LECTURE 15

Last time

- Testing triangle-freeness
- Regularity Lemma

Today

- Testing triangle-freeness
- Triangle-removal lemma
- Testing other properties of dense graphs
- Behrend’s construction

Project progress reports due Thursday
Sign up for project meetings

Testing Triangle-Freeness

Input: parameters ε, 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 ε-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 ε

Tester

Algorithm (Input: ε, n; query access to adjacency matrix of $G=(V,E)$)

1. Repeat s times:
2. Sample vertices v_1, v_2, v_3 uniformly at random
3. **Reject** if they form a triangle.
4. **Accept**.

How many repetitions suffice?

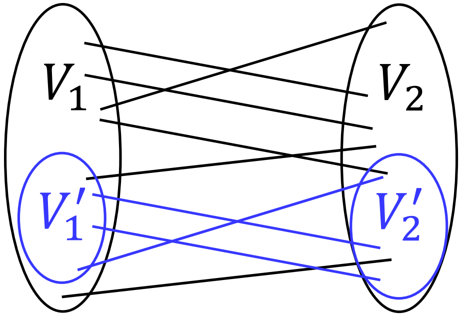
Triangle-Removal Lemma

$\forall \varepsilon \exists \delta = \delta(\varepsilon)$ such that every n-node graph that is ε-far from triangle-free contains at least $\delta \cdot \binom{n}{3}$ triangles.

- It is easy to see that if G is ε-far from triangle-free then it has at leas $\varepsilon \binom{n}{2}$ triangles. This is asymptotically better.
- By Witness Lemma, setting $s = 2/\delta$ yields a tester.

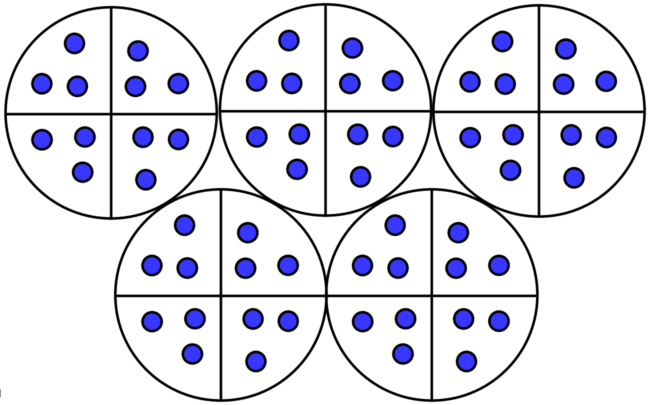
Definitions from Last Lecture

- The edge density of the pair (V_1, V_2), denoted $d(V_1, V_2)$, is $\frac{|e(V_1,V_2)|}{|V_1|\cdot|V_2|}$.



- A pair (V_1, V_2) of disjoint subsets of vertices is γ-regular if $\forall V_1' \subseteq V_1, V_2' \subseteq V_2$, such that $|V_1'| > \gamma|V_1|$ and $|V_2'| > \gamma|V_2|$,
  $|d(V_1, V_2) - d(V_1', V_2')| < \gamma$.

- An equipartition of a graph is a partition of its vertices into sets that differ in size by at most 1.



- A partition $\mathcal{B}$ is a refinement of a partition $\mathcal{A}$ if every set in $\mathcal{B}$ is a subset of set in $\mathcal{A}$.

Regularity Lemma

Every large graph G has an equipartition where

- (almost) all pairs of sets are regular,
- the number of parts is not too large.

Regularity Lemma [Szemerédi 78]

$\forall a, \forall \gamma > 0, \exists T = T(a, \gamma)$ such that if G is a graph with more that T nodes and $\mathcal{A}$ is an equipartition of G into a sets then there is an equipartition $\mathcal{B}$ of G into b sets which is a refinement of $\mathcal{A}$ satisfying:

1. $a \le b < T$;
2. at most $\gamma \binom{b}{2}$ pairs of sets in $\mathcal{B}$ are not γ-regular.

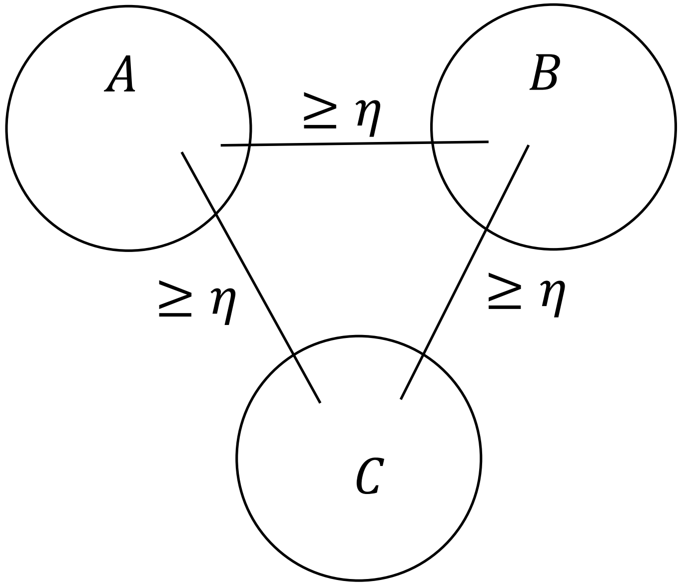
Important: T does not depend on the size of the graph

- But the dependence of T on γ is a tower $2^{2^{\cdots 2}}$ of height $\text{poly}\left(\frac{1}{\gamma}\right)$

Triangles in a Graph with Three Regular Pairs

Lemma [Kolmos Simonovits]

$\forall\eta$>0, if A, B, C are disjoint subsets of V and each pair of them is γ^{Δ}-regular with density at least η then G contains at least $\delta^{\Delta}\,|A|\cdot|B|\cdot|C|$ triangles, where $\gamma^{\Delta}=\gamma^{\Delta}(\eta)=\frac{\eta}{2}$ and $\delta^{\Delta}=\delta^{\Delta}(\eta)=\frac{1}{8}(1-\eta)\eta^{3}$.



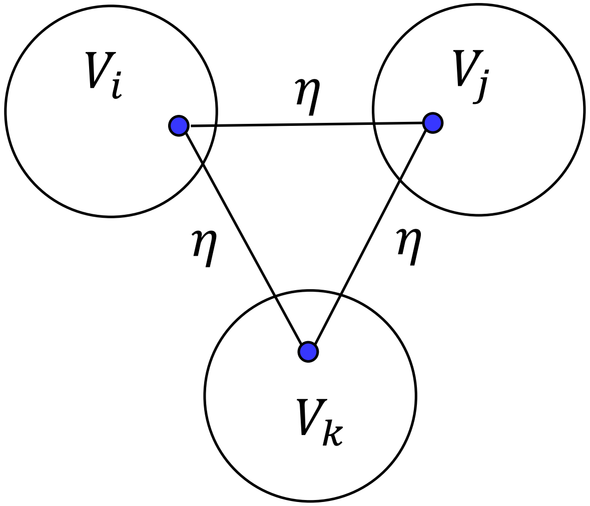
Proof of the Triangle-Removal Lemma: Idea

Triangle-Removal Lemma

$\forall \varepsilon \exists \delta = \delta(\varepsilon)$ such that every n-node graph that is ε-far from triangle-free contains at least $\delta \cdot \binom{n}{3}$ distinct triangles.

Main Idea: Consider a graph G which is ε-far from being triangle-free.

- We apply the Regularity Lemma to get a regular partition.
- We carefully remove fewer than $\varepsilon \binom{n}{2}$ edges, and show that there remains a triangle consisting of edges between regular dense pairs.
- We apply [Kolmos Simonovits] to get many triangles.



Proof of the Triangle-Removal Lemma

Triangle-Removal Lemma

$\forall \varepsilon \exists \delta = \delta(\varepsilon)$ such that every n-node graph that is ε-far from triangle-free contains at least $\delta \cdot \binom{n}{3}$ distinct triangles.

Proof: Consider a graph G which is ε-far from being triangle-free.

- Start with an equipartition $\mathcal{A}$ of G with $4/\varepsilon$ sets.

Apply the regularity lemma with $a = 4/\varepsilon$ and $\gamma = \min(\varepsilon/4, \gamma^{\Delta}(\varepsilon/4)) = \varepsilon/8$

- By Regularity Lemma, $\mathcal{A}$ can be refined into equipartition $\mathcal{B}=\{V_1, \dots, V_b\}$:
  1. $\frac{4}{\varepsilon} \leq b \leq T$
  2. at most $\gamma \cdot \binom{b}{2}$ pairs among $V_1, \dots, V_b$ are not γ-regular

$|V_i| = \frac{n}{b} \in \left[\frac{n}{T}, \frac{\varepsilon n}{4}\right]$ for all $i \in [b]$

- An edge (u, v), where $u \in V_i$ and $v \in V_j$ is useful if it satisfies:
  1. $i \neq j$
  2. (V_i, V_j) is γ-regular
  3. the density $d(V_i, V_j) \geq \varepsilon/4$

Claim. Graph G has less than $\varepsilon \binom{n}{2}$ non-useful edges.

Proof of Claim

- An edge (u, v), where $u \in V_i$ and $v \in V_j$ is useful if it satisfies:
  1. $i \neq j$
  2. (V_i, V_j) is γ-regular
  3. the density $d\left(V_i, V_j\right) \geq \varepsilon/4$

Claim. Graph G has less than $\varepsilon \binom{n}{2}$ non-useful edges.

Edges violating	Number of such edges
Condition 1	
Condition 2	
Condition 3	

Total: $\frac{7\varepsilon}{8} \cdot \binom{n}{2} < \varepsilon \binom{n}{2}$

Proof of the Triangle-Removal Lemma

Triangle-Removal Lemma

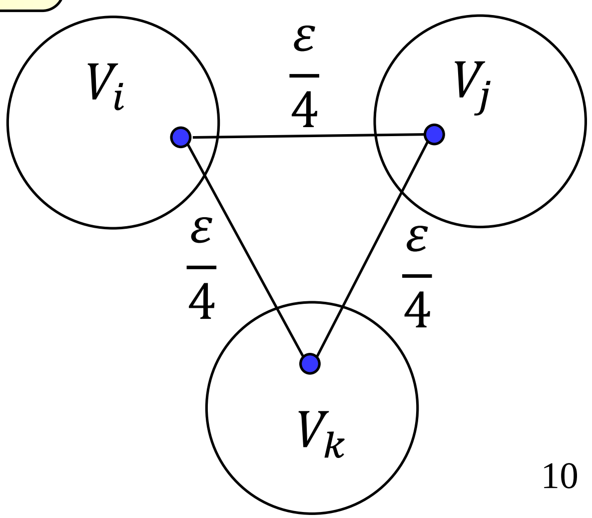
$\forall \varepsilon \exists \delta = \delta(\varepsilon)$ such that every n-node graph that is ε-far from triangle-free contains at least $\delta \cdot \binom{n}{3}$ distinct triangles.

Proof: Consider a graph G which is ε-far from being triangle-free.

- An edge (u, v), where $u \in V_i$ and $v \in V_j$ is useful if it satisfies:
  1. $i \neq j$
  2. (V_i, V_j) is $\varepsilon/8$-regular
  3. the density $d(V_i, V_j) \geq \varepsilon/4$

Claim. Graph G has less than $\varepsilon \binom{n}{2}$ non-useful edges.

Triangle of useful edges



- When we remove all non-useful edges, there is still a triangle!
- By [Kolmos Simonovits], there are at least

$$\delta^{\Delta}\left(\frac{\varepsilon}{4}\right) \cdot |V_i| \cdot |V_j| \cdot |V_k| \geq \frac{1}{8}\left(1 - \frac{\varepsilon}{4}\right)\left(\frac{\varepsilon}{4}\right)^3 \cdot \frac{n^3}{T^3}$$

triangles.

Testing Other Properties

Testing Subgraph-Freeness [Alon 02]

Let H be a fixed graph on h nodes.

Let $\mathcal{P}_H$ be the property that G does not contain a copy of H as a subgraph.

1. If H is bipartite:
   - There is a 2-sided error tester for $\mathcal{P}_H$ with $O\left(\frac{1}{\varepsilon}\right)$ queries.
   - There is a 1-sided error tester for $\mathcal{P}_H$ with $O\left(h^2\left(\frac{1}{2\varepsilon}\right)^{h^2/4}\right)$ queries.

Polynomial in $1/\varepsilon$ for fixed H.

2. If H is not bipartite, then there exists c > 0, such that every 1-sided error tester for $\mathcal{P}_H$ makes $\Omega(\left(\frac{c}{\varepsilon}\right)^{c \log \frac{c}{\varepsilon}})$ queries.

Super-polynomial in $1/\varepsilon$.

- We will prove part (2) for triangles.

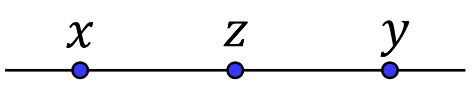
Main Tool for Proving the Lower Bound

Dense Sets of Integers with no Arithmetic Progression

Behrend's Theorem

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

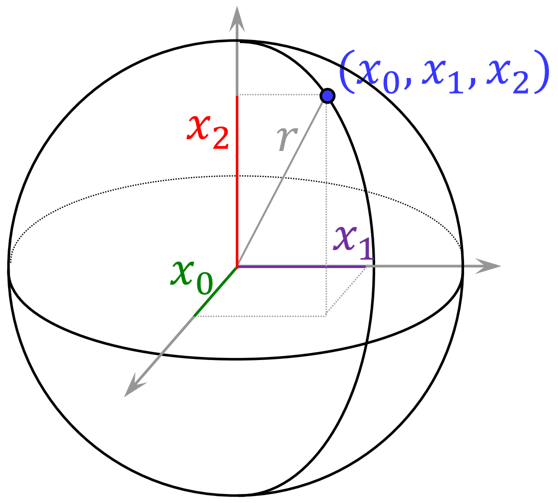
- Behrend's bound [Behrend 46] is slightly better.



- The best known is $\Omega\left(\frac{m}{2^{2\sqrt{2}\sqrt{\log_2 m}}} \log_2^{1/4} m\right)$ [Elkin 10]

Proof idea: Represent integers in $[m]$ as k-digit numbers base d, where k and d are parameters.

- For a number x, view its digits as coordinates of a point $(x_0, x_1, \ldots, x_{k-1})$
- Pick points that lie on the same sphere: i.e., with fixed $x_0^2 + x_1^2 + \cdots + x_{k-1}^2$
- Then no three of them lie on the same line, which ensures that no point is the average of two other points.



Proof of Behrend's Theorem

Behrend's Theorem

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

Proof: For an integer $B > 0$, define a set

$$S_B = \left\{ \sum_{i=0}^{k-1} x_i d^i : \text{each } x_i \in \left\{0, \dots, \frac{d}{2} - 1\right\} \text{ and } B = \sum_{i=0}^{k-1} x_i^2 \right\}$$

- All numbers in sets S_B are less than d^k.

  We set $d^k = m$ to ensure $S_B \subseteq [m]$ $\forall B$.

Claim

For all B, the only solution to $x + y = 2z$ for $x, y, z \in S_B$ is $x = y = z$.

Proof of Claim

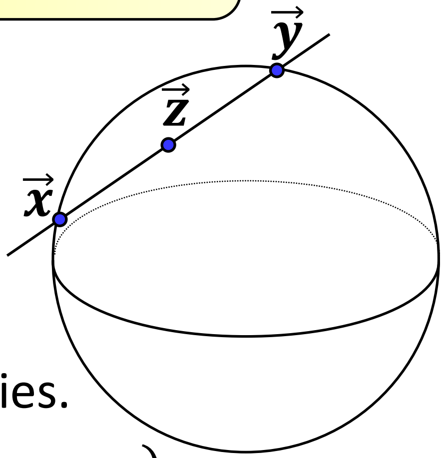
For an integer $B > 0$, define a set

$$S_B = \left\{ \sum_{i=0}^{k-1} x_i d^i : \text{each } x_i \in \left\{0, \ldots, \frac{d}{2} - 1\right\} \text{ and } B = \sum_{i=0}^{k-1} x_i^2 \right\}$$

Claim

For all B, the only solution to $x + y = 2z$ for $x, y, z \in S_B$ is $x = y = z$.

Proof: Suppose $x + y = 2z$ for some $x, y, z \in S_B$.



- Representing x, y, z base d, we get

$$\sum_{i=0}^{k-1} x_i d^i + \sum_{i=0}^{k-1} y_i d^i = 2 \sum_{i=0}^{k-1} z_i d^i$$

- Since x_i, y_i, z_i are less than $d/2$ for all i, there are no carries.
  That is, $(x_0, x_1, \ldots, x_{k-1}) + (y_0, y_1, \ldots, y_{k-1}) = 2(z_0, z_1, \ldots, z_{k-1})$
  But these three points are on a sphere,
  so one can be the average of the other two only if they are identical.

Proof of Behrend's Theorem: Setting Parameters

Behrend's Theorem

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

Proof: For an integer $B > 0$, define a set

$$S_B = \left\{ \sum_{i=0}^{k-1} x_i d^i : \text{each } x_i \in \left\{0, \dots, \frac{d}{2} - 1\right\} \text{ and } B = \sum_{i=0}^{k-1} x_i^2 \right\}$$

- Set $d^k = m$ and $d = 2^{\sqrt{1/2 \cdot \log m}}$. Then $k =$
- How many possibilities for B?
- How many numbers are in all sets S_B?
- By Pigeonhole Principle, at least one of the sets has size at least