1 Lecture 8 (ii): Sparse regularity and KŁR Conjecture (02/03/2019)

1.1 Szemerédi regularity lemma

Szemerédi Regularity Lemma [7] is one of the most powerful tools in understanding large dense graphs. Szemerédi first used the lemma in his celebrated theorem on long arithmetic progressions in dense subset of integers [6]. Regularity Lemma gives us a rough structure of graphs: it says that every large graph can be partitioned into a bounded number of parts such that the graphs between almost every pair of parts are random-like. To make this precise we need some definitions.

Let G be a graph and $X, Y \subseteq V(G)$. We define the *edge density* between X and Y $d(X,Y) = e(X,Y)/(|X||Y|)$. We say the pair (X,Y) is *ε-regular* if for all $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geqslant \varepsilon|X|$ and $|Y'| \geqslant \varepsilon|Y|$, we have $|d(X',Y') - d(X,Y)| < \varepsilon$. A vertex partition $\mathcal{P} = \{V_i\}_{i=1}^k$ is *equitable* if $V_i \cap V_j = \varnothing$ for every $1 \leqslant i < j \leqslant k$, and $\big||V_i| - |V_j|\big| \leqslant 1$. An equitable vertex partition $V_1, \ldots, V_K$ with K parts is *ε-regular* if all but at most εK^2 pairs of parts (V_i, V_j) $(1 \leqslant i < j \leqslant K)$ are ε-regular.

Theorem 1.1 (Szemerédi Regularity Lemma). *For every $\varepsilon > 0$ and every integer m, there exists an integer $M = M(m, \varepsilon)$ such that every graph G has an ε-regular partition into K parts, where $m \leqslant K \leqslant M$.*

One of the most important applications of Szemerédi regularity lemma is the counting lemma. The following is the counting lemma for counting complete subgraphs.

Theorem 1.2 (Counting Lemma). *Given $\varepsilon > 0$, integer $r > 0$ and real number d such that $10\varepsilon < d \leqslant 1$. There exists $m(d, \varepsilon)$, such that if $|V_1| = \cdots = |V_r| = m$ and each pair (V_i, V_j) is ε-regular with edge density at least d for every $i \neq j$. Then the number of K_r in the graph is at least $(d - 3\varepsilon)^{\binom{r}{2}} m^r$.*

Counting lemma asserts that the number of subgraphs is what we expect if the ambient graph is random. Note that the counting lemma works not only for counting complete graphs, it works for counting all subgraphs (or induced subgraphs) of constant order. One can also relax the condition of counting lemma, from the ambient graph has an ε-partition, to a weak ε-partition. For more details we refer to [4].

1.2 KŁR Conjecture

Although Szemerédi regularity lemma works for all graphs, but when graph is sparse, i.e. the number of edges is $o(n^2)$, it does not provide any useful information of the graph, since all the edges are hidden in the error term. Kohayakawa and Rödl [3] proved a sparse analogue of regularity lemma. We first introduce some definitions.

Definition 1.3. Given $p \in (0,1)$, $\varepsilon > 0$, a pair of vertices (V_1, V_2) is (ε, p)-*regular* if for every $W_1 \subseteq V_1$ and $W_2 \subseteq V_2$, with $|W_1| \geqslant \varepsilon|V_1|$ and $|W_2| \geqslant \varepsilon|V_2|$, we have $|d(W_1, W_2) - d(V_1, V_2)| < \varepsilon p$. Given a graph G and a vertex partition $\mathcal{P} = V_1, \ldots, V_K$, we say $\mathcal{P}$ is (ε, p)-*regular* if $\mathcal{P}$ is equitable, and all but at most εK^2 pairs are (ε, p)-regular.

Definition 1.4. Given η, k, we say G is (η, p, k)-*upper uniform* if for every $W \subseteq V(G)$ with $|W| \geqslant \eta n$, we have $d(W) \leqslant kp$.

Roughly speaking, a graph G satisfies upper uniform property means G is "nowhere dense". All the random graphs $\mathcal{G}(n,p)$ satisfies this condition.

Theorem 1.5 (Sparse Regularity Lemma)**.** *Given ε, K, r_0 there exist η, R, such that for every $p \in (0,1)$, every (η, p, K)-upper uniform graph with at least r_0 vertices has an (ε, p)-regular partition into r parts, where $r_0 \leqslant r \leqslant R$.*

Remark. Scott [5] removed the (ε, p, K)-upper uniform condition in the above theorem.

With sparse regularity lemma in hand, one may hope if we also have an analogue of counting lemma (or embedding lemma) in sparse graphs. The answer is negative. Indeed, given a graph H, consider the random blow-up of H, that is, the random graph G obtained from H by replacing each vertex of H by an independent set of size n and each edge of H by a random bipartite graph which edge occurs with probability p. With high probability, the number of copies of H in G will be about $p^{e(H)}n^{v(H)}$. Hence if $p^{e(H)}n^{v(H)} \ll pn^2$, then one can remove all copies of H from G by deleting a tiny proportion of all edges. We denote the resulted graph by G'. Note that when p is small enough, although G' satisfies (ε, p)-regular condition, it does not contain any copy of H.

Nevertheless, it still seemed likely that except a tiny fraction, a counting lemma or embedding lemma for sparse graphs should hold. This was formalized in the following conjecture of Kohayakawa, Łuczak, and Rödl [2], usually referred to as the KŁR conjecture. Before stating the conjecture we first introduce some definitions and notation.

Given a graph H, let us denote by $\mathcal{G}(H, n, m, p, \varepsilon)$ the collection of all graphs G constructed in the following way. The vertex set of G is a disjoint union $V_1 \cup \cdots \cup V_{v(H)}$ of sets of size n, one for each vertex of H. For each edge $\{i, j\}$ of H, we add to G an (ε, p)-regular bipartite graph with m edges between the sets V_i and V_j. These are the only edges of G. let $\mathcal{G}^*(H, n, m, p, \varepsilon)$ denote the collection of graphs in $\mathcal{G}(H, n, m, p, \varepsilon)$ that contain no copy of H. Now we are able to state the KŁR conjecture.

Theorem 1.6 (The KŁR conjecture). *For every graph H and every positive β, there exist positive constants C, n_0, and ε such that the following holds. For every $n \in \mathbb{N}$ with $n > n_0$ and $m \in \mathbb{N}$ with $m > Cn^{2-1/m_2(H)}$,*

$$|\mathcal{G}^*(H, n, m, m/n^2, \varepsilon)| \leqslant \beta^m \binom{n^2}{m}^{e(H)}.$$

In the rest of this notes, we will present a sketch of the proof of Theorem 1.6. For the detailed version see [1]. We first recall the container lemma.

Theorem 1.7 (Hypergraph container lemma). *For every $k \in \mathbb{N}$ and all positive c and ε, there exists a positive constant C such that the following holds. Let $\mathcal{H}$ be a k-uniform hypergraph and let $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$ be an increasing family of sets such that $|A| \geqslant \varepsilon v(\mathcal{H})$ for all $A \in \mathcal{F}$. Suppose that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense and $p \in (0, 1)$ is such that, for every $\ell \in [k]$,*

$$\Delta_\ell(\mathcal{H}) \leqslant c \cdot p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}.$$

Then there exists a family $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leqslant Cp \cdot v(\mathcal{H})}$ and functions $f \colon \mathcal{S} \to \overline{\mathcal{F}}$ and $g \colon \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that for every $I \in \mathcal{I}(\mathcal{H})$,

$$g(I) \subseteq I \qquad \text{and} \qquad I \setminus g(I) \subseteq f(g(I)).$$

We will deduce from Theorem 1.7 the KŁR conjecture, Theorem 1.6. Let H be an arbitrary graph and let $\mathcal{H}$ be the $e(H)$-uniform hypergraph of canonical copies of H in the complete blow-up of H.

Given a graph H and integers $n_1, \ldots, n_{v(H)}$, let us denote by $\mathcal{G}(H; n_1, \ldots, n_{v(H)})$ the collection of all graphs G constructed in the following way. The vertex set of G is a disjoint union $V_1 \cup \ldots \cup V_{v(H)}$ of sets of sizes $n_1, \ldots, n_{v(H)}$, respectively, one for each vertex of H. The only edges of G lie between those pairs of sets (V_i, V_j) such that $\{i, j\}$ is an edge of H. We observe that $\mathcal{G}(H, n, m, p, \varepsilon) \subseteq \mathcal{G}(H; n, \ldots, n)$ for all m, p, and ε.

Now we are going to define an appropriate family $\mathcal{F}$ and showing that $\mathcal{H}$ is $(\mathcal{F},\varepsilon)$-dense. The following lemma suggests the right choice of $\mathcal{F}$. The lemma is well-known, and so we omit the proof.

Lemma 1.8. *Let H be a graph and let $\delta\colon (0,1] \to (0,1)$ be an arbitrary function. There exist positive constants α_0, ξ, and N such that for every collection of integers $n_1,\dots,n_{v(H)}$ satisfying $n_1,\dots,n_{v(H)} \geqslant N$ and every graph $G \in \mathcal{G}(H; n_1,\dots,n_{v(H)})$, one of the following holds:*

(a) G contains at least $\xi n_1 \dots n_{v(H)}$ canonical copies of H.

(b) There exist a positive constant α with $\alpha \geqslant \alpha_0$, an edge $\{i,j\} \in E(H)$, and sets $A_i \subseteq V_i$, $A_j \subseteq V_j$ such that $|A_i| \geqslant \alpha n_i$, $|A_j| \geqslant \alpha n_j$, and $d_G(A_i,A_j) < \delta(\alpha)$.

Our next lemma allows us to count (ε,p)-regular subgraphs of a graph that has a 'hole', as in Lemma 1.8(b). Recall that $\mathcal{G}(K_2,n,m,p,\varepsilon)$ denotes the collection of all (ε,p)-regular bipartite graphs with m edges and n vertices in each part. Given such G, let $V_1(G)$ and $V_2(G)$ denote the two parts. For each $\beta \in (0,1)$, define a function $\delta\colon (0,1] \to (0,1)$ by setting

$$\delta(x) = \frac{1}{4e}\left(\frac{\beta}{2}\right)^{2/x^2} \tag{1}$$

for each $x \in (0,1]$. The following lemma says that a graph $\tilde{G}$ that has a hole of size αn and density at most $\delta(\alpha)$ has very few subgraphs in $\mathcal{G}(K_2,n,m,m/n^2,\varepsilon)$. We omit the proof.

Lemma 1.9. *For every positive α_0 and β, there exists a positive constant ε such that the following holds. Let $\tilde{G} \subseteq K_{n,n}$ be such that there exist subsets $A \subseteq V_1(\tilde{G})$ and $B \subseteq V_2(\tilde{G})$ with*

$$\min\{|A|,|B|\} \geqslant \alpha n \quad \textit{and} \quad d_G(A,B) < \delta(\alpha)$$

for some $\alpha \in [\alpha_0,1]$, and let $S \subseteq \tilde{G}$. Then, for every m with $|S|/\varepsilon \leqslant m \leqslant n^2$, there are at most

$$\beta^m \binom{n^2}{m-|S|}$$

subgraphs of $\tilde{G}$ that belong to $\mathcal{G}(K_2,n,m,m/n^2,\varepsilon)$ and contain S.

We now deduce Theorem 1.6 from Theorem 1.7.

Proof of Theorem 1.6. Let H be a fixed graph, let $n \in \mathbb{N}$, and let $H(n)$ be the largest graph in the family $\mathcal{G}(H;n,\dots,n)$, i.e., the complete blow-up of H, where each vertex of H is replaced by an independent set of size n and each edge of H is replaced by the complete bipartite graph $K_{n,n}$. Let $\mathcal{H}$ be the $e(H)$-uniform hypergraph on the vertex set $E(H(n))$ whose edges are all $n^{v(H)}$ canonical copies of H in $H(n)$.

Fix an arbitrary positive constant β, let $\delta\colon (0,1] \to (0,1)$ be the function defined in (1) with β replaced by $\beta/2$, i.e., set $\delta(x) = \frac{1}{4e}\left(\frac{\beta}{4}\right)^{2/x^2}$ for each $x \in (0,1]$, and let $\alpha_0 = (\alpha_0)_{1.8}(H,\delta)$, $\xi = \xi_{1.8}(H,\delta)$, and $N = N_{1.8}(H,\delta)$. Let $\mathcal{F}$ be the family of all subgraphs of $H(n)$, i.e., graphs in $\mathcal{G}(H; n, \dots, n)$, for which (b) in Lemma 1.8 is not satisfied. Clearly $\mathcal{F}$ is an upset, and so, by Lemma 1.8, $\mathcal{H}$ is $(\mathcal{F},\xi)$-dense provided that $n \geqslant N$.

Now, since $\mathcal{H}$ is contained in the hypergraph of all copies of H in the complete graph on $v(H)n$ vertices and contains a positive proportion of those copies, it follows easily that $\mathcal{H}$ satisfies the assumptions of Theorem 1.7 with $p = n^{2-1/m_2(H)}$ and $\varepsilon = \xi$, for some constant c depending only on H. Therefore, by hypergraph container lemma, there is a constant C', a family $\mathcal{S} \subseteq \binom{E(H(n))}{\leqslant C' n^{2-1/m_2(H)}}$, and functions $f\colon \mathcal{S} \to \overline{\mathcal{F}}$ and $g\colon \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that

$$g(I) \subseteq I \qquad \text{and} \qquad I \setminus g(I) \subseteq f(g(I))$$

for every $I \in \mathcal{I}(\mathcal{H})$.

Let ε be a sufficiently small positive constant such that, in particular, $\varepsilon \leqslant \varepsilon_{1.9}(\alpha_0, \beta/2)$, let $C = C'/\varepsilon$, and suppose that $m \geqslant Cn^{2-1/m_2(H)}$. Let $\mathcal{G}^* = \mathcal{G}^*(H, n, m, m/n^2, \varepsilon)$ and note that $\mathcal{G}^* \subseteq \mathcal{I}(\mathcal{H})$. We are required to bound from above the number of graphs in $\mathcal{G}^*$.

To this end, fix an $S \in \mathcal{S}$, let

$$\mathcal{G}_S^* = \big\{ G \in \mathcal{G}^* \colon g(G) = S \big\},$$

and let $G_S = f(S)$. For each $\{i,j\} \in E(H)$, let $s(i,j) = e_S(V_i, V_j)$ and note that $\sum_{ij \in E(H)} s(i,j) = |S|$. Since

$$|S| \leqslant C' n^{2-1/m_2(H)} \leqslant \varepsilon \cdot C n^{2-1/m_2(H)} \leqslant \varepsilon m,$$

then $s(i,j) \leqslant \varepsilon m$ for every $\{i,j\} \in E(H)$.

Now, since $G_S \in \overline{\mathcal{F}}$, it follows that there exist an $\alpha \in [\alpha_0, 1]$, an edge $\{i,j\} \in E(H)$, and sets $A_i \subseteq V_i$, $A_j \subseteq V_j$ such that $|A_i|, |A_j| \geqslant \alpha n$ and $d_{G_S}(A_i, A_j) < \delta(\alpha)$. By Lemma 1.9, it follows that there are at most

$$\left(\frac{\beta}{2}\right)^m \binom{n^2}{m - s(i,j)}$$

choices for the edges between V_i and V_j such that $G[V_i, V_j] \in \mathcal{G}(K_2, n, m, m/n^2, \varepsilon)$ and $S[V_i, V_j] \subseteq G[V_i, V_j] \subseteq S \cup G_S[V_i, V_j]$. It follows immediately that

$$|\mathcal{G}_S^*| \leqslant \left(\frac{\beta}{2}\right)^m \prod_{ij \in E(H)} \binom{n^2}{m - s(i,j)}.$$

Summing over sets $S \in \mathcal{S}$, we obtain

$$|\mathcal{G}^*| \leqslant \sum_{S \in \mathcal{S}} \left(\frac{\beta}{2}\right)^m \prod_{ij \in E(H)} \left(\frac{m}{n^2 - m}\right)^{s(i,j)} \binom{n^2}{m} = \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{S \in \mathcal{S}} \left(\frac{m}{n^2 - m}\right)^{|S|}$$
$$\leqslant \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{s \leqslant \varepsilon m} \binom{e(H)n^2}{s} \left(\frac{2m}{n^2}\right)^s \leqslant \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{s \leqslant \varepsilon m} \left(\frac{2e \cdot e(H)m}{s}\right)^s.$$

Now, since ε was chosen to be sufficiently small, it follows that the summand above is increasing in s on $(0, \varepsilon m]$ and hence

$$|\mathcal{G}^*| \leqslant \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} m \left(\frac{2e \cdot e(H)}{\varepsilon}\right)^{\varepsilon m} \leqslant \beta^m \binom{n^2}{m}^{e(H)},$$

as required. □

References

[1] J. Balogh, R. Morris, and W. Samotij. *Independent sets in hypergraphs*, J. Amer. Math. Soc. 28, 669–709, 2015.

[2] Y. Kohayakawa, T. Łuczak, and V. Rödl. *On K_4-free subgraphs of random graphs*, Combinatorica, 17, 173–213, 1997.

[3] Y. Kohayakawa, V. Rödl. *Szemerédi's regularity lemma and quasirandomness*, in Recent Advances in Algorithms and Combinatorics, CMS Books Math./Ouvrages Math. SMC 11, Springer, New York, 289–351, 2003.

[4] L. Lovász and B. Szegedy. *Limits of dense graph sequences*, J. Combin. Theory B, 96, 933–957, 2006.

[5] A. Scott. *Szemerédi's regularity lemma for matrices and sparse graphs*, Combinatorics, Probability and Computing, 20, 455–466, 2011.

[6] E. Szemerédi. *On sets of integers containing no k elements in arithmetic progression.* Acta Arith., 27, 199–245, 1975.

[7] E. Szemerédi. *Regular partitions of graphs.* In Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), volume 260 of Colloq. Internat. CNRS, pages 399–401. CNRS, Paris, 1978.