Sublinear Algorithms

LECTURE 14

Last time

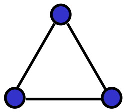
- Approximate Max-Cut

Today

- Testing triangle-freeness
- Regularity Lemma

Project progress reports are due Thursday
Sign up for project meetings

Recall

- We discussed testing bipariteness.
- A graph is bipartite iff it has no odd cycles.
- In particular, a bipartite graph has no triangles. 

Today: Testing triangle-freeness

(a special case of [Alon Fischer Krivelevich Szegedy 09])

Main tool: Regularity Lemma [Szemerédi 78]

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).

- [AFKS09]: 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?

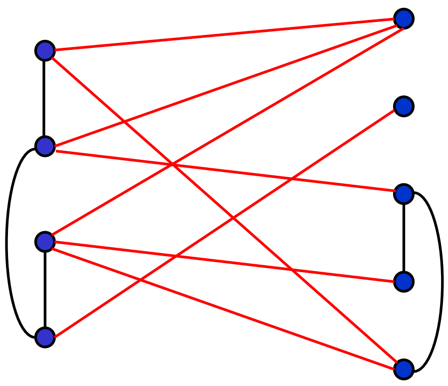
Triangle-Removal Lemma

$\forall\varepsilon\exists\delta = \delta(\varepsilon)$ 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. The lemma is asymptotically better in n.
- By Witness Lemma, setting $s = \frac{2}{\delta}$ yields a tester.

The Regularity Lemma: Density

- Let V_1, V_2 be nonempty disjoint subsets of V.
  $e(V_1, V_2)$ = set of edges between V_1 and V_2



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

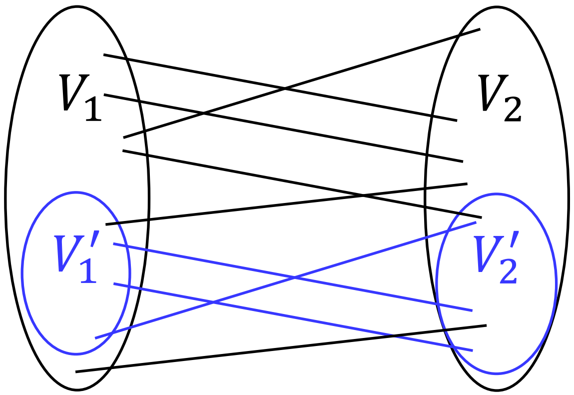
The probability that a random pair of nodes from different sets is an edge.

- This is the same definition as in the last lecture, except for normalization, generalized to non-partitions.

Regularity

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 $|V_2'| > \gamma|V_2|$,

$$|d(V_1, V_2) - d(V_1', V_2')| < \gamma.$$



If the subsets are large then the set pair and the subset pair have similar densities

We expect subsets in a random graph to have this property.

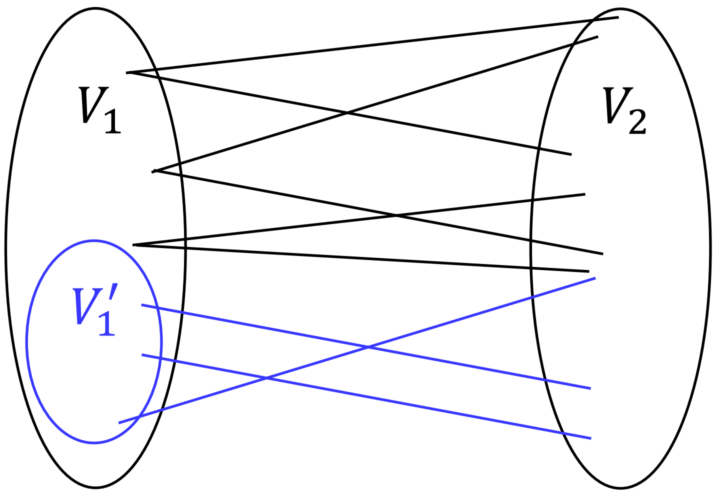
Connections in Regular Pairs

Claim (Most nodes in regular pairs have many neighbors)

Suppose (V_1, V_2) is a γ-regular pair of density $\geq \eta$.

Consider the set V_1' of nodes in V_1, each of which has at most $(\eta - \gamma)|V_2|$ neighbors in V_2. Then $|V_1'| < \gamma|V_1|$.

Proof:



$d(V_1', V_2) =$

$d(V_1, V_2) \geq \eta$

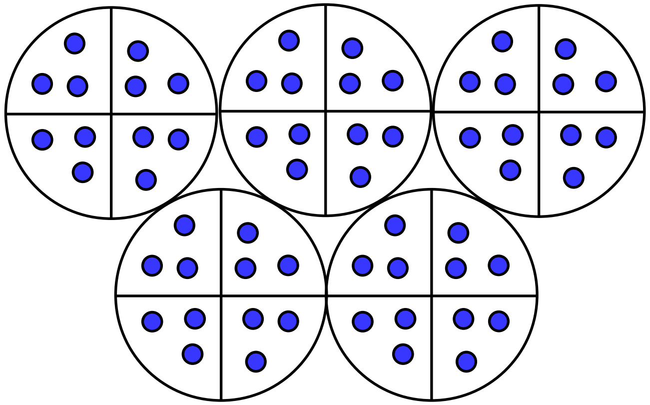
$|d(V_1, V_2) - d(V_1', V_2)| \geq$

By γ-regularity of (V_1, V_2), we conclude that $|V_1'| < \gamma|V_1|$

Equipartions

- 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:

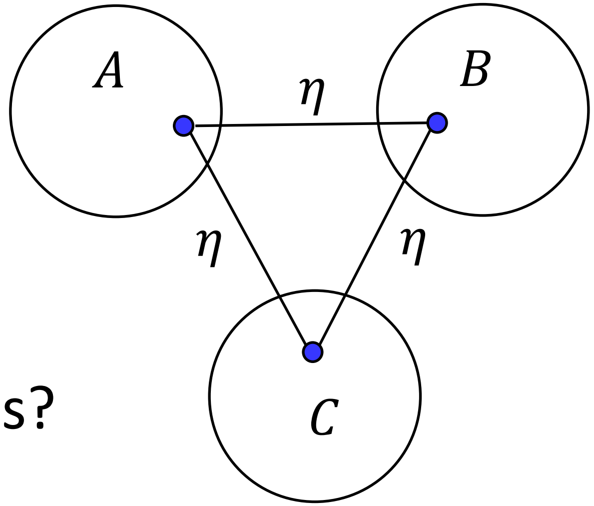
1. $a \le b < T$;
2. at most $\gamma \binom{b}{2}$ pairs of sets in $\mathcal{B}$ are not γ-regular.

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 poly$\left(\frac{1}{\gamma}\right)$

Triangles in a Random Tripartite Graph

- Consider disjoint sets A, B, C of vertices
- Suppose that each pair of nodes from different sets becomes an edge with probability η



- What is the expected number of triangles?

- Let X_{uvw} be an indicator that u, v, w form a triangle.

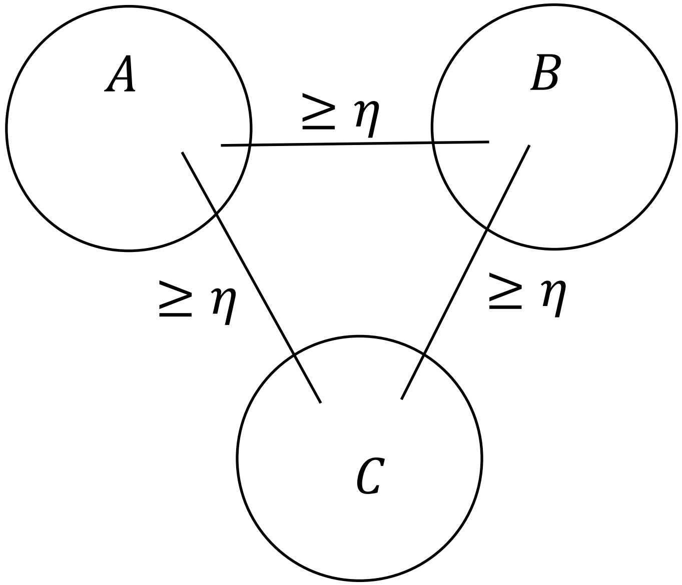
$$\mathbb{E}\left[\sum_{u\in A, v\in B, w\in C} X_{uvw}\right] = \sum_{u\in A, v\in B, w\in C} \mathbb{E}[X_{uvw}] = \eta^3 |A| \cdot |B| \cdot |C|$$

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: A'= the set of nodes in A, each of which has <$(\eta - \gamma^{\Delta})$ neighbors in B.



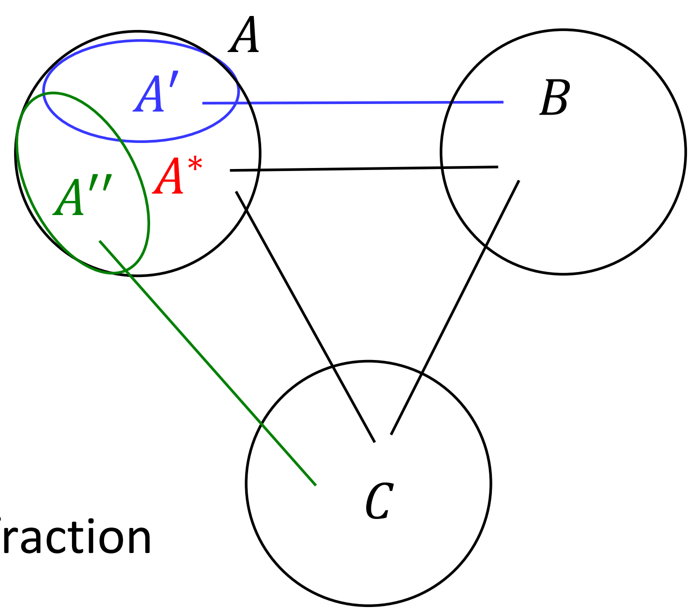
Triangles in a Graph with Three Regular Pairs

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: A'= the set of nodes in A, each of which has <$(\eta-\gamma^{\Delta})\,|B|$ neighbors in B.

- By Claim (most nodes in regular pairs have many neighbors), $|A'|<\gamma^{\Delta}|A|$.
- A''= the set of nodes in A, each of which has <$(\eta-\gamma^{\Delta})\,|C|$ neighbors in C.
- Analogously, $|A''|<\gamma^{\Delta}|A|$.
- $A^{*}=A-A'-A''$
- $|A^{*}|\geq\left(1-2\gamma^{\Delta}\right)|A|$
- Each node in A^{*} is adjacent to $\geq(\eta-\gamma^{\Delta})$ fraction of nodes in B and in C



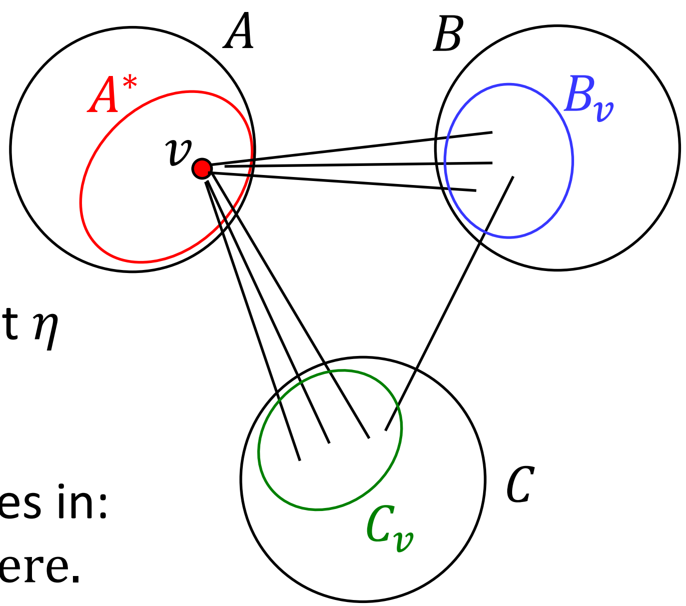
Triangles in a Graph with Three Regular Pairs

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 $\delta^{\Delta}\,|A|\cdot|B|\cdot|C|$ disjoint triangles, where $\gamma^{\Delta}=\gamma^{\Delta}(\eta)=\frac{\eta}{2}$ and $\delta^{\Delta}=\delta^{\Delta}(\eta)=\frac{1}{8}(1-\eta)\eta^{3}$.

Proof: Each $v \in A^*$ is adjacent to $\geq (\eta-\gamma^{\Delta})$ fraction of nodes in B and in C

$|A^*| \geq \left(1-2\gamma^{\Delta}\right)|A|$



- Consider a node $v \in A^*$
- Let B_v be the set of neighbors of v in B.
- Let C_v be the set of neighbors of v in C.
- But (B, C) is γ^{Δ}-regular with density at least η
- Therefore, $d(B_v, C_v) > \eta-\gamma^{\Delta}=\frac{\eta}{2}$
- Number of triangles each $v \in A^*$ participates in:

Type equation here.

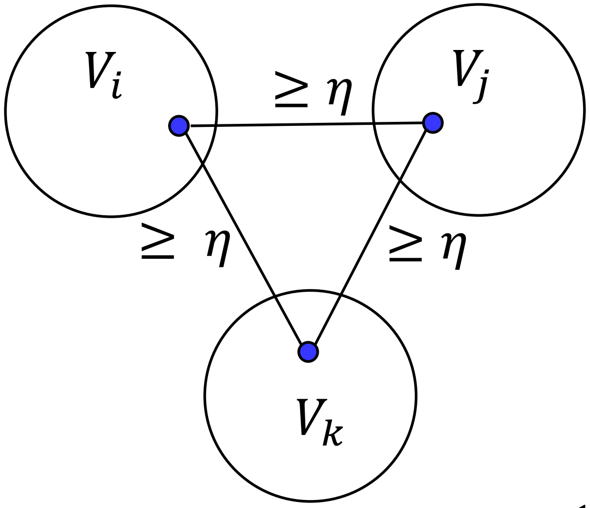
Proof of the Triangle-Removal Lemma: Idea

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 $\mathcal{B}=\{V_1, \ldots, V_b\}$:
  1. $\frac{4}{\varepsilon} \le b \le T$
  2. at most $\gamma \cdot \binom{b}{2}$ pairs among $V_1, \ldots, V_b$ 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 $d(V_i, V_j) \ge \varepsilon/4$

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

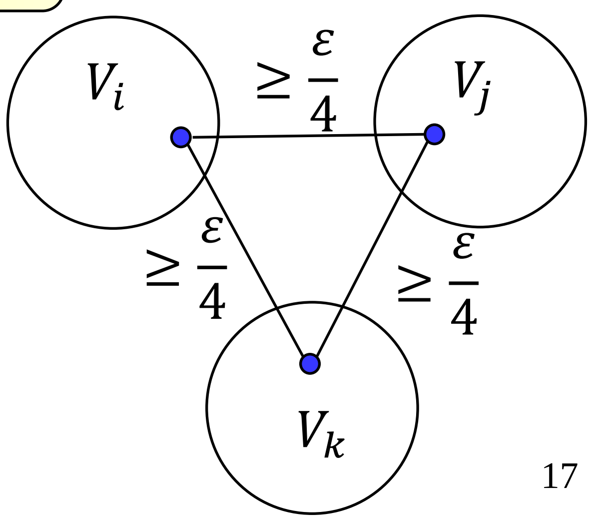
$\forall \varepsilon \exists \delta = \delta(\varepsilon)$ such that every n-node graph that is ε-far from triangle-free contains at least $\delta \cdot \binom{n}{3}$ distinct triangles.

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.