Regularity Lemmas for Hypergraphs and Quasi-randomness

Fan R.K. Chung
Bellcore, Morristown NJ 07060-1910

ABSTRACT

We give a simple proof for Szemerédi's Regularity Lemma and its generalization for k-uniform hypergraphs. For fixed k, there are altogether $k-1$ different versions of the regularity lemma for k-uniform hypergraphs. The connection between regularity lemmas for hypergraphs and quasi-random classes of hypergraphs is also investigated.

1. INTRODUCTION

Szemerédi's regularity lemma [9] asserts that if the vertex set of a graph is partitioned in the right way, then the edges joining the sets behave in some random-like manner. To be precise, we need some definitions. Let X and Y denote two disjoint subsets of the vertex set $V(G)$ of a graph G. $E(X, Y)$ consists of all pairs $\{x, y\}$ where $x \in X$ and $y \in Y$ and $\{x, y\}$ is in the edge set $E(G)$ of G. The density of edges between X and Y is defined to be

$$\delta(X, Y) = \frac{|E(X, Y)|}{|X|\,|Y|} .$$

Regularity Lemma. *For every $\epsilon > 0$ and $w > 0$, there exist $t(\epsilon, w)$ and $n(\epsilon, w)$, such that for every graph G, the vertex set can be partitioned into sets $V_0, V_1, \ldots, V_t$, where $w < t < t(\epsilon, w)$, so that $|V_0| < \epsilon n$ and $|V_i| = m$ for $i \geq 1$, and for all but at most $\epsilon\binom{t}{2}$ pairs i, j with $i \neq j$, for every $X \subseteq V_i$ and $Y \subseteq V_j$ satisfying $|X|, |Y| > \epsilon m$, we have*

$$|\delta(X, Y) - \delta(V_i, V_j)| < \epsilon .$$

Random Structures and Algorithms, Vol. 2, No. 2 (1991)

Since the regularity lemma can be used to extract the underlying structure of graphs, it has found numerous applications in many areas ranging from extremal problems [3, 10], to lower bounds for complexity [5, 6]. After the regularity lemma was introduced in the 70s, it is of interest to find its analog for hypergraphs. Peter Frankl and V. Rödl have announced their results generalizing the regularity lemma to hypergraphs [4]. We give here a simple proof of several versions of regularity lemma for hypergraphs. Some of these versions are stronger and some are simpler than the above regularity lemma as stated.

A k-uniform G hypergraph (or a k-graph for short) consists of a vertex set V and an edge set as a subset of the set $\binom{V}{k}$ of k-elements subsets of V. Roughly speaking, for any k-graph and any $r < k$, if the family of all r-subsets of the vertex set is partitioned in the right way, then the hyper-edges "joining" parts of the partitions behave in some random-like fashion. The statements will be given in Section 2. Section 3 contains the proof. We also consider the relation of regularity lemmas and equivalence classes of graph properties, so-called "quasi-random" graphs, in Section 4.

2. REGULARITY LEMMAS FOR k-GRAPHS

Suppose G is a k-graph with vertex set $V = V(G)$ and edge set $E = E(G)$. Let r denote a positive integer smaller than k. Let $S_1, \ldots, S_{\binom{k}{r}}$ be disjoint subsets of $\binom{V}{r}$, the set of all r-subsets of V. In other words, S_i can be viewed as r-graphs. We define $E_G(S_1, \ldots, S_{\binom{k}{r}}) = y \in E(G) : |\binom{y}{r} \cap S_i| = 1$ for $i = 1, \ldots, \binom{k}{r}$. If G is the special graph with edge set $\binom{V}{k}$, we write $E_G(S_1, \ldots, S_{\binom{k}{r}}) = E(S_1, \ldots, S_{\binom{k}{r}})$. We also denote $e_G(S_1, \ldots, S_{\binom{k}{r}}) = |E_G(S_1, \ldots, S_{\binom{k}{r}})|$. The (k, r)-density $\delta_{k,r}$ is defined as follows:

$$\delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}}) = \frac{e_G(S_1, S_2, \ldots, S_{\binom{k}{r}})}{e(S_1, S_2, \ldots, S_{\binom{k}{r}})} .$$

For $r = 1$, $e(S_1, \ldots, S_k) = |S_1| \cdots |S_k|$. We note that $\delta_{2,1}$ coincides with what was described in Section 1 (for $k = 2$ and $r = 1$).

We say $\{S_1, \ldots, S_{\binom{k}{r}}\}$ is (k, r)-ϵ-regular if for every choice of $T_i \subseteq S_i$ with

$$e(T_1, \ldots, T_{\binom{k}{r}}) \geq \epsilon e(S_1, \ldots, S_{\binom{k}{r}}),$$

we have

$$|\delta_{k,r}(T_1, \ldots, T_{\binom{k}{r}}) - \delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}})| < \epsilon .$$

For general k, there are $k - 1$ different versions of the regularity lemma. Namely, for each $1 \leq r \leq k$, there is one version as follows:

Theorem 2.1. *Suppose $1 \le r < k$. For every $\epsilon > 0$, there exists $t(\epsilon)$ such that for every k-graph G, $\binom{V}{r}$ can be partitioned into sets $S_1, \ldots, S_t$ for some $t < t(\epsilon)$ so that all but ϵn^k edges of G are contained in $E(S_{i_1}, \ldots, S_{i_{\binom{k}{r}}})$* for some $i_1, \ldots, i_{\binom{k}{r}}$, where $1 \le i_1 < \cdots < i_{\binom{k}{r}} \le t$ and $\{S_{i_1}, \ldots, S_{i_{\binom{k}{r}}}\}$ is (k, r)-ϵ-regular.

Let P denote a partition of $\binom{V}{r}$ into $P_1, \ldots, P_t$. The cardinality of P is t and we write $|P| = t$. P is said to be a (k, r)-ϵ-regular partition for a k-graph G if all but ϵn^k edges of G are contained in $E_G(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})$ for some distinct $P_{i_j} \in P$, $j = 1, \ldots, \binom{k}{r}$ where $\{P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}\}$ is (k, r)-ϵ-regular.

A partition Q of $\binom{V}{r}$ is said to be a refinement of a partition P if members of P consist of disjoint unions of members of Q. Theorem 2.1 follows from the following stronger version.

Theorem 2.2. *Suppose $1 \le r < k$. For every $\epsilon > 0$ and every partition P of $\binom{V}{r}$, every k-graph G has a (k, r)-ϵ-regular partition Q which is a refinement of P and $|Q|$ is bounded above by a constant depending only on ϵ and $|P|$. Furthermore, any refinement of Q is (k, r)-ϵ-regular.*

It is easy to see that Theorem 2.2 implies the original regularity lemma as a special case. We can choose the refinement of Q in Theorem 2.2 so that all its members have approximately equal size (or all but one have equal size and the exceptional one has fewer than the fraction ϵ of the total).

3. PROOFS OF THE THEOREMS

The proof of Theorem 2.1 consists of a sequence of lemmas.

Lemma 1. *For positive values c_i and x_i, for $i = 1, \ldots, N$ with $\Sigma_{i=1}^N c_i = 1$, we have*

$$\sum_{i=1}^{N} c_i x_i^2 \ge \left(\sum_{i=1}^{N} c_i x_i\right)^2$$

Furthermore, if

$$\sum_{i=1}^{M} c_i x_i \ge \left(\sum_{i=1}^{N} c_i x_i\right)\left(\sum_{i=1}^{M} c_i\right) + \gamma$$

then

$$\sum_{i=1}^{N} c_i x_i^2 \ge \left(\sum_{i=1}^{N} c_i x_i\right)^2 + \frac{\gamma^2}{\left(\sum_{i=1}^{M} c_i\right)\left(\sum_{i=M+1}^{N} c_i\right)}$$

Proof. This follows from the inequality $\Sigma\, a_i^2\, \Sigma\, b_i^2 \geq (\Sigma\, a_i b_i)^2$ by taking $a_i = \sqrt{c_i}$ and $b_i = \sqrt{c_i} x_i$. The rest of Lemma 1 is an easy exercise. ∎

Let G be a k-graph on n vertices and P denote a partition of $\binom{V}{r}$ into $P_1, \ldots, P_t$. The index of P is defined by

$$ind\ P = \sum \frac{(e_G(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}))^2}{e(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})}$$

where the sum ranges over all choices of $\binom{k}{r}$ of the P's. It is easy to see that

$$\frac{ind\ P}{\binom{n}{k}} = \sum \frac{e(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})}{\binom{n}{k}} \cdot \left(\frac{e_G(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})}{e(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})} \right)^2 \leq \sum \frac{e(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})}{\binom{n}{k}} \leq 1\,. \tag{1}$$

Lemma 2. *Suppose Q is a refinement of P. Then*

$$ind\ Q \geq ind\ P\,.$$

Proof. It suffices to show that for fixed $P_1, \ldots, P_{\binom{k}{r}}$,

$$\sum_{Q_1 \subseteq P_1} \sum_{Q_2 \subseteq P_2} \cdots \sum_{Q_{\binom{k}{r}} \subseteq P_{\binom{k}{r}}} \frac{(e_G(Q_1, \ldots, Q_{\binom{k}{r}}))^2}{e(Q_1, \ldots, Q_{\binom{k}{r}})} \geq \frac{(e_G(P_1, \ldots, P_{\binom{k}{r}}))^2}{e(P_1, \ldots, P_{\binom{k}{r}})}\,.$$

We use Lemma 1 by taking c's to be

$$\frac{e(Q_1, \ldots, Q_{\binom{k}{r}})}{e(P_1, \ldots, P_{\binom{k}{r}})}$$

and x's to be

$$\frac{e_G(Q_1, \ldots, Q_{\binom{k}{r}})}{e(Q_1, \ldots, Q_{\binom{k}{r}})}\,.$$

Therefore

$$\sum_{i=1}^{N} c_i x_i^2 = \sum_{\substack{Q_j \subseteq P_j \\ j=1,\ldots,\binom{k}{r}}} \frac{(e_G(Q_1, \ldots, Q_{\binom{k}{r}}))^2}{e(Q_1, \ldots, Q_{\binom{k}{r}})\, e(P_1, \ldots, P_{\binom{k}{r}})}$$

$$\geq \left(\sum \frac{e_G(Q_1, \ldots, Q_{\binom{k}{r}})}{e(P_1, \ldots, P_{\binom{k}{r}})} \right)^2$$

$$\geq \left(\frac{e_G(P_1, \ldots, P_{\binom{k}{r}})}{e(P_1, \ldots, P_{\binom{k}{r}})} \right)^2$$

Lemma 2 is proved. ∎

When P is (k, r)-ϵ-irregular, Lemma 2 can be further improved. Let $S_1, \ldots, S_p$, $T_1, \ldots, T_q$ be subsets of $\binom{V}{r}$. We say $U = \{U_1, \ldots, U_s\}$ is a refinement of $S_1, \ldots, S_p$ by $T_1, \ldots, T_q$ if (i) $U_i \cap U_j = \emptyset$ if $i \neq j$, (ii) S_i is the disjoint union of U's, and (iii) $U_i \subset T_j$ or $U_i \subset V - T_j$, for all i and j.

Lemma 3. *Suppose $S_1, \ldots, S_{\binom{k}{r}}$ are disjoint subsets of $\binom{V}{r}$. If there exist $T_j \subset S_j$ for $j = 1, \ldots, \binom{k}{r}$, so that*

$$|\delta_{k,r}(T_1, \ldots, T_{\binom{k}{r}}) - \delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}})| > \epsilon ,$$

then any refinement U of $S_1, \ldots, S_{\binom{k}{r}}$ by $T_1, \ldots, T_{\binom{k}{r}}$ satisfies

$$ind\, U \geq ind(S_1, \ldots, S_{\binom{k}{r}}) + \epsilon^2 e(T_1, \ldots, T_{\binom{k}{r}})$$

where,

$$ind\, U = \sum_{U_j \in U} \frac{e_G(U_1, \ldots, U_{\binom{k}{r}}))^2}{e(U_1, \ldots, U_{\binom{k}{r}})} \text{ and } ind(S_1, \ldots, S_{\binom{k}{r}}) = \frac{(e_G(S_1, \ldots, S_{\binom{k}{r}})^2}{e(S_1, \ldots, S_{\binom{k}{r}})}$$

Furthermore, if $\{S_1, \ldots, S_{\binom{k}{r}}\}$ is (k, r)-ϵ-irregular, we have

$$ind\, U \geq ind(S_1, \ldots, S_{\binom{k}{r}}) + \epsilon^3 e(S_1, \ldots, S_{\binom{k}{r}}) .$$

Proof. We follow the proof of Lemma 2 and set c's to be

$$\frac{e(U_1, \ldots, U_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})} ,$$

and x's to be

$$\frac{e_G(U_1, \ldots, U_{\binom{k}{r}})}{e(U_1, \ldots, U_{\binom{k}{r}})} .$$

Let the first M of the x's to be those with U's contained in T's. We then have

$$\sum_{i=1}^{M} c_i = \frac{e(T_1, \ldots, T_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})}$$

and

$$\begin{aligned}
\sum_{i=1}^{M} c_i x_i &= \sum \frac{e_G(U_1, \ldots, U_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})} \cdot \frac{e_G(U_1, \ldots, U_{\binom{k}{r}})}{e(U_1, \ldots, U_{\binom{k}{r}})} \\
&= \frac{e_G(T_1, \ldots, T_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})} \\
&= \delta_{k,r}(T_1, \ldots, T_{\binom{k}{r}}) \cdot \frac{e(T_1, \ldots, T_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})}
\end{aligned}$$

Without loss of generality, we assume

$$\delta_{k,r}(T_1, \ldots, T_{\binom{k}{r}}) > \delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}}) + \epsilon .$$

Therefore

$$\begin{aligned}
\sum_{i=1}^{M} c_i x_i &> (\delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}}) + \epsilon) \cdot \frac{e(T_1, \ldots, T_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}})} \\
&= \left(\sum_{i=1}^{N} c_i x_i + \epsilon\right)\left(\sum_{i=1}^{M} c_i\right)
\end{aligned}$$

Therefore, by Lemma 1 we have

$$\begin{aligned}
\frac{ind\, U}{e(S_1, \ldots, S_{\binom{k}{r}})} &= \sum_{i=1}^{N} c_i x_i^2 \geq \left(\sum_{i=1}^{N} c_i x_i\right)^2 + \frac{\epsilon^2 \cdot \sum_{i=1}^{M} c_i}{\sum_{i=M+1}^{N} c_i} \\
&= (\delta_{k,r}(S_1, \ldots, S_{\binom{k}{r}}))^2 + \frac{\epsilon^2 e(T_1, \ldots, T_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}}) - e(T_1, \ldots, T_{\binom{k}{r}})}
\end{aligned}$$

$$\begin{aligned}
ind\, U &\geq ind(S_1, \ldots, S_{\binom{k}{r}}) + \frac{\epsilon^2 e(T_1, \ldots, T_{\binom{k}{r}}) e(S_1, \ldots, S_{\binom{k}{r}})}{e(S_1, \ldots, S_{\binom{k}{r}}) - e(T_1, \ldots, T_{\binom{k}{r}})} \\
&\geq ind(S_1, \ldots, S_{\binom{k}{r}}) + \epsilon^2 e(T_1, \ldots, T_{\binom{k}{r}})
\end{aligned}$$

If $\{S_1, \ldots, S_{\binom{k}{r}}\}$ is (k, r)-ϵ-irregular, there exist such T_i's that $e(T_1, \ldots, T_{\binom{k}{r}}) \geq \epsilon e(S_1, \ldots, S_{\binom{k}{r}})$. Therefore, Lemma 3 is proved.

Lemma 4. *Let ϵ be a positive value and P denote a partition of $\binom{V}{r}$ into $P_1, \ldots, P_t$. Let G be a k-graph on n vertices. Suppose more than $\delta\binom{n}{k}$ edges are in $E_G(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})$ where $\{P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}\}$ is (k, r)-ϵ-irregular. Then there is a partition Q of $\binom{V}{r}$ which is a refinement of P and $|Q| \leq t \cdot 2^{\binom{t}{r}}$ such that*

$$ind\ Q \geq ind\ P + \epsilon^3\delta\binom{n}{k}.$$

Proof. For $P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}$ which is (k, r)-ϵ-irregular, we apply Lemma 3 and we obtain Q which is a refinement of P by all (k, r)-ϵ-irregular subsets $T_{i_1}, \ldots, T_{i_{\binom{k}{r}}}$. We have

$$\begin{aligned} ind\ Q &= \sum_{Q_{i_1}, \ldots, Q_{i_{\binom{k}{r}}}} ind(Q_{i_1}, \ldots, Q_{i_{\binom{k}{r}}}) \\ &\geq \sum_{P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}} ind(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}) + \epsilon^3 \sum_{P_{i_j}\ \epsilon\text{-irregular}} e(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}}) \\ &\geq ind\ P + \epsilon^3\delta\binom{n}{k} \end{aligned}$$

It is not hard to see that the refinement Q of P by all choices of $T_i, \ldots, T_{i_{\binom{k}{r}}}$ has cardinality at most $|P|2^{\binom{t}{r}}$. Lemma 4 is proved ∎

We are now ready to prove Theorem 2.2. From (1) we have

$$\frac{ind\ Q}{\binom{n}{k}} \leq 1$$

for any partition Q of $\binom{V}{r}$.

Proof of Theorem 2.2. We start with a given partition P. First we refine P into P' with $|P'| \leq |P|\epsilon^{-1}$ so that all but $\epsilon n^k/2$ edges are in $E_G(P_{i_1}, \ldots, P_{i_{\binom{k}{r}}})$ where P_{i_j} are in P' and are distinct. (This can be easily done by, for example, restricting P'_{i_j} to a vertex set with less than ϵn vertices.) We now apply Lemma 4 repeatedly if the resulting partition Q is not (k, r)-ϵ-regular. In each iteration, $\frac{ind\ Q}{\binom{n}{k}}$

increases by at most $\epsilon^4/2$. Therefore after at most $2\epsilon^{-4}$ steps we have a refinement Q of P with $|Q| \le f(2\epsilon^{-4})$, where $f(i+1) = 2^{\left[\frac{f(i)}{r}\right]}$ and $f(0) = |P'|$ such that Q is (k, r)-ϵ-regular. Furthermore, any refinement of Q is also (k, r)-ϵ-regular. This completes the proof in Theorem 2.2.

4. SZEMERÉDI-PARTITIONS AND QUASI-RANDOMNESS

In [2] a large class of graph properties, shared by random graphs, are shown to be mutually equivalent in the sense that any graph that satisfies one of the properties must satisfy all the properties. Some examples of the properties (for graphs $G = G(n)$ on n vertices) are the following:

P_1: G has at least $(1 + o(1)n^2/4$ edges and at most $(1 + o(1))n^4/16$ 4-cycles;

P_2: For a fixed s, $s \ge 4$, every fixed graph on s vertices occurs in G almost equally as often as induced subgraphs of G;

P_3: For any subset S of vertices of G, the number $e(S)$ of edges induced by S satisfies $e(S) = \frac{1}{4}|S|^2 + o(n^2)$;

P_4: For almost all choices of vertices u and v, the number $s(u, v)$ of vertices adjacent to either both u and v, or neither u nor v, satisfies $s(u, v) = (1 + o(1))n/2$.

Remark 1. The description of properties P_i contains occurrences of the asymptotic "little-oh" notation $o(1)$. The use of these $o(1)$s can be viewed in two essentially equivalent ways.

In the first way, suppose we have two properties P and P', each with occurrences of $o(1)$, so that $P = P(o(1))$ and $P' = P'(o(1))$. The implication "$P \to P'$" means for each $\epsilon > 0$ there is a $\delta > 0$ so that if a graph on n vertices satisfies $P(\delta)$, then it must also satisfy $P'(\epsilon)$, provided $n > n_0(\epsilon)$. We say P and P' are equivalent if $P \to P'$ and $P' \to P$. The above equivalence class of graph properties including P_i, $i = 1, \ldots, 5$, is termed "quasi-random."

In the second way, we can think of considering an infinite family $\mathcal{F}$ of graphs $G(n)$. We say $\mathcal{F}$ satisfied $P(o(1))$ if for every $\epsilon > 0$, there is an $n_0(\epsilon)$ so that all graphs $G(n)$ satisfy $P(\epsilon)$ if $n \ge n_0(\epsilon)$.

Remark 2. The properties P_i's are satisfied by random graphs with edge density $1/2$. Generalizations to properties for graphs with edge density ρ can be similarly derived.

Recently, Simonovits and Sós [7] have added one more property to the above equivalence class:

P_S: For every $\epsilon > 0$ there exists two integers, $t(\epsilon)$ and $n_0(\epsilon)$ such that the vertex set of G can be partitioned into t almost equinumerous classes $V_1, \ldots, V_t$, for $\epsilon^{-1} \le t \le t(\epsilon)$ satisfying the property that for all but $\epsilon\binom{t}{2}$ pairs (i, j), $1 \le i < j \le t$, for $A \subseteq V_i$, $B \subseteq V_j$ with $|A| \ge \epsilon \frac{n}{t}$, $|B| \ge \epsilon \frac{n}{t}$, we have

$$\left|\frac{e(A,B)}{|A|\,|B|}-\frac{1}{2}\right|<\epsilon\,.$$

The theory of quasi-random can be generalized to k-graphs. In [2] a large class of properties for k-graphs are shown to be equivalent. Furthermore, in [1] a hierarchy of equivalence classes $\mathcal{A}_r$ of graph properties for k-graph was established. Namely, we have

$$\mathcal{A}_0 \supsetneqq \mathcal{A}_1 \supsetneqq \cdots \supsetneqq \mathcal{A}_k$$

where in $\mathcal{A}_0$ there is the property that the number of edges is approximately a half, and in $\mathcal{A}_1$, there is the "almost regular" property. In general, for $i \geq 2$, $\mathcal{A}_r$ includes the following properties of a k-graph G on n vertices.

$Q_1^{(r)}$: For every $(r-1)$-graph H with vertex set $V = V(G)$,

$$|e(G,H) - e(\bar{G},H)| < o(n^k)\,.$$

where $\bar{G}$ denotes the complement of G with edge set $\{x \in \binom{V}{k}: x \not\subseteq E(G)\}$ and $e(G,H) = |E(G,H)|$ denotes the number of edges of G induced by H, i.e.,

$$E(G,H) = \left\{x \in E(G): \binom{x}{r-1} \subseteq E(H)\right\}.$$

The above property was shown to be equivalent to the following property which involves an invariant, so-called "deviation" (that can be interpreted as counting the occurrence of a special type of subgraphs on $k+i$ vertices).

$Q_2^{(r)}$: $dev_r(G) = o(1)$
where $dev_r(G)$ denotes the i-deviation of G which can be defined as follows:

$$dev_r(G) = \frac{1}{n^{k+r}} \sum_{u_1,\ldots,u_{k+r}} \prod_{\epsilon_1}\prod_{\epsilon_2}\cdots\prod_{\epsilon_r} \mu_G(\epsilon_1,\ldots,\epsilon_r,u_{2r+1},\ldots,u_{k+r})$$

where $\epsilon_j \in \{u_{2j-1}, u_{2j}\}$ for $j \leq r$ and $\mu_G(w_1,\ldots,w_k)$ is defined to be 0 if two of the w's are equal; to be -1 if $\{w_1,\ldots,w_k\}$ is an edge in G and to be 1 otherwise.

The main result of this section is to extend the result of Simonovits and Sós [7] to hypergraphs. By using the Regularity Lemmas for hypergraphs as described in previous section, we will show the following property is equivalent to $Q_1^{(r)}$ and therefore is (k,r)-quasi-random.

$Q_S^{(r)}(\epsilon)$: For every $\epsilon > 0$ there exists $t(\epsilon)$ such that $\binom{V}{i-1}$ can be partitioned into t classes $S_1, \ldots, S_t$ with $t \le t(\epsilon)$, so that all but ϵn^k edges are in $E_G(P_{j_1}, \ldots, P_{j_{\binom{k}{r-1}}})$ for $\binom{k}{r-1}$-tuples $j_1, \ldots, j_{\binom{k}{r-1}}$ satisfying

$$\left| \frac{e_G(X_1, \ldots, X_{\binom{k}{r-1}})}{e(X_1, \ldots, X_{\binom{k}{r-1}})} - \frac{1}{2} \right| < \epsilon .$$

Theorem. $Q_1^{(r)} \Leftrightarrow Q_S^{(r)}$.

Proof. To prove $Q_1^{(r)} \Rightarrow Q_s^{(r)}$, we assume $G(n)$ satisfies the property that for every $(r-1)$-graph H

$$|e(G, H) - e(\bar{G}, H)| < \epsilon' n^k , \tag{2}$$

where ϵ' will be chosen later.

We now partition $\binom{V}{r-1}$ into (k, ϵ)-regular subsets $S_1, \ldots, S_t$ where $t < t(\epsilon)^{-1}$. It suffices to show that for any choice of $\binom{k}{r-1}$-tuples $i_1, \ldots, i_{\binom{k}{r-1}}$, $1 \le i_j \le t$, (all i_js are distinct), for $X_j = S_{i_j}$ with $e(X_1, \ldots, X_{\binom{k}{r-1}}) > \epsilon t^{-\binom{k}{r-1}} n^k$ we have

$$\left| \frac{e_G(X_1, \ldots, X_{\binom{k}{r-1}})}{e(X_1, \ldots, X_{\binom{k}{r-1}})} - \frac{1}{2} \right| < \epsilon . \tag{3}$$

We need some definition. Let $H(a_1, \ldots, a_l)$ denote the $(r-1)$-graph with the vertex set V and edge set $X_{a_1} \cup \cdots \cup X_{a_l}$. Let $E_G(l)$ denote the union of $E(G, H(a_1, \ldots, a_l))$ where $\{a_1, \ldots, a_l\}$ ranges over all l-subsets of $\{1, \ldots, \binom{k}{r-1}\}$ and $E(l') = \emptyset$ if $l' \le 0$. We use the lower case e to denote the cardinality of the set. It is not difficult to check that

$$e_G(X_1, \ldots, X_{\binom{k}{r-1}}) = e_G\left(\binom{k}{r-1}\right) - e_G\left(\binom{k}{r-1} - 1\right) + e_G\left(\binom{k}{r-1} - 2\right) - \cdots .$$

$E_{\bar{G}}(l)$ denote the union of $e(\bar{G}, H(a_1, \ldots, a_l))$ for $\{a_1, \ldots, a_l\}$ ranges over l-subsets of $\{1, \ldots, \binom{k}{r-1}\}$. We have, in a similar way, that

$$e_{\bar{G}}(X_1, \ldots, X_{\binom{k}{r-1}}) = e_{\bar{G}}\left(\binom{k}{r-1}\right) - e_{\bar{G}}\left(\binom{k}{r-1} - 1\right) + e_{\bar{G}}\left(\binom{k}{r-1} - 2\right) - \cdots .$$

By (1), we get

$$|e_G(l) - e_{\bar{G}}(l)| \le \binom{\binom{k}{r-1}}{l} \cdot \epsilon n^k$$

Therefore

$$|e_G(X_1, \ldots, X_{\binom{k}{r-1}}) - e_{\bar{G}}(X_1, \ldots, X_{\binom{k}{r-1}})| \le \epsilon' n^k \cdot 2^{\binom{k}{r-1}}$$

Since

$$e_G(X_1, \ldots, X_{\binom{k}{r-1}}) + e_{\bar{G}}(X_1, \ldots, X_{\binom{k}{r-1}}) = e(X_1, \ldots, X_{\binom{k}{r-1}}) \ge \epsilon t^{-\binom{k}{r-1}} n^k$$

we have, by choosing $\epsilon' = \epsilon^2 t^{-k^{r-1}} n^k$,

$$\left| \frac{e_G(X_1, \ldots, X_{\binom{k}{r-1}})}{e(X_1, \ldots, X_{\binom{k}{r-1}})} - \frac{1}{2} \right| < \epsilon$$

(3) is proved.

We now proceed to show $Q_S^{(r)} \Rightarrow Q_1^{(r)}$. For given $\epsilon' = \epsilon/4 > 0$, we apply $Q_S^{(r)}$ to find a partition of $\binom{V}{i-1}$, say $S_1, \ldots, S_t$ which is (k, r)-ϵ'-regular. Let H be a $(r-1)$-graph on V. Let $H(j)$ denote the set $E(H) \cap S_j$, for $j = 1, \ldots, t$.

$Q_S^{(r)}$ implies that for all but $\epsilon' n^k$ edges are in $E_G(H(a_1), \ldots, H(a_{\binom{k}{r-1}}))$ such that

$$\left| \frac{e_G(H(a_1), \ldots, H(a_{\binom{k}{r-1}}))}{e(H(a_1), \ldots, H(a_{\binom{k}{r-1}}))} - \frac{1}{2} \right| < \epsilon' .$$

That is

$$|e_G(H(a_1), \ldots, H(a_{\binom{k}{r-1}})) - e_{\bar{G}}(H(a_1), \ldots, H(a_{\binom{k}{r-1}}))| < 2\epsilon' n^k .$$

We have

$$|e(G, H) - e(\bar{G}, H)|$$

$$\le 3\epsilon' n^k + \sum_{\substack{a_1, \ldots, a_{\binom{k}{r-1}} \\ \epsilon\text{-regular}}} |e_G(a_1, \ldots, a_{\binom{k}{r-1}}) - e_{\bar{G}}(a_1, \ldots, a_{\binom{k}{r-1}})|$$

$$\le \epsilon n^k$$

This completes the proof of Theorem 4.1. ∎

REFERENCES

[1] F. R. K. Chung, Quasi-random classes of hypergraphs, *Random Struct. Alg.*, **1**, 363–382 (1990).

[2] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica*, **9**, 345–362 (1989).

[3] C. Chvatal, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr., The Ramsey number of a graph with bounded maximum degree, *J. Combinat. Theory Ser. B* **34**, 239–243 (1983).

[4] P. Frankl and V. Rödl, private communication (1986).

[5] A. Hajnal, W. Maass, and G. Turán, On the communication complexity of graph properties, *20th STOC*, (1988), pp. 186–191.

[6] W. Paul, N. Pippenger, E. Szemerédi, and W.T. Trotter, Jr., On determinism versus non-determinism and related problems, *24th FOCS*, (1983), 429–438.

[7] M. Simonovits and V. T. Sós, Szemerédi-partition and quasi-randomness, *Random Struct. Alg.*, **2**, 1–10 (1991).

[8] E. Szemerédi, On a set containing no k elements in arithmetic progression, *Acta Arithmetica*, **27**, 199–245 (1975).

[9] E. Szemerédi, Regular partitions of graphs, *Problémes Combinatoires et Théorie des Graphes*, J.-C. Bermond et al., Eds., CNRS, Paris, (1978), pp. 399–401.

[10] E. Szemerédi and W. T. Trotter, Jr., Extremal problems in discrete geometry, *Combinatorica*, **3**, 381–392 (1983).

Received July 25, 1990