far from triangle-free (at least $\varepsilon \binom{n}{2}$ edges need to be removed to get rid of all triangles).

- **[Alon Fischer Krivelevich Szegedy 09]:** Time that depends only on $\varepsilon$

- **Today**

**Lower Bound for Testing Triangle-Freeness** [Alon 02]

Testing triangle-freeness with 1-sided error requires super-polynomial dependence on $1/\varepsilon$.

$$\Omega \left( \left( \frac{c}{\varepsilon} \right)^c \log \frac{c}{\varepsilon} \right)$$

queries for some $c > 0$
Canonical Tester for Dense Graphs

 Canonical Tester (Input: $\varepsilon, n$; query access to adjacency matrix of $G=(V,E)$)

1. Sample $s$ nodes uniformly at random.
2. Query all pairs of sampled nodes.
3. **Accept** or **reject** based on available information.

- Consider any property $\mathcal{P}$ of graphs that does not depend on the names of the nodes. That is, if $G \in \mathcal{P}$ and $G'$ is isomorphic to $G$ then $G' \in \mathcal{P}$.

**Exercise**: Show that if there is an $\varepsilon$-tester $T$ for $\mathcal{P}$ with query complexity $q(\varepsilon,n)$, then there is a canonical $\varepsilon$-tester $T'$ for $\mathcal{P}$ with query complexity $O(q^2(\varepsilon,n))$. Moreover, if $T$ has 1-sided error, so does $T'$.

A lower bound $q$ for canonical tester implies a lower bound $\sqrt{q}$ for every tester.

Sufficient to prove our lower bound $\Omega \left( \frac{c}{\varepsilon} c \log \frac{c}{\varepsilon} \right)$ for 1-sided error canonical testers.
Exercise: Show that if there is an $\epsilon$-tester $T$ for $\mathcal{P}$ with query complexity $q(\epsilon,n)$, then there is a canonical $\epsilon$-tester $T'$ for $\mathcal{P}$ with query complexity $O(q^2(\epsilon,n))$. Moreover, if $T$ has 1-sided error, so does $T'$. 
Exercise: Show that if there is an $\varepsilon$-tester $T$ for $\mathcal{P}$ with query complexity $q(\varepsilon,n)$, then there is a canonical $\varepsilon$-tester $T'$ for $\mathcal{P}$ with query complexity $O(q^2(\varepsilon,n))$. Moreover, if $T$ has 1-sided error, so does $T'$. 
**Goal for Proving the Lower bound**

- A 1-sided error tester can reject only if it finds a triangle.
- Suppose we construct a graph $G$ that is $\varepsilon$-far from being triangle free, where $p$ fraction of triples are triangles for some small $p$.
- Consider a canonical tester $T$ that samples $q$ vertices.
- Let $X$ be the number of triangles the tester catches.

$$\mathbb{E}[X] = p \binom{q}{3} = \Theta(p \cdot q^3)$$

- Suppose $q$ is set so that $\mathbb{E}[X] \leq 1/2$
- By Markov, $\Pr[T \text{ rejects } G] \leq \Pr[X \geq 1] \leq \mathbb{E}[X] \leq \frac{1}{2} < \frac{2}{3}$
- So, for $T$ to reject with high enough probability, $q = \Omega\left(p^{-\frac{1}{3}}\right)$

Sufficient to ensure $p = O\left(\left(\frac{\varepsilon}{c}\right)^c \log \frac{c}{\varepsilon}\right)$
### Behrend’s Theorem

For all integer $m \geq 1$, there exists a set $S \subseteq [m]$ such that $|S| \geq \frac{m}{8 \sqrt[3]{\log_2 m}}$ and the only solution to $x + y = 2z$ for $x, y, z \in S$ is $x = y = z$.

- We will use such a set $S$ to construct a graph that is
  - far from triangle free
  - has relatively few triangles
Initial Graph Construction

- Let $S \subset [m]$ be a set from Behrend’s Thm
- We construct a tripartite graph with $m$, $2m$, and $3m$ nodes in the three parts
- Intended triangles

\[ i \quad i + s \quad i + 2s \]

- No other triangles:
  If $(i, i + x, i + x + y)$ is a triangle, then $x \in S$, $y \in S$, and $x + y = 2z$ for $z \in S$
  But then $x = y = z$ by construction of $S$
- All triangles are edge disjoint: each edge participates in exactly one triangle.
Parameters of the Initial Construction

- Number of nodes, $n$
  
  $6m$

- Number of edges
  
  $3m \cdot |S|$

- Number of (edge-disjoint) triangles, $T$
  
  $m \cdot |S|$

- Distance to triangle-freeness

  Necessary and sufficient to remove one edge from each triangle, because they are edge-disjoint.

\[
\frac{T}{\binom{n}{2}} = \Theta \left( \frac{m \cdot |S|}{m^2} \right) = \Theta \left( \frac{|S|}{m} \right) = \Theta \left( \frac{1}{8 \sqrt{\log m}} \right)
\]

Not constant!
**Blowup of a Graph**

To construct a $b$-**blowup** of a graph,

- make $b$ copies of each node;
- make two copies (of different nodes) adjacent iff their originals are adjacent.

![Blowup Diagram](image-url)
Parameters of the Blowup Construction

- Number of nodes, $n$
  \[ 6mb \]
- Number of edges
  \[ 3m \cdot |S| \cdot b^2 \]
- Number of triangles
  \[ m \cdot |S| \cdot b^3 \]
- Number of (edge-disjoint) triangles, $T$
  \[ m \cdot |S| \cdot b^2 \]
- Distance to triangle-freeness
  \[ \frac{T}{\binom{n}{2}} = \Theta \left( \frac{m \cdot |S| \cdot b^2}{(m \cdot b)^2} \right) = \Theta \left( \frac{|S|}{m} \right) = \Theta \left( \frac{1}{8 \sqrt{\log m}} \right) \]
- Given $\varepsilon$ and $n$, pick $m$ so that $\varepsilon = \Theta \left( \frac{1}{8 \sqrt{\log m}} \right)$ and $b = \frac{n}{6m}$
- Fraction of triples that are triangles:
  \[ \approx \frac{m \cdot |S| \cdot b^3}{n^3} \approx \frac{m \cdot |S|}{m^3} = \frac{|S|}{m^2} = \frac{\varepsilon}{m} < \frac{1}{m} \]
Conclusion: Triangle-Freeness

- The query complexity of testing triangle-freeness with 1-sided error depends only on $\varepsilon$ (and is independent of the size of the graph).

- However, the dependence is super-polynomial in $1/\varepsilon$. 