Extremal and Probabilistic Graph Theory

Instructor: Jie Ma, Scribed by Tianchi Yang

Mar 25th 2020, Wednesday

1 Lecture 10. Induced Removal Lemma

Lemma 1.1 (Strong regularity lemma). *For any sequence of constants $\varepsilon_0 \geq \varepsilon_1 \geq \cdots > 0$, there exists an integer M such that every graph has two equipartitions P, Q satisfying P is ε_0-regular, Q is $\varepsilon_{|P|}$-regular, Q refines P, $|Q| \leq M$ and $q(Q) \leq q(P) + \varepsilon_0$.*

Proof. The proof of regularity lemma can also show the following statement: for $\varepsilon > 0$, there is a constant $M(\varepsilon)$ such that for any equipartition P of $V(G)$, we can get a refinement P' of P where each part of P refines to at most $M(\varepsilon)$ parts and P' is ε-regular.

By repeatedly applying the above statement, we obtain a sequence of equipartitions $P_0, P_1, \cdots$ of $V(G)$, satisfying each P_{i+1} refines P_i and P_{i+1} is $\varepsilon_{|P_i|}$-regular with $|P_{i+1}| \leq |P_i| M(\varepsilon_{|P_i|})$. Since $q(P_i)$ is non-decreasing and $0 \leq q(\cdot) \leq 1$, we can find an $i \leq 1/\varepsilon$ satisfies $q(P_{i+1}) \leq q(P_i) + \varepsilon_0$. Thus $P = P_i$ and $Q = P_{i+1}$ are the desired sets. ∎

Recall the definition of the potential funcion $q(\cdot)$. Given an n-vertex graph G and $A, B \subseteq V(G)$, the potential function is $q(A,B) = \frac{|A||B|}{n^2} d^2(A,B)$. For a partition $P = \{V_1, \cdots, V_k\}$ of $V(G)$, let $q(P) = \sum_{i,j} q(V_i, V_j)$. Here we note that $q(V_i, V_i) = d(V_i, V_i) = 0$.

Now let $\mathcal{A} = \{A_1, \cdots, A_s\}$ be a partition of A and $\mathcal{B} = \{B_1, \cdots, B_t\}$ be a partition of B. Define $q(\mathcal{A}, \mathcal{B}) = \sum_{1 \leq i \leq s} \sum_{1 \leq j \leq t} q(A_i, B_j)$. Choose $x \in A$ uniformly at random, and choose $y \in B$ uniformly at random, independently of each other. Let $A_x \in \mathcal{A}$ be the part containing x, and $B_y \in \mathcal{B}$ be the part containing y. Let $z = d(A_x, A_y)$ be the random variable. Then

$$\mathbb{E}[z] = \sum_{i \in [s]} \sum_{j \in [t]} \frac{|A_i|}{|A|} \frac{|B_j|}{|B|} d(A_i, B_j) = \sum_{i \in [s]} \sum_{j \in [t]} \frac{e(A_i, B_j)}{|A||B|} = \frac{e(A,B)}{|A||B|} = d(A,B),$$

and

$$\mathbb{E}[z^2] = \sum_{i \in [s]} \sum_{j \in [t]} \frac{|A_i|}{|A|} \frac{|B_j|}{|B|} d^2(A_i, B_j) = \frac{n^2}{|A||B|} q(\mathcal{A}, \mathcal{B}).$$

Lemma 1.2. *For any sequence of constant $\varepsilon_0 \geq \varepsilon_1 \geq \cdots > 0$, there exists a $\delta_{1.2} > 0$ such that every n-vertex graph G has an equipartition $V_1, V_2, \cdots, V_k$ and $W_i \subseteq V_i$ for $i \in [k]$ satisfying*

- $|W_i| \geq \delta_{1.2} n$.
- (W_i, W_j) *is ε_k-regular for all $1 \leq i < j \leq k$.*
- *All but at most $\varepsilon_0 \binom{k}{2}$ of the pairs $\{i, j\}$ with $1 \leq i < j \leq k$ satisfy $|d(V_i, V_j) - d(W_i, W_j)| \leq \varepsilon_0$.*

Proof. By the regularity lemma, there exist two equipartitions P and Q of $V(G)$ such that the following hold. $P = \{V_1, V_2, \cdots, V_k\}$ is ε_0^3-regular, Q is ε_k-regular, $Q = \{W_i^j\}_{i \in [k], j \in [t]}$, where $\{W_i^j\}_{j \in [t]}$ is a partition of V_i, $|W_i^j| \geq \delta n$ and $q(Q) \leq q(P) + \varepsilon_0^3/4$.

Choose $x \in V(G)$ uniformly at random, and choose $y \in V(G)$ uniformly at random, independently of each other. Let $z_P = d(V_i, V_j)$ where $V_i \ni x$ and $V_j \ni y$. Similarly, let $z_Q = d(W_i^\alpha, W_j^\beta)$ where $W_i^\alpha \ni x$ and $W_j^\beta \ni y$. So we get $\mathbb{E}[z_P^2] = q(P,P) = 2q(P)$ and $\mathbb{E}[z_Q^2] = q(Q,Q) = 2q(Q)$. Also we have

$$\mathbb{E}[z_P z_Q] = \sum_{i\in[k]}\sum_{j\in[k]} \frac{|V_i|}{n}\frac{|V_j|}{n} d(V_i,V_j)\mathbb{E}[d(W_i^\alpha, W_j^\beta)] = \sum_{i\in[k]}\sum_{j\in[k]} \frac{|V_i|}{n}\frac{|V_j|}{n} d^2(V_i,V_j) = q(P,P) = \mathbb{E}[z_P^2],$$

then

$$\mathbb{E}[(z_Q - z_P)^2] = \mathbb{E}[z_Q^2] + \mathbb{E}[z_P^2] - 2\mathbb{E}[z_Q z_P] = \mathbb{E}[z_Q^2] - \mathbb{E}[z_P^2] = 2q(Q) - 2q(P) \le \varepsilon_0^3/2.$$

This implies that $|d(V_i, V_j) - d(W_i^\alpha, W_j^\beta)| \le \varepsilon_0$ holds for all (W_i^α, W_j^β) but at most $\varepsilon_0(tk)^2/2$ pairs.

Now for each $1 \le i \le k$, choose W_i^α uniformly at random, independently of each other. So we get k disjoint sets. Since there are at most $\varepsilon_0(tk)^2/2$ pairs (W_i^β, W_j^θ) with $|d(V_i,V_j) - d(W_i^\beta, W_j^\theta)| \ge \varepsilon_0$, the expected number of such pairs in the sets we chosen is no more than $\varepsilon_0(tk)^2 t^{k-2}/t^k = \varepsilon_0 k^2/2$. So with probability at least $1/2$, the sets we chosen contain no more than $\varepsilon_0 k^2$ such pairs. Meanwhile, we know G has at most $\varepsilon_k(tk)^2$ non-ε_k-regular (W_i^α, W_j^β) pairs. Thus with probability more than $1/2$, all (W_i^α, W_j^β) pairs we chosen are ε_k-regular, since $\varepsilon_k(tk)^2 t^{k-2}/t^k < 1/2$ where $\varepsilon_k < 1/(2k^2)$. Together, there exists a choice k sets, say $W_1, W_2, \cdots, W_k$, such that all pairs are ε_k-regular and all but at most $\varepsilon_0\binom{k}{2}$ pairs satisfy $|d(V_i,V_j) - d(W_i,W_j)| \le \varepsilon_0$, as desired. ∎

Lemma 1.3. *For every l, γ there exist $\delta_{1.3} = \delta(t,\gamma)$ such that for every graph G there is a disjoint vertex sets $W_1, W_2, \cdots, W_\ell$ satisfying*

- $|W_i| \ge \delta_{1.3} n$.
- *All $\binom{\ell}{2}$ pairs are γ-regular.*
- *Either all pairs are with densities at least $1/2$, or all pairs are with densities less than $1/2$.*

Proof. By regularity lemma, we can get an equipartion $\{V_i, \cdots, V_k\}$ of the of G such that all but at most $(r-1)^{-1}\binom{k}{2}$ pairs of them are γ-regular, with k to be chosen later. Then apply Turán's Theorem, there exist r sets V_i such that all pairs of them are γ-regular. Finally use Ramsey Theorem, we can take ℓ sets among these r sets such that all pairs are with densities at least $1/2$, or all pairs are with densities less than $1/2$. Obviously, we know such k, r exists. ∎

We are ready to prove the induced graph removal lemma.

Theorem 1.4 (Induced removal lemma)**.** *For any graph H and $\varepsilon > 0$, there is a constant $\delta > 0$ such that any n-vertex graph has at most $\delta n^{|V(H)|}$ induced copies of H can be made induced H-free by deleting or adding εn^2 edges.*

Proof. Let $h = |V(H)|$. Apply Lemma 1.2 to G with $\varepsilon_i = \varepsilon/3$, we can get an equipartition $V_1, \cdots, V_2$ and $W_i \subseteq V_i$ for $i \in [k]$ with the following properties. All pairs (W_i, W_j) are $\varepsilon/3$-regular and all but at most $\varepsilon\binom{k}{2}/3$ of the them satisfy $|d(V_i,V_j) - d(W_i,W_j)| \le \varepsilon/3$. Next we use Lemma 1.3 on the induced subgraph $G[W_i]$ to obtain sets $W_{i,1}, W_{i,2}, \cdots, W_{i,h}$, where $\ell = h$ and $\gamma = \varepsilon/3$.

Next we construct a new graph G_1 from G by add and removing the following edges.

- For $1 \leq i < j \leq k$ such that $|d(V_i, V_j) - d(W_i, W_j)| \geq \varepsilon/3$, remove all edges in (W_i, W_j) if $d(W_i, W_j) < 1/2$ and add all missing edges while $d(W_i, W_j) > 1/2$.

- For the rest $1 \leq i < j \leq k$, if $d(W_i, W_j) \leq 2\varepsilon/3$, remove all edges in (W_i, W_j), and if $d(V_i, V_j) \geq 1 - 2\varepsilon/3$, add all missing edges in (V_i, V_j).

- For a fixed i, if all densities of pairs from $W_{i,1}, W_{i,2}, \cdots, W_{i,h}$ are less than $1/2$, all edges in $G[V_i]$ are removed. Otherwise, all those densities are at least $1/2$, then we add all missing edges in $G[V_i]$.

By doing this, the total number of edges removed and added is at most εn^2. Now if the resulting graph G_1 is induced H-free, then we are done. So we may assume G_1 has an induced copy of H, and the vertices of it are from h parts with repetition, say $V_{i_1}, V_{i_2}, \cdots, V_{i_h}$. Let G_2 be the induced subgraph of G with vertices $W_{i_1,1}, W_{i_2,2}, \cdots, W_{i_h,h}$, clearly these sets are disjoint and pairwise ε-regular. Also we can check that they satisfy the density condition required for graph counting lemma, so there are at least δn^h induced copies of H in G where $\delta = (\delta_{1.2}\delta_{1.3}\varepsilon/3)^h$, a contradiction. ∎