APPLICATIONS OF THE REGULARITY LEMMA FOR UNIFORM HYPERGRAPHS

VOJTĚCH RÖDL AND JOZEF SKOKAN

ABSTRACT. In this note we discuss several combinatorial problems that can be addressed by the Regularity Method for hypergraphs. Based on recent results of Nagle, Schacht and the authors, we give here solutions to these problems.

In particular, we prove the following: Let $K_t^{(k)}$ be the complete k-uniform hypergraph on t vertices and suppose an n-vertex k-uniform hypergraph $\mathcal{H}$ contains only $o(n^t)$ copies of $K_t^{(k)}$. Then one can delete $o(n^k)$ edges of $\mathcal{H}$ to make it $K_t^{(k)}$-free.

Similar results were recently obtained by W. T. Gowers.

1. INTRODUCTION

In 1976, Szemerédi proved the Regularity Lemma [Sze78], a theorem which asserts that any graph can be partitioned into bounded number of random-like blocks (ε-regular pairs).

The Regularity Lemma proved to be a very powerful tool in graph theory with many applications (see [KS96, KSSS02] for a survey). Many of these applications are based on the fact that random-like blocks ensured by the Regularity Lemma allow to find small subgraphs. A regularity lemma for 3-uniform hypergraphs that allows the same phenomenon (i.e. finding fixed size subhypergraphs) was considered in [FR02]. This lemma was extended to the case of k-uniform hypergraphs in [RSb].

This paper presents several applications of the lemma from [RSb] combined with the result of [NRSa] and provides complete solutions to the following problems.

1.1. **Erdős-Stone type problem.**

Let $\mathcal{G}$ and $\mathcal{H}$ be two k-uniform hypergraphs. We say that $\mathcal{H}$ is $\mathcal{G}$-free if $\mathcal{H}$ does not contain a subgraph isomorphic to $\mathcal{G}$. Erdős, Frankl, and Rödl [EFR86] proved the following theorem.

Theorem 1.1. *For every $\varepsilon > 0$ and a fixed graph G with chromatic number χ, there exists $n_0(\varepsilon, G) \in \mathbb{N}$ so that every G-free graph H on $n > n_0(\varepsilon, G)$ vertices can be made K_χ-free by removing εn^2 edges.*

As an extension of Theorem 1.1, they proposed to study the following question: For integers $t \geq k \geq 2$, $s \geq 1$, let $K_t^{(k)}$ be the complete k-uniform hypergraph on t vertices and $K_t^{(k)}(s)$ be the complete t-partite k-uniform hypergraph with s vertices in each partite class. Note that $K_t^{(k)}(1) = K_t^{(k)}$.

For $k < t$, denote by $\varphi(k, t, s, n)$ the maximum number of edges needed to be deleted from a $K_t^{(k)}(s)$-free k-uniform hypergraph on n vertices to get a $K_t^{(k)}$-free k-uniform hypergraph. Erdős, Frankl, and Rödl [EFR86] conjectured that for fixed $t > k \geq 2$ and $s \geq 1$ the function $\varphi(k, t, s, n) = o(n^k)$ as n tends to infinity. So far the above conjecture was confirmed to be true for $k = 3$, $t = 4$ in [FR02] and for $k = 3$, $t > 4$ and $k = 4$, $t = 5$ it follows from results in [NR03] and [RSa], respectively. Based on the recent results of Nagle, Rödl, Schacht and Skokan [NRSa, RSb], in this paper, we establish the conjecture for all suitable choices of t, k, and s.

Theorem 1.2. *Suppose $t > k \geq 2$, $s \geq 1$, and let $\mathcal{H}$ be a $K_t^{(k)}(s)$-free k-uniform hypergraph on n vertices. Let $\varepsilon > 0$ be an arbitrary real and $n > n_0(\varepsilon, k, t, s)$. Then it is possible to remove εn^k edges from $\mathcal{H}$ so that the remaining hypergraph is $K_t^{(k)}$-free. In other words,*

$$\varphi(k, t, s, n) = o(n^k) \,.$$

For graphs, i.e. when $k = 2$, this theorem implies that the Turán number $\mathrm{ex}(n, K_t^{(2)}(s))$ (the maximum number of edges in a $K_t^{(2)}(s)$-free graph on n vertices) does not differ from the Turán number $\mathrm{ex}(n, K_t^{(2)})$ by more than εn^2 for n sufficiently large. This combined with well-known Turán Theorem [Tur41] yields

$$\mathrm{ex}(n, K_t^{(2)}(s)) = \left(1 - \frac{1}{t-1} + o(1)\right)\binom{n}{2}. \tag{1.1}$$

Since (1.1) is the statement of the Erdős-Stone Theorem [ES46], Theorem 1.2 can be viewed as a generalization of the Erdős-Stone Theorem to hypergraphs.

In this paper we also prove the following theorem.

Theorem 1.3. *For all integers $t > k \geq 2$ and $\varepsilon > 0$ there exist $\delta = \delta(t, k, \varepsilon) > 0$ and $n_0 = n_0(t, k, \varepsilon) \in \mathbb{N}$ such that the following statement holds.*

Suppose that an n-vertex k-uniform hypergraph $\mathcal{H}$, with $n > n_0$, contains only δn^t copies of $K_t^{(k)}$. Then one can delete εn^k edges of $\mathcal{H}$ to make it $K_t^{(k)}$-free.

As it turns out, it suffices to establish Theorem 1.3 in order to verify Theorem 1.2. We formally prove this observation in Section 2.

Proposition 1.4. *Theorem* 1.3 *implies Theorem* 1.2.

Theorem 1.2 and Theorem 1.3 have several applications. Some of them regard density theorems, among which are Szemerédi's theorem (see Section 1.2 and [FR02]) and related results due to Furstenberg and Katznelson [Gow, Sola, RSTT]. It also has applications in discrete geometry [Solb] and to extremal hypergraph problems [NRSb].

Below we will discuss some of these as well as some other applications in more detail.

1.2. **Szemerédi's Density Theorem.**

Let $r_k(n)$ be the maximum cardinality of a set $A \subseteq [n] := \{1, \ldots, n\}$ containing no arithmetic progression of length k. Answering an old question of Erdős and Turán [ET36], in 1975 Szemerédi [Sze75] established that $r_k(n) = o(n)$ for any fixed integer k.

There are several extremal hypergraph problems, which are closely related to the value of $r_k(n)$. Perhaps first such a problem (related to a well-known $(6, 3)$-configuration) was suggested by Brown, Erdős and Sós [BES73, SEB73] and considered by Ruzsa and Szemerédi in [RS78]. Some other problems of this type were discussed in [RS78, ENR90, Ele]. The extremal problem related to the configuration $\mathcal{F}(k)$ (defined below) was investigated in [FR02] (see also [Röd91]). The particular configuration $\mathcal{F}(k)$ was originally suggested by Frankl.

Let $A_i = \{a_i, b_i\}$ be pairwise disjoint 2-element sets for $i \in [k]$. Define $F_i = \{a_1, \ldots, a_k, b_i\} \setminus \{a_i\}$ and $\mathcal{F} = \mathcal{F}(k) = \{F_1, \ldots, F_k\}$. Note that $|F_j \cap A_i| = 1$ for $1 \leq i, j \leq k$, that is, $\mathcal{F}$ is a k-partite k-uniform hypergraph. Also, $F_i \cap F_j = \{a_1, \ldots, a_k\} \setminus \{a_i, a_j\}$; in particular, $|F_i \cap F_j| = k - 2 < k - 1$ holds for $1 \leq i < j \leq k$. We note that the triple system $\mathcal{F}(3)$ is the $(6, 3)$-configuration considered in [RS78].

Let $\tilde{\mathrm{ex}}(n, \mathcal{F}(k))$ denote $\max |\mathcal{H}|$, $\mathcal{H} \subset [X]^k$, $|X| = n$, such that

(i) $|H \cap H'| \leq k - 2$ holds for all distinct $H, H' \in \mathcal{H}$, and
(ii) $\mathcal{H}$ is $\mathcal{F}(k)$-free.

Note that for any $\mathcal{H}$ satisfying *(i)*

$$|\mathcal{H}| \leq \frac{\binom{n}{k-1}}{\binom{k}{k-1}} \leq \frac{n^{k-1}}{k}$$

must hold. In [FR02, Proposition 2.1-2.2] it was shown that

$$c_k n^{k-2} \times r_k(n) \leq \tilde{\mathrm{ex}}(n, \mathcal{F}(k)) \leq \varphi(k - 1, k, 2, n), \tag{1.2}$$

where c_k is a constant only depending on k. Consequently, Theorem 1.2 implies $r_k(n) = o(n)$, i.e., the famous Density Theorem of Szemerédi.

1.3. **Székely's jack problem.**

The following problem was formulated by Székely [Szé97] (see also [Mat02, pages 226-7]).

For a point $\boldsymbol{c} = (c_1, c_2, \ldots, c_k) \in \{1, 2, \ldots, n\}^k$ we define a *jack* $J(\boldsymbol{c})$ with *center* $\boldsymbol{c}$ as the set of points that differ from c in at most one coordinate. For i, $1 \leq i \leq k$, and fixed $c_1, c_2, \ldots, c_{i-1}, c_{i+1}, \ldots, c_k \in \{1, 2, \ldots, n\}$, we also define a *line* as a set of n points of the form

$$\{(c_1, c_2, \ldots, c_{i-1}, x, c_{i+1}, \ldots, c_k), 1 \leq x \leq n\}.$$

Let $LS(n, k)$ be the maximum cardinality of a system $\mathcal{J}$ of jacks for which

 (1) no two distinct jacks share a common line, and
 (2) $\bigcap_{i=1}^{k} J_i = \emptyset$ for all distinct jacks $J_1, \ldots, J_k \in \mathcal{J}$.

Clearly $LS(n, k) \leq n^{k-1}$, but Székely conjectured that $LS(n, k)/n^{k-1}$ tends to 0 as $n \to \infty$.

One can show that $LS(n, k)$ is closely related to $\tilde{\mathrm{ex}}(kn, \mathcal{F}(k))$. Indeed, in Section 3 we show the following.

Proposition 1.5. *For every integer $k > 1$*

$$\frac{k!}{k^k} \tilde{\mathrm{ex}}(kn, \mathcal{F}(k)) \leq LS(n, k) \leq \tilde{\mathrm{ex}}(kn, \mathcal{F}(k)).$$

Hence, in view of (1.2) and Theorem 1.2 we infer the following.

Theorem 1.6. $LS(n, k) = o(n^{k-1})$.

1.4. **Organization.**

The paper is organized as follows: in the next section we show Proposition 1.4, i.e. how Theorem 1.3 implies Theorem 1.2. Proposition 1.5 is verified in Section 3. In Section 4, we describe the notation and statement of our main tool - the Hypergraph Regularity Lemma. Other results needed in our proof are presented in Section 5. Then, in Section 6, we prove Theorem 1.3.

2. Proof of Proposition 1.4

In the proof of this proposition, we make use of the following lemma, which comes from the theorem of Erdős from [Erd64] by a supersaturation argument (see also [ES83]).

Lemma 2.1. *For every $c > 0$ and positive integers $t \geq 2$ and $s \geq 1$ there exists $n_1 = n_1(c, t, s)$ and $c' > 0$ such that whenever $\mathcal{G}$ is a t-uniform hypergraph with $n > n_1$ vertices and at least cn^t edges, then $\mathcal{G}$ contains $c'n^{ts}$ copies of $K_t^{(t)}(s)$.*

Proof of Proposition 1.4. Let $\varepsilon > 0$ and $k, s, t \in \mathbb{N}$ be given. We must show that for any $K_t^{(k)}(s)$-free k-uniform hypergraph $\mathcal{H}$ on n vertices, n sufficiently large, it is possible to delete εn^k edges from $\mathcal{H}$ to obtain $K_t^{(k)}$-free k-uniform hypergraph. Consequently, $\varphi(k, t, s, n) \leq \varepsilon n^k$ holds.

We start with defining the constants. With intention to apply Theorem 1.3 later, let $\delta > 0$ and $n_0 = n_0(t, k, \varepsilon)$ be the numbers guaranteed by Theorem 1.3. Furthermore, let $n_1 = n_1(\delta, t, s)$ be the number guaranteed by Lemma 2.1 applied with $c = \delta$.

Suppose $\mathcal{H}$ is an arbitrary $K_t^{(k)}(s)$-free k-uniform hypergraph on $n > \max\{n_0, n_1\}$ vertices. Let $\mathcal{G}$ be a t-uniform hypergraph with vertex set $V(\mathcal{G}) = V(\mathcal{H})$ and edge set formed by all cliques $K_t^{(k)}$ of $\mathcal{H}$. Then $\mathcal{G}$ is $K_t^{(t)}(s)$-free because $\mathcal{H}$ is $K_t^{(k)}(s)$-free. By Lemma 2.1, we obtain $|\mathcal{G}| \le \delta n^t$ and, therefore, $\mathcal{H}$ contains at most δn^t copies of $K_t^{(k)}$ as a subgraphs.

Applying Theorem 1.3 yields that $\mathcal{H}$ can be made $K_t^{(k)}$-free by omitting εn^k edges. $\qquad\square$

3. Proof of Proposition 1.5

We start with the second inequality. Let $\mathcal{J}$ be the system of jacks satisfying *(1)* and *(2)* with the maximum size. Our goal is to construct a k-partite k-uniform hypergraph $\mathcal{H}$ satisfying *(i)*, *(ii)*, and $|\mathcal{H}| = |\mathcal{J}|$.

Let $V_1, \ldots, V_k$ be k copies of $\{1, \ldots, n\}$. Then we define $\mathcal{H}$ by setting

$$
\begin{aligned}
V(\mathcal{H}) &= V_1 \cup \ldots \cup V_k, \\
E(\mathcal{H}) &= \big\{\{a_1, \ldots, a_k\} \colon J(a_1, \ldots, a_k) \in \mathcal{J}, a_i \in V_i, i = 1, \ldots, k\big\}.
\end{aligned}
$$

Clearly, $\mathcal{H}$ is a k-partite k-uniform hypergraph on kn vertices with $|\mathcal{H}| = |\mathcal{J}|$. We prove that $\mathcal{H}$ also satisfies *(i)* and *(ii)*. Since two jacks share a line if and only if their centers differ by at most one coordinate, *(1)* implies *(i)*.

Suppose now that $\mathcal{H}$ contains a copy of $\mathcal{F}(k)$ and $F_i = \{a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_k\}$, where $a_i, b_i \in V_i$, $i \in [k]$. By the definition of $\mathcal{H}$, $J_i = J(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_k) \in \mathcal{J}$. Then, however, $(a_1, a_2, \ldots, a_k) \in \bigcap_{i=1}^k J_i$, which is a contradiction to *(2)*. Consequently

$$
LS(n, k) \le \tilde{\mathrm{ex}}(kn, \mathcal{F}(k)).
$$

On the other hand, let $\tilde{\mathcal{H}}$ be a k-uniform hypergraph on kn vertices satisfying *(i)* and *(ii)* such that $|\tilde{\mathcal{H}}| = \tilde{\mathrm{ex}}(nk, \mathcal{F}(k))$. It is a well-known fact that $\tilde{\mathcal{H}}$ contains a k-partite subgraph with k-partition $V(\mathcal{H}) = V_1 \cup \ldots \cup V_k$ such that each partite set has size n and $|\mathcal{H}| \ge (k!/k^k)|\tilde{\mathcal{H}}| = (k!/k^k)\tilde{\mathrm{ex}}(nk, \mathcal{F}(k))$. Let $\mathcal{J}$ be a system of jacks defined by

$$
\mathcal{J} = \big\{J(a_1, \ldots, a_k) \colon \{a_1, \ldots, a_k\} \in E(\mathcal{H})\big\}.
$$

Then *(i)* implies that every two centers of jacks in $\mathcal{J}$ differ by at least two coordinates and, thus, no two jacks in $\mathcal{J}$ share a line. Hence $\mathcal{J}$ satisfies *(1)*.

Suppose that $(a_1, a_2, \ldots, a_k) \in \bigcap_{i=1}^k J_i$ for some distinct jacks $J_1, \ldots, J_k \in \mathcal{J}$. By reordering, we may assume that the center of J_i differs from $(a_1, a_2, \ldots, a_k)$ at the i-th coordinate. Therefore, $J_i = J(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_k)$ for some $b_i \in [n]$. Consequently, $\{F_1, \ldots, F_k\}$, where $F_i =$

$\{a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_k\}$, forms a copy of $\mathcal{F}(k)$ in $\mathcal{H} \subset \tilde{\mathcal{H}}$, which is a contradiction to *(ii)*. Thus,

$$\frac{k!}{k^k}\tilde{\text{ex}}(kn, \mathcal{F}(k)) \le LS(n, k),$$

and we conclude Proposition 1.5 holds.

4. Hypergraph Regularity Lemma

In this section, we present our main tool – the Hypergraph Regularity Lemma from [RSb]. To this end, we need to introduce some notation. For the detailed description of this notation we refer the reader to [RSb].

4.1. **Cylinders and Complexes.**

This paper deals mainly with ℓ-partite k-uniform hypergraphs. We shall refer to such hypergraphs as (ℓ, k)-*cylinders*.

Definition 4.1 (cylinder). *Let $\ell \ge k \ge 2$ be two integers, V be a set, $|V| \ge \ell$, and $V = V_1 \cup \cdots \cup V_\ell$ be a partition of V.*

A k-set $K \in [V]^k$ is crossing if $|V_i \cap K| \le 1$ for every $i \in [\ell]$. We shall denote by $K_\ell^{(k)}(V_1, \ldots, V_\ell)$ the complete (ℓ, k)-cylinder with vertex partition $V_1 \cup \cdots \cup V_\ell$, i.e. the set of all crossing k-sets. Then, an (ℓ, k)-cylinder $\mathcal{G}$ is any subset of $K_\ell^{(k)}(V_1, \ldots, V_\ell)$.

Definition 4.2. *For an (ℓ, k)-cylinder $\mathcal{G}$, where $k > 1$, we shall denote by $\mathcal{K}_j(\mathcal{G})$, $k \le j \le \ell$, the j-uniform hypergraph with the same vertex set as $\mathcal{G}$ and whose edges are precisely those j-element subsets of $V(\mathcal{G})$ that span cliques of order j in $\mathcal{G}$.*

Clearly, the quantity $|\mathcal{K}_j(\mathcal{G})|$ counts the total number of cliques of order j in an (ℓ, k)-cylinder $\mathcal{G}$, $1 < k \le j \le \ell$, and $\mathcal{K}_k(\mathcal{G}) = \mathcal{G}$.

For formal reasons, we find it convenient to extend the above definitions to the case when $k = 1$.

Definition 4.3. *We define an $(\ell, 1)$-cylinder $\mathcal{G}$ as a partition $V_1 \cup \cdots \cup V_\ell$. For an $(\ell, 1)$-cylinder $\mathcal{G} = V_1 \cup \cdots \cup V_\ell$ and $1 \le j \le \ell$, we set $\mathcal{K}_j(\mathcal{G}) = K_\ell^{(j)}(V_1, \ldots, V_\ell)$.*

The concept of "cliques in 1-uniform hypergraphs" is certainly artificial. It fits well, however, to our general description of a complex (see Definition 4.6).

For an (ℓ, k)-cylinder $\mathcal{G}$ and a subset L of vertices in $\mathcal{G}$, where $k \le |L| \le \ell$, we say that L *belongs to* $\mathcal{G}$ if L induces a clique in $\mathcal{G}$.

We will often face a situation when we need to describe that one cylinder 'lies on' another cylinder. To this end, we define the term *underlying cylinder*.

Definition 4.4 (**underlying cylinder**). *Let $\mathcal{F}$ be an $(\ell, k-1)$-cylinder and $\mathcal{G}$ be an (ℓ, k)-cylinder with the same vertex set. We say that $\mathcal{F}$ underlies $\mathcal{G}$ if $\mathcal{G} \subset \mathcal{K}_k(\mathcal{F})$.*

Note that if $k = 2$ and $\mathcal{F} = V_1 \cup \cdots \cup V_\ell$, then $\mathcal{G}$ is an ℓ-partite graph with ℓ-partition $V_1 \cup \cdots \cup V_\ell$.

Definition 4.5 (**density**). *Let $\mathcal{G}$ be a k-uniform hypergraph and $\mathcal{F}$ be a $(k, k-1)$-cylinder. We define a density of $\mathcal{F}$ with respect to $\mathcal{G}$ by*

$$d_{\mathcal{G}}(\mathcal{F}) = \begin{cases} \frac{|\mathcal{G} \cap \mathcal{K}_k(\mathcal{F})|}{|\mathcal{K}_k(\mathcal{F})|} & \text{if } |\mathcal{K}_k(\mathcal{F})| > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

Through this paper, we will work with a sequence of underlying cylinders. To accommodate this situation, we introduce the notion of *complex*.

Definition 4.6 (**complex**). *Let ℓ and k, $\ell \geq k \geq 1$, be two integers. An (ℓ, k)-complex $\boldsymbol{\mathcal{G}}$ is a system of cylinders $\left\{\mathcal{G}^{(j)}\right\}_{j=1}^{k}$ such that*

(a) *$\mathcal{G}^{(1)}$ is an $(\ell, 1)$-cylinder, i.e. $\mathcal{G}^{(1)} = V_1 \cup \cdots \cup V_\ell$,*
(b) *(ℓ, j)-cylinder $\mathcal{G}^{(j)}$ underlies $(\ell, j+1)$-cylinder $\mathcal{G}^{(j+1)}$ for every $j \in [k-1]$, i.e. $\mathcal{G}^{(j+1)} \subset \mathcal{K}_{j+1}(\mathcal{G}^{(j)})$.*

4.2. **Regularity of Cylinders and Complexes.**

Now we define the notion of *regularity* of cylinders.

Definition 4.7. *Let $r \in \mathbb{N}$, $\mathcal{G}$ be a k-uniform hypergraph, and $\tilde{\mathcal{F}}$ be a system of $(k, k-1)$-cylinders $\mathcal{F}_1, \ldots, \mathcal{F}_r$ with the same vertex set as $\mathcal{G}$. We define a density of $\tilde{\mathcal{F}}$ with respect to $\mathcal{G}$ by*

$$d_{\mathcal{G}}(\tilde{\mathcal{F}}) = \begin{cases} \frac{|\mathcal{G} \cap \bigcup_{j=1}^{r} \mathcal{K}_k(\mathcal{F}_j)|}{|\bigcup_{j=1}^{r} \mathcal{K}_k(\mathcal{F}_j)|} & \text{if } |\bigcup_{j=1}^{r} \mathcal{K}_k(\mathcal{F}_j)| > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

Now we define a regular cylinder.

Definition 4.8 (($\boldsymbol{\delta, d, r}$)-**regular cylinder**). *Let $r \in \mathbb{N}$, $\mathcal{F}$ be a $(k, k-1)$-cylinder, and $\mathcal{G}$ be a k-uniform hypergraph. We say that $\mathcal{G}$ is (δ, d, r)-regular with respect to $\mathcal{F}$ if the following condition is satisfied: whenever $\tilde{\mathcal{F}} = \{\mathcal{F}_1, \ldots, \mathcal{F}_r\}$ is a system of subcylinders of $\mathcal{F}$ such that*

$$\left| \bigcup_{j=1}^{r} \mathcal{K}_k(\mathcal{F}_j) \right| \geq \delta \left| \mathcal{K}_k(\mathcal{F}) \right|,$$

then

$$d - \delta \leq d_{\mathcal{G}}(\tilde{\mathcal{F}}) \leq d + \delta.$$

We say that $\mathcal{G}$ is (δ, d, r)-irregular with respect to $\mathcal{F}$ if it is not (δ, d, r)-regular with respect to $\mathcal{F}$. If $r = 1$, we simply say that $\mathcal{G}$ is (δ, d)-regular with respect to $\mathcal{F}$.

We extend the above definition to the case of an $(\ell, k-1)$-cylinder $\mathcal{F}$.

Definition 4.9. *Let $k, \ell, r \in \mathbb{N}$, $\ell \geq k$, $\mathcal{F}$ be an $(\ell, k-1)$-cylinder with an ℓ-partition $\bigcup_{i=1}^{\ell} V_i$, and $\mathcal{G}$ be a k-uniform hypergraph. We say that $\mathcal{G}$ is (δ, d, r)-regular with respect to $\mathcal{F}$ if the restriction $\mathcal{G}\left[\bigcup_{j \in I} V_j\right]$ is (δ, d, r)-regular with respect to $\mathcal{F}\left[\bigcup_{j \in I} V_j\right]$ for all $I \in [\ell]^k$.*

Now we are ready to introduce the concept of regularity for an (ℓ, k)-complex $\mathcal{G}$.

Definition 4.10 ($(\boldsymbol{\delta}, \boldsymbol{r})$-regular complex). *Let $r \in \mathbb{N}$, and let $\boldsymbol{d} = (d_2, \ldots, d_k)$ and $\boldsymbol{\delta} = (\delta_2, \ldots, \delta_k)$ be two vectors of positive real numbers such that $0 < \delta_j < d_j \leq 1$ for all $j = 2, \ldots k$. We say that an (ℓ, k)-complex $\mathcal{G}$ is $(\boldsymbol{\delta}, \boldsymbol{d}, r)$-regular if*

(a) *$\mathcal{G}^{(2)}$ is (δ_2, d_2)-regular with respect to $\mathcal{G}^{(1)}$, and*
(b) *$\mathcal{G}^{(j+1)}$ is $(\delta_{j+1}, d_{j+1}, r)$-regular with respect to $\mathcal{G}^{(j)}$ for every $j \in [k-1] \backslash \{1\}$.*

4.3. Partitions.

Fix an arbitrary integer $k > 1$. For every $j \in [k-1]$, let $a_j \in \mathbb{N}$ and $\psi_j \colon [V]^j \to [a_j]$ be a mapping. Clearly, mapping ψ_1 defines a partition $V = V_1 \cup \ldots \cup V_{a_1}$, where $V_i = \psi_1^{-1}(i)$ for all $i \in [a_1]$.

For every $j \in [a_1]$, denote by $\mathrm{Cross}_j(\psi_1)$ the set of all crossing sets $J \in [V]^j$, i.e. sets for which $|J \cap V_h| \leq 1$ for all $h \in [a_1]$.

Let $([a_1])_<^j = \{(\lambda_1, \ldots, \lambda_j) \colon 1 \leq \lambda_1 < \ldots < \lambda_j \leq a_1\}$ be the set of vectors naturally corresponding to the totally ordered j-element subsets of $[a_1]$. More generally, for a totally ordered set Π of cardinality at least j, let $(\Pi)_<^j$ be the family of totally ordered j-element subsets of Π.

For every $j \in [k-1]$, we consider the projection π_j of $\mathrm{Cross}_j(\psi_1)$ to $([a_1])_<^j$, mapping a set $J \in \mathrm{Cross}_j(\psi_1)$ to the set $\pi_j(J) = (\lambda_1, \ldots, \lambda_j) \in ([a_1])_<^j$ so that $|J \cap V_{\lambda_h}| = 1$ for every $h \in [j]$.

Moreover, for every $1 \leq h \leq j$, let

$$\Psi_h(J) = (x_{\pi_h(H)} = \psi_h(H))_{H \in [J]^h}$$

be a vector with $\binom{j}{h}$ entries indexed by elements from $(\pi_j(J))_<^h$. For our purposes it will be convenient to assume that the entries of $\Psi_h(J)$ are ordered lexicographically with respect to their indices. Notice that

$$\Psi_1(J) \in ([a_1])_<^j \text{ and } \Psi_h(J) \in \underbrace{[a_h] \times \ldots \times [a_h]}_{\binom{j}{h}-\text{times}} = [a_h]^{\binom{j}{h}} \text{ for } h > 1.$$

We define

$$\Psi^{(j)}(J) = (\Psi_1(J), \Psi_2(J), \ldots, \Psi_j(J)).$$

Then $\Psi^{(j)}(J)$ is a vector with $2^j - 1$ entries. Also observe that if we set $\boldsymbol{a} = (a_1, a_2, \ldots, a_{k-1})$ and

$$A(j, \boldsymbol{a}) = ([a_1])_<^j \times \prod_{h=2}^{j} [a_h]^{\binom{j}{h}}, \tag{4.5}$$

then $\Psi^{(j)}(J) \in A(j, \boldsymbol{a})$ for every crossing set $J \in \mathrm{Cross}_j(\psi_1)$. In other words, to each crossing set J we assign a vector $(x_{\pi_h(H)})_{H \subset J}$ with each entry $x_{\pi_h(H)}$ corresponding to a non-empty subset H of J such that $x_{\pi_h(H)} = \psi_h(H) \in [a_h]$, where $h = |H|$.

For two crossing sets $J_1, J_2 \in \mathrm{Cross}_j(\psi_1)$, let us write

$$J_1 \sim J_2 \text{ if } \Psi^{(j)}(J_1) = \Psi^{(j)}(J_2). \tag{4.6}$$

The equivalence relation (4.6) defines a partition of $\mathrm{Cross}_j(\psi_1)$ into at most

$$|A(j, \boldsymbol{a})| = \binom{a_1}{j} \times \prod_{h=2}^{j} a_h^{\binom{j}{h}}$$

parts. Now we describe these parts explicitly using $(2^j - 1)$-dimensional vectors from $A(j, \boldsymbol{a})$.

To this end, we need the following notation. Let $\boldsymbol{x}^{(j)} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j) \in A(j, \boldsymbol{a})$, where $\boldsymbol{x}_1 \in ([a_1])_<^j$ is a totally ordered set and $\boldsymbol{x}_u = (x_\Upsilon)_{\Upsilon \in (\boldsymbol{x}_1)_<^u} \in [a_u]^{\binom{j}{u}}$, $1 < u \leq j$. For a given h-element subset Ξ of $\boldsymbol{x}_1 = (x_1, \ldots, x_j)$ we are interested in a vector $\boldsymbol{x}^{(j)}(\Xi)$ which is "the restriction of $\boldsymbol{x}^{(j)}$ to Ξ". More precisely, we define $\boldsymbol{x}^{(j)}(\Xi)$ as the vector consisting of precisely those entries of $\boldsymbol{x}^{(j)}$ that are indexed by subsets of Ξ. Finally, $\boldsymbol{x}^{(j)}(\Xi) = (\boldsymbol{x}_1^\Xi, \boldsymbol{x}_2^\Xi, \ldots, \boldsymbol{x}_h^\Xi)$, where for $1 \leq u \leq h$,

$$\boldsymbol{x}_u^\Xi = (x_\Upsilon)_{\Upsilon \in (\Xi)_<^u}$$

is the $\binom{h}{u}$-dimensional vector consisting of those entries of $\boldsymbol{x}_u$ that are labeled with ordered u-element subsets of Ξ.

For example, if $\boldsymbol{x}^{(4)} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)$, where

$$\boldsymbol{x}_1 = (2, 3, 5, 7), \ \boldsymbol{x}_2 = (x_{(2,3)}, x_{(2,5)}, x_{(2,7)}, x_{(3,5)}, x_{(3,7)}, x_{(5,7)}),$$
$$\boldsymbol{x}_3 = (x_{(2,3,5)}, x_{(2,3,7)}, x_{(2,5,7)}, x_{(3,5,7)}), \ \boldsymbol{x}_4 = (x_{(2,3,5,7)}),$$

and $\Xi = (2, 5, 7)$, then

$$\boldsymbol{x}_1^\Xi = (2, 5, 7), \ \boldsymbol{x}_2^\Xi = (x_{(2,5)}, x_{(2,7)}, x_{(5,7)}), \ \boldsymbol{x}_3^\Xi = (x_{(2,5,7)}).$$

Definition 4.11. *For each $h \in [j]$ and $\boldsymbol{x}^{(j)} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j) \in A(j, \boldsymbol{a})$, we set*

$$\mathcal{P}^{(h)}(\boldsymbol{x}^{(j)}) = \bigcup_{\Xi \in (\boldsymbol{x}_1)_<^h} \left\{ P \in \mathrm{Cross}_h(\psi_1) \colon \Psi^{(h)}(P) = (\boldsymbol{x}_1^\Xi, \ldots, \boldsymbol{x}_h^\Xi) \right\}. \tag{4.7}$$

Then, the following claim holds.

Claim 4.12. *For every $j \in [k-1]$ and every $\boldsymbol{x}^{(j)} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j) \in A(j, \boldsymbol{a})$, the following is true.*

    (a) *For all $h \in [j]$, $\mathcal{P}^{(h)}(\boldsymbol{x}^{(j)})$ is a (j, h)-cylinder;*

    (b) *$\boldsymbol{\mathcal{P}}(\boldsymbol{x}^{(j)}) = \{\mathcal{P}^{(h)}(\boldsymbol{x}^{(j)})\}_{h=1}^{j}$ is a (j, j)-complex.*

Now we define formally the notion of a partition.

Definition 4.13 (Partition). *Let k be a positive integer, V be a non-empty set, $\boldsymbol{a} = \boldsymbol{a}_{\mathscr{P}} = (a_1, a_2, \ldots, a_{k-1})$ be a vector of positive integers, and $\psi_j \colon [V]^j \to [a_j]$ be a mapping, $j \in [k-1]$. Set $\boldsymbol{\psi} = \{\psi_j \colon j \in [k-1]\}$. Then, we define a partition $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ of $\mathrm{Cross}_{k-1}(\psi_1)$ by[1]*

$$\mathscr{P} = \left\{ \mathcal{P}^{(k-1)}(\boldsymbol{x}) \colon \boldsymbol{x} \in A(k-1, \boldsymbol{a}) \right\}. \tag{4.8}$$

We also define the rank of $\mathscr{P}$ by

$$\mathrm{rank}(\mathscr{P}) = |A(k-1, \boldsymbol{a})|. \tag{4.9}$$

Remark 4.14. Without loss of generality, we may assume that mappings $\psi_j \colon [V]^j \to [a_j]$ are onto for all $j \in [k-1]$. Then we have

$$\binom{a_1}{k-1} \times \prod_{h=2}^{k-1} a_h^{\binom{k-1}{h}} = \mathrm{rank}(\mathscr{P}) \geq a_h$$

for all $h \in [k-1]$.

It follows from Definition 4.13 that

$$\mathscr{P}^{(j)} = \mathscr{P}(j, \boldsymbol{a}, \boldsymbol{\psi}) = \left\{ \mathcal{P}^{(j)}(\boldsymbol{x}^{(j)}) \colon \boldsymbol{x}^{(j)} \in A(j, \boldsymbol{a}) \right\} \tag{4.10}$$

is a partition of $\mathrm{Cross}_j(\psi_1)$ for every $j \in [k-1]$. Therefore, with every partition $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ of $\mathrm{Cross}_{k-1}(\psi_1)$ we have associated a system of partitions $\{\mathscr{P}^{(j)}\}_{j=1}^{k-1}$ defined by (4.10). This system represents the "underlying structure" of $\mathcal{P}$ in the following sense:

Every $\mathcal{P} \in \mathscr{P}$ can be written as $\mathcal{P}^{(k-1)}(\boldsymbol{x})$ for some $\boldsymbol{x} \in A(k-1, \boldsymbol{a})$ (see (4.8)). Since $\mathscr{P} = \mathscr{P}^{(k-1)}$, every $\mathcal{P} \in \mathscr{P}$ uniquely defines $(k-1, k-1)$-complex $\boldsymbol{\mathcal{P}}(\boldsymbol{x}) = \{\mathcal{P}^{(h)}(\boldsymbol{x})\}_{h=1}^{k-1}$ (see Claim 4.12) such that

    • $\mathcal{P} = \mathcal{P}^{(k-1)}(\boldsymbol{x}) \in \boldsymbol{\mathcal{P}}(\boldsymbol{x})$,

    • $\mathcal{P}^{(h)}(\boldsymbol{x})$ consists of $\binom{k-1}{h}$ elements of $\mathscr{P}^{(h)}$ for every $h \in [k-1]$, and

    • $\mathcal{P}^{(h+1)}(\boldsymbol{x}) \subseteq \mathcal{K}_{h+1}(\mathcal{P}^{(h)}(\boldsymbol{x}))$ for every $h \in [k-2]$.

4.4. Polyads.

A regular pair played a central role in the definition of a regular partition for graphs (cf. [Sze78]). In [FR02], where the regularity lemma for triples was considered, this role was played by a 'triad' (which corresponds to a $(3, 2)$-cylinder). In order to define a regular partition $\mathscr{P}$ for a k-uniform

---

[1]If there is no danger of confusion, we will omit the superscript $^{(k-1)}$ in $\boldsymbol{x}^{(k-1)} \in A(k-1, \boldsymbol{a})$ to simplify the text.

hypergraph, we extend these two concepts by introducing *polyads*. Polyads are $(k, k-1)$-cylinders consisting of selected k members of $\mathscr{P}$.

We describe first the environment in which we work.

Setup 4.15. *Let k be a positive integer, V be a non-empty set, $\boldsymbol{a} = (a_1, a_2, \ldots, a_{k-1})$ be a vector of positive integers, $\boldsymbol{\psi} = \{\psi_j \colon j \in [k-1]\}$ be a set of mappings $\psi_j \colon [V]^j \to [a_j]$, $j \in [k-1]$. Let $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ be the partition of $\mathrm{Cross}_{k-1}(\psi_1)$ (see Definition 4.13).*

Recall that for every crossing set $K \in \mathrm{Cross}_k(\psi_1)$ and $h \in [k-1]$, we defined $\Psi_h(K)$ as the $\binom{k}{h}$-dimensional vector

$$\Psi_h(K) = (x_{\pi_h(H)} = \psi_h(H))_{H \in (K)^h_<},$$

where $\pi_h(H) = (\lambda_1, \ldots, \lambda_h) \in ([a_1])^h_<$ is such that $|H \cap V_{\lambda_u}| = 1$ for every $u \in [h]$. We set

$$\hat{\boldsymbol{\Psi}}^{(k-1)}(K) = (\Psi_1(K), \Psi_2(K), \ldots, \Psi_{k-1}(K))$$

and observe that $\hat{\boldsymbol{\Psi}}^{(k-1)}(K)$ is a vector having $\sum_{h=1}^{k-1} \binom{k}{h} = 2^k - 2$ entries. We define set $\hat{A}(k-1, \boldsymbol{a})$ of $(2^k - 2)$-dimensional vectors by

$$\hat{A}(k-1, \boldsymbol{a}) = \hat{A}_{\mathscr{P}}(k-1, \boldsymbol{a}) = ([a_1])^k_< \times \prod_{h=2}^{k-1} [a_h]^{\binom{k}{h}}. \qquad (4.11)$$

Then $\hat{\boldsymbol{\Psi}}^{(k-1)}(K) \in \hat{A}(k-1, \boldsymbol{a})$ for each crossing set $K \in \mathrm{Cross}_k(\psi_1)$.

Let $\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})$. Then we write vector $\hat{\boldsymbol{x}}$ as $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \ldots, \hat{\boldsymbol{x}}_{k-1})$, where $\hat{\boldsymbol{x}}_1 \in ([a_1])^k_<$ is an ordered set and $\hat{\boldsymbol{x}}_u = (\hat{x}_\Upsilon)_{\Upsilon \in (\hat{\boldsymbol{x}}_1)^u_<} \in [a_u]^{\binom{k}{u}}$, is a $\binom{k}{u}$-dimensional vector with entries from $[a_u]$ for every $u > 1$.

Given an ordered set $\Xi \subseteq \hat{\boldsymbol{x}}_1$ with $1 \le |\Xi| = h \le k-1$, we set $\hat{\boldsymbol{x}}_u^\Xi = (\hat{x}_\Upsilon)_{\Upsilon \in (\Xi)^u_<}$ for each $u \in [h]$. We also define

$$\hat{\mathcal{P}}^{(h)}(\hat{\boldsymbol{x}}) = \bigcup_{\Xi \in (\hat{\boldsymbol{x}}_1)^h_<} \left\{ P \in \mathrm{Cross}_h(\psi_1) \colon \Psi^{(h)}(P) = (\hat{\boldsymbol{x}}_1^\Xi, \ldots, \hat{\boldsymbol{x}}_h^\Xi) \right\} \qquad (4.12)$$

for each $h \in [k-1]$, and set $\hat{\boldsymbol{P}}(\hat{\boldsymbol{x}}) = \left\{ \hat{\mathcal{P}}^{(h)}(\hat{\boldsymbol{x}}) \right\}_{h=1}^{k-1}$. Similarly to Claim 4.12, we can prove the following.

Claim 4.16. *For every vector $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \ldots, \hat{\boldsymbol{x}}_{k-1}) \in \hat{A}(k-1, \boldsymbol{a})$, the following statements are true.*

    (a) *For all $h \in [k-1]$, $\hat{\mathcal{P}}^{(h)}(\hat{\boldsymbol{x}})$ is a (k, h)-cylinder;*

    (b) $\hat{\boldsymbol{P}}(\hat{\boldsymbol{x}}) = \left\{ \hat{\mathcal{P}}^{(h)}(\hat{\boldsymbol{x}}) \right\}_{h=1}^{k-1}$ *is a $(k, k-1)$-complex.*

In this paper, $(k, k-1)$-cylinders $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$ will play a special role and we will call them polyads.

Definition 4.17 (**Polyad**). *Let $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ be the partition of* $\mathrm{Cross}_{k-1}(\boldsymbol{\psi}_1)$ *as described in the Setup 4.15. Then, for each vector* $\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})$, *we refer to* $(k, k-1)$-*cylinder* $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$ *as a polyad. We also define the set* $\hat{\mathscr{P}}$ *of all polyads of* $\mathscr{P}$ *by*

$$\hat{\mathscr{P}} = \left\{ \hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}}) \colon \hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a}) \right\}. \tag{4.13}$$

For every polyad $\hat{\mathcal{P}} \in \hat{\mathscr{P}}$ there exists a unique vector $\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})$ such that $\hat{\mathcal{P}} = \hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$. Hence (see also Claim 4.16), each polyad $\hat{\mathcal{P}} \in \hat{\mathscr{P}}$ uniquely defines $(k, k-1)$-complex $\hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{x}}) = \left\{ \hat{\mathcal{P}}^{(i)}(\hat{\boldsymbol{x}}) \right\}_{i=1}^{k-1}$ such that $\hat{\mathcal{P}} \in \hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{x}})$.

Remark 4.18. Similarly to Remark 4.14, if $\psi_j \colon [V]^j \to [a_j]$, $j \in [k-1]$, are mappings defining $\mathscr{P}$, then we have

$$\binom{a_1}{k} \times \prod_{h=2}^{k-1} a_h^{\binom{k}{h}} \geq |\hat{\mathscr{P}}|.$$

For a polyad $\hat{\mathcal{P}}^{(k-1)} \in \hat{\mathscr{P}}$, we define its *(relative) volume* by

$$\mathrm{Vol}(\hat{\mathcal{P}}^{(k-1)}) = \frac{|\mathcal{K}_k(\hat{\mathcal{P}}^{(k-1)})|}{\binom{n}{k}}. \tag{4.14}$$

4.5. Regular partitions.

Definition 4.19 (**equitable $(\mu, \boldsymbol{\delta}, \boldsymbol{d}, r)$-partition**). *Let $\boldsymbol{\delta} = (\delta_2, \ldots, \delta_{k-1})$ and $\boldsymbol{d} = (d_2, \ldots, d_{k-1})$ be two arbitrary but fixed vectors of real numbers between 0 and 1, μ be a number in interval $(0, 1]$ and r be a positive integer. We say that a partition $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ is an equitable $(\mu, \boldsymbol{\delta}, \boldsymbol{d}, r)$-partition if all but at most $\mu\binom{n}{k}$ many k-tuples $K \in [V]^k$ belong to $(\boldsymbol{\delta}, \boldsymbol{d}, r)$-regular complexes $\hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{x}}) = \left\{ \hat{\mathcal{P}}^{(j)}(\hat{\boldsymbol{x}}) \right\}_{j=1}^{k-1}$, where $\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})$. More precisely,*

$$\sum_{\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})} \left\{ \mathrm{Vol}(\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})) \colon \hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{x}}) \text{ is } (\boldsymbol{\delta}, \boldsymbol{d}, r)\text{-regular} \right\} > 1 - \mu. \tag{4.15}$$

For $k = 3$, the above definition describes the equitable partition considered by Frankl and Rödl [FR02, page 139].

The following definition describes a type of partition we are looking for.

Definition 4.20 (**regular partition**). *Let $\mathcal{H}$ be a k-uniform hypergraph with vertex set V, $|V| = n$, and let $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ be any equitable $(\mu, \boldsymbol{\delta}, \boldsymbol{d}, r)$-partition of* $\mathrm{Cross}_{k-1}(\boldsymbol{\psi}_1)$.

A polyad $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$ is called (δ_k, r)-regular (w.r.t $\mathcal{H}$) if

(a) *complex* $\hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{x}}) = \left\{ \hat{\mathcal{P}}^{(j)}(\hat{\boldsymbol{x}}) \right\}_{j=1}^{k-1}$ *is $(\boldsymbol{\delta}, \boldsymbol{d}, r)$-regular, and*

(b) $\mathcal{H}$ is (δ_k, r)-regular[2] with respect to $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$.

We say $\mathscr{P}$ is (δ_k, r)-regular[2] (w.r.t $\mathcal{H}$) if all but at most $\delta_k \binom{n}{k}$ many k-tuples $K \in [V]^k$ are in (δ_k, r)-regular polyads $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$. In other words,

$$\sum_{\hat{\boldsymbol{x}} \in \hat{A}(k-1, \boldsymbol{a})} \left\{ \mathrm{Vol}(\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})) \colon \hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}}) \text{ is } (\delta_k, r)\text{-regular} \right\} > 1 - \delta_k. \quad (4.16)$$

In our theorem, instead of having fixed $\delta_2, \ldots, \delta_{k-1}$, we will prescribe how they depend on densities $d_2, \ldots, d_{k-1}$. Thus, we need to adapt Definitions 4.19 and 4.20.

Definition 4.21 (functionally equitable partition). *Let μ be a number in interval $(0, 1]$, $\delta_{k-1}(d_{k-1})$, $\delta_{k-2}(d_{k-2}, d_{k-1})$, ..., $\delta_2(d_2, \ldots, d_{k-1})$, and $r = r(t, d_2, \ldots, d_{k-1})$ be non-negative functions. Set $\boldsymbol{\delta} = (\delta_2, \ldots, \delta_{k-1})$.*

A partition $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ of $\mathrm{Cross}_{k-1}(\psi_1)$ is a functionally equitable $(\mu, \boldsymbol{\delta}, r)$-partition if there exists a vector $\boldsymbol{d} = (d_2, \ldots, d_{k-1})$ such that $\mathscr{P}$ is an equitable $(\mu, \boldsymbol{\delta}(\boldsymbol{d}), \boldsymbol{d}, r(a_1, \boldsymbol{d}))$-partition (see Definition 4.19).

Definition 4.22 (regular functionally equitable partition). *Let a k-uniform hypergraph $\mathcal{H}$ and a number δ_k, where $0 < \delta_k \le 1$, be given. We say that a functionally equitable $(\mu, \boldsymbol{\delta}, r)$-partition $\mathscr{P}$ is (δ_k, r)-regular (w.r.t. $\mathcal{H}$) if $\mathscr{P}$ is $(\delta_k, r(a_1, \boldsymbol{d}))$-regular (w.r.t. $\mathcal{H}$), where $\boldsymbol{d}$ is the vector from Definition 4.21.*

In [RSb], a regularity lemma for k-uniform hypergraphs was proved.

Theorem 4.23 (Hypergraph Regularity Lemma). *For every integer $k \in \mathbb{N}$, all numbers $\delta_k > 0$ and $\mu > 0$, and any non-negative functions $\delta_{k-1}(d_{k-1})$, $\delta_{k-2}(d_{k-2}, d_{k-1})$, ..., $\delta_2(d_2, \ldots, d_{k-1})$ and $r = r(d_2, \ldots, d_{k-1})$, there exist integers N_k and L_k such that the following holds.*

For every k-uniform hypergraph $\mathcal{H}$ with at least N_k vertices there exists a partition $\mathscr{P} = \mathscr{P}(k-1, \boldsymbol{a}, \boldsymbol{\psi})$ of $\mathrm{Cross}_{k-1}(\psi_1)$ so that

   (i) *$\mathscr{P}$ is a functionally equitable $(\mu, \boldsymbol{\delta}, r)$-partition,*
   (ii) *$\mathscr{P}$ is (δ_k, r)-regular (w.r.t. $\mathcal{H}$), and*
   (iii) *$\mathrm{rank}(\mathscr{P}) = |A(k-1, \boldsymbol{a})| \le L_k$.*

Remark 4.24. The proof of Theorem 4.23 is by induction and implicitly uses the Regularity Lemma of Szemerédi as the base case for the induction. Since the proof doesn't change the sizes of vertex classes once we apply the induction assumption, we may assume that every two vertex classes of the partition guaranteed by the Hypergraph Regularity Lemma differ in sizes by at most 1 (this is guaranteed by the Regularity Lemma of Szemerédi). In other words, if $\mathscr{P}$ is a partition of $\mathrm{Cross}_{k-1}(\psi_1)$, then

$$|\psi_1^{-1}(1)| \le |\psi_1^{-1}(2)| \le \ldots \le |\psi_1^{-1}(a_1)| \le |\psi_1^{-1}(1)| + 1.$$

---

[2]δ_2-regular for $k = 2$

5. Counting Lemma

In our proof, we will need the following theorem, Theorem 5.1, which was proved for general k in [NRSa] (special cases of Theorem 5.1 were shown in [FR02, NR03, RSa, Sko00]).

Before stating Theorem 5.1 we introduce the following setup. Let $\boldsymbol{\mathcal{H}} = \{\mathcal{H}^{(1)}, \ldots, \mathcal{H}^{(k)}\}$ be an (t, k)-complex such that

(1) $\mathcal{H}^{(1)} = V_1 \cup \ldots \cup V_t$, $|V_1| = \cdots = |V_t| = m$,
(2) $\{\mathcal{H}^{(1)}, \ldots, \mathcal{H}^{(k-1)}\}$ is a $((\delta_2, \ldots, \delta_{k-1}), (d_2, \ldots, d_{k-1}), r)$-regular $(t, k-1)$-complex,
(3) $\mathcal{H}^{(k)} = \bigcup_{I \in [t]^k} \mathcal{H}_I^{(k)}$, where $\mathcal{H}_I^{(k)}$ is the restriction of $\mathcal{H}^{(k)}$ on $\bigcup_{i \in I} V_i$,
(4) $\mathcal{H}_I^{(k)}$ is (δ_k, d_I, r)-regular (w.r.t. $\mathcal{H}^{(k-1)}$).

Theorem 5.1 (Counting Lemma). *For any $\nu > 0$ and integers $t > k \geq 2$ the following statement holds. There exist functions $\delta_k'(d_k)$, $\delta_{k-1}'(d_{k-1}, d_k)$, $\ldots$, $\delta_2'(d_2, \ldots, d_k)$, $r'(d_2, \ldots, d_k)$ and a constant m_0 so that for every choice of $d_2, \ldots, d_k \in (0, 1]$ and $m \geq m_0$ the following holds. Whenever $\boldsymbol{\mathcal{H}} = \{\mathcal{H}^{(1)}, \ldots, \mathcal{H}^{(k)}\}$ is a (t, k)-complex satisfying (1)-(4) above with $\delta_k = \delta_k'(d_k)$, $\delta_{k-1} = \delta_{k-1}'(d_{k-1}, d_k), \ldots, \delta_2 = \delta_2'(d_2, \ldots, d_k)$, $r = r'(d_2, \ldots, d_k)$, and $d_I \geq d_k$ for every $I \in [t]^k$, then $\mathcal{H}^{(k)}$ contains at least*

$$(1 - \nu) \prod_{i=2}^{k} d_i^{\binom{t}{i}} \times m^t$$

copies of $K_t^{(k)}$.

6. Proof of Theorem 1.3

Now we prove Theorem 1.3. First, we outline its proof which although conceptually simple, is somewhat tedious and technical. First we apply the Hypergraph Regularity Lemma (Theorem 4.23) to $\mathcal{H}$ with $\delta_k \ll \varepsilon$. Then we delete all k-tuples in irregular and sparse polyads. Our choice of δ_k will guarantee that at most εn^k edges are deleted. We conclude the proof by showing that $\mathcal{H}' = \mathcal{H} - \{\text{deleted edges}\}$ is $K_t^{(k)}$-free.

Proof of Theorem 1.3. Suppose that $\varepsilon > 0$ and $t > k \geq 2$ are given. Set $\nu = 1/2$ and let $\delta_k'(d_k)$, $\delta_{k-1}'(d_{k-1}, d_k)$, $\ldots$, $\delta_2'(d_2, \ldots, d_k)$, and $r'(d_2, \ldots, d_k)$ be functions and m_0 be the constant guaranteed by Theorem 5.1. We also set $d_k = \varepsilon/100$. With intention to apply Theorem 4.23 we choose

$$\begin{aligned}
\delta_k &= \min\{\varepsilon/100, \delta_k'(d_k)\}, \\
\mu &= \varepsilon/100, \\
\delta_i(d_i, \ldots, d_{k-1}) &= \min\{\delta_i'(d_i, \ldots, d_{k-1}, \varepsilon/100), d_i/2\} \text{ for } i = 2, \ldots, k-1, \\
r(t, d_2, \ldots, d_{k-1}) &= r'(d_2, \ldots, d_{k-1}, \varepsilon/100),
\end{aligned}$$

and obtain integers N_k and L_k. Set

$$\delta = \frac{1}{2} \times \left(\frac{1}{2^k L_k} \right)^{2^t} \times \left(\frac{\varepsilon}{100} \right)^{\binom{t}{k}} \times \left(\frac{1}{L_k} \right)^t,$$

and $n_0 = \max\{N_k, m_0 L_k\}$.

Suppose that $\mathcal{H}$ is a k-uniform hypergraph with $n > n_0$ vertices and with at most δn^t copies of $K_t^{(k)}$. Applying Theorem 4.23 to $\mathcal{H}$ yields a partition $\mathscr{P}$ of $\mathrm{Cross}_{k-1}(\psi_1)$ and a vector $\boldsymbol{d} = (d_2, \ldots, d_{k-1}) \in (0,1]^{k-2}$ such that

(i) $\mathscr{P}$ is an equitable $(\mu, \boldsymbol{\delta}(\boldsymbol{d}), \boldsymbol{d}, r(a_1, \boldsymbol{d}))$-partition,
(ii) $\mathscr{P}$ is $(\delta_k, r(a_1, \boldsymbol{d}))$-regular with respect to $\mathcal{H}$.

Moreover, it follows from (iii) and Remark 4.14 that

(iii) $a_1, \ldots, a_{k-1} \le \mathrm{rank}(\mathscr{P}) \le L_k$.

For $i = 2, \ldots, k-1$, we formally fix the constant δ_i by applying the function $\delta_i'(d_i, \ldots, d_{k-1})$ to $d_i, \ldots, d_{k-1}$ coming from the vector $\boldsymbol{d}$. Similarly, we set $r = r'(d_2, \ldots, d_{k-1}, \varepsilon/100)$.

We now delete all edges from $\mathcal{H}$ which are

(a) not in $(\delta_k, r(a_1, \boldsymbol{d}))$-regular polyads (there are at most $(\delta_k + \mu)\binom{n}{k}$ such edges by (i) and (ii)), or
(b) are in polyads $\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}})$ whose densities $d_{\mathcal{H}}(\hat{\mathcal{P}}^{(k-1)}(\hat{\boldsymbol{x}}))$ are smaller than $\varepsilon/100$ (there are at most $(\varepsilon/100)\binom{n}{k}$ such edges).

Note that the edges considered in (a) include all non-crossing (w.r.t. $\mathscr{P}$) edges of $\mathcal{H}$. In (a) and (b) we removed at most $(\delta_k + \mu + \varepsilon/100)\binom{n}{k} \le \varepsilon n^k$ edges. We claim, this yields a subhypergraph $\mathcal{H}'$ without a copy of $K_t^{(k)}$.

To the contrary, suppose there is a copy $\mathcal{F}$ of $K_t^{(k)}$ in $\mathcal{H}'$. Let $V(\mathcal{F}) = \{v_1, v_2, \ldots, v_t\} \subseteq V(\mathcal{H}')$ and suppose $v_\alpha \in V_{h_\alpha}$ for $\alpha = 1, \ldots, t$. Due to the construction of $\mathcal{H}'$ we observe that there exists a (t, k)-complex $\boldsymbol{\mathcal{H}} = \{\mathcal{H}^{(1)}, \mathcal{H}^{(2)}, \ldots, \mathcal{H}^{(k)}\}$ so that

(1) $\mathcal{H}^{(1)} = \bigcup_{\alpha=1}^t V_{h_\alpha}$, where $V_{h_\alpha} \cap V_{h_\beta} = \emptyset$ for $\alpha \ne \beta$ and $|V_{h_1}| = \ldots = |V_{h_t}| = m = n/a_1 > m_0$,
(2) $\{\mathcal{H}^{(1)}, \mathcal{H}^{(2)}, \ldots, \mathcal{H}^{(k-1)}\}$ is a $((\delta_2, \ldots, \delta_{k-1}), (d_2, \ldots, d_{k-1}), r)$-regular $(t, k-1)$-complex,
(3) $\mathcal{H}^{(k)} = \bigcup_{I \in [t]^k} \mathcal{H}_I^{(k)}$, where $\mathcal{H}_I^{(k)} = \mathcal{H}^{(k)}\left[\bigcup_{\alpha \in I} V_{h_\alpha} \right]$, and
(4) $\mathcal{H}_I^{(k)}$ is (δ_k, d_I, r)-regular w.r.t. $\mathcal{H}^{(k-1)}$ and $d_I \ge d_k = \varepsilon/100$ for all $I \in [t]^k$.

By our choice of constants and (1)–(4), the (t, k)-complex $\boldsymbol{\mathcal{H}}$ satisfies the assumptions of Theorem 5.1 and, therefore, $\mathcal{H}^{(k)}$ contains at least

$$\frac{1}{2} \prod_{i=2}^k d_i^{\binom{t}{i}} \times m^t \tag{6.17}$$

copies of $K_t^{(k)}$. In order to complete the argument, we will prove that

$$\frac{1}{2} \prod_{i=2}^{k} d_i^{\binom{t}{i}} \times \left(\frac{1}{a_1}\right)^t > \delta, \tag{6.18}$$

which in view of (6.17) contradicts the assumption that $\mathcal{H} \supseteq \mathcal{H}' \supseteq \mathcal{H}^{(k)}$ contains less than δn^t copies of $K_t^{(k)}$. Hence, it is left to verify (6.18). For that we first show the following.

Claim 6.1. $d_j > \frac{1}{2^k L_k}$ for $j = 2, \ldots, k-1$

Proof. Let $2 \le j \le k-1$ and suppose $d_j \le 1/(2^{k+1} L_k)$. Recall, that $\operatorname{rank}(\mathscr{P}) \le L_k$ and hence (using $\binom{a_1}{j} \le 2^{k-1}\binom{a_1}{k-1}$ and $j-1 \le k-2$) one can show that

$$|\hat{A}(j-1, \boldsymbol{a})| \le \binom{a_1}{j} \prod_{h=2}^{j-1} a_h^{\binom{j}{h}} \le 2^{k-1}\binom{a_1}{k-1} \prod_{h=2}^{k-2} a_h^{\binom{k-1}{h}} \times a_{k-1}$$
$$\le 2^{k-1} \operatorname{rank}(\mathscr{P}) \le 2^{k-1} L_k.$$

We now bound the number of j-tuples in (δ_j, d_j, r)-regular polyads of $\mathscr{P}$. For that we observe $m^j = (n/a_1)^j \le (n/j)^j \le \binom{n}{j}$ and consequently the number of j-tuples in (δ_j, d_j, r)-regular polyads is at most

$$(d_j + \delta_j) \times m^j \times |\hat{A}(j-1, \boldsymbol{a})| \le \frac{3d_j}{2} \times \binom{n}{j} \times |\hat{A}(j-1, \boldsymbol{a})| \tag{6.19}$$
$$\le \frac{3}{2^{k+1} L_k} \times \binom{n}{j} \times 2^{k-1} L_k \le \frac{3}{4} \binom{n}{j}.$$

On the other hand, at most $\mu\binom{n}{k}/\binom{n-j}{k-j} = \mu\binom{n}{j}/\binom{k}{j} \le \mu\binom{n}{j}$ j-tuples are not in regular (j, j)-complexes of the partition $\mathscr{P}$. Indeed, each j-tuple not belonging to a $((\delta_2, \ldots, \delta_{j-1}), (d_2, \ldots, d_{j-1}), r)$-regular $(j, j-1)$-complexes can be extended to $\binom{n-j}{k-j}$ k-tuples. Each such k-tuple necessarily is either not crossing (w.r.t. $\mathscr{P}$) or belongs to a $(\boldsymbol{\delta}, \boldsymbol{d}, r)$-irregular polyad. Since $\mathscr{P}$ is $(\mu, \boldsymbol{\delta}, \boldsymbol{d}, r)$-equitable, there are at most $\mu\binom{n}{k}$ such k-tuples (not belonging to regular polyads).

Moreover, since $\mu < 1/4$ the above observation combined with (6.19) yields a contradiction and hence the claim follows. $\square$

Finally, combining Claim 6.1, the choice of $d_k = \varepsilon/100$ and the fact that $a_1 < L_k$, we infer (6.18). $\square$

7. Concluding remarks

Along the lines of the proof of Theorem 1.3, one can prove the following extension of Theorem 1.1 to hypergraphs which was proposed by Füredi [Für] (see also [Für95]).

Theorem 7.1. *For all integers $t > k \geq 2$ and $\varepsilon > 0$ there exist $\delta = \delta(t, k, \varepsilon) > 0$ and $n_2 = n_2(t, k, \varepsilon) \in \mathbb{N}$ such that the following statement holds.*

Given an t-vertex k-uniform hypergraph $\mathcal{F}$, suppose that an n-vertex k-uniform hypergraph $\mathcal{H}$, with $n > n_2$, contains only δn^t copies of $\mathcal{F}$ as a subgraph. Then one can delete εn^k edges of $\mathcal{H}$ to make it $\mathcal{F}$-free.

Sketch of the proof. Note that for $\mathcal{F} = K_t^{(k)}$ we obtain Theorem 1.3. For the proof of Theorem 7.1 we need a slightly more sophisticated version of the Counting Lemma (Theorem 5.1), which can be fairly easily derived from Theorem 5.1. More precisely we would need a version of the Counting Lemma, which allows us to count arbitrary hypergraphs of fixed size and not only cliques.

The philosophy of the proof of Theorem 7.1 is then very similar. Again, we first regularize a given hypergraph $\mathcal{H}$ (with an appropriate choice of constants and functions) to obtain a regular partition $\mathscr{P}$. Then, as done in the proof of Theorem 1.3, we delete all non-crossing (w.r.t. $\mathscr{P}$) edges of $\mathcal{H}$, as well as, all edges which belong to "sparse" or irregular polyads of $\mathscr{P}$. This way we obtain the hypergraph $\mathcal{H}'$. The choice of constants ensures that $\mathcal{H}'$ differs from $\mathcal{H}$ by only εn^k edges. Now suppose $\mathcal{H}'$ contains a copy of $\mathcal{F}$. We observe that this copy of $\mathcal{F}$ does not necessarily have to be crossing with respect to the vertex partition of $\mathscr{P}$. This is the main difference from the $\mathcal{F} = K_t^{(k)}$ case. However, this copy of $\mathcal{F}$ will "witness" the existence of a homomorphic image $\mathcal{F}'$ of $\mathcal{F}$ with the property that there are $\Omega(n^{|V(\mathcal{F}')|})$ copies of $\mathcal{F}'$ in $\mathcal{H}' \subseteq \mathcal{H}$. (For that we will employ the adjusted Counting Lemma mentioned above.) Then a simple supersaturation argument yields the existence of $\Omega(n^t)$ copies of $\mathcal{F}$ in $\mathcal{H}' \subseteq \mathcal{H}$, which with the right choice of constants gives a contradiction. □

References

[BES73]   W. G. Brown, P. Erdős, and V. T. Sós, *Some extremal problems on r-graphs*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971), Academic Press, New York, 1973, pp. 53–63.

[EFR86]   P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs and Combinatorics **2** (1986), no. 2, 113–121.

[Ele]   G. Elekes, oral communication.

[ENR90]   P. Erdős, J. Nešetřil, and V. Rödl, *On Pisier type problems and results (combinatorial applications to number theory)*, Mathematics of Ramsey theory, Algorithms Combin., vol. 5, Springer, Berlin, 1990, pp. 214–231.

[Erd64]   P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel Journal of Mathematics **2** (1964), 183–190.

[ES46]   P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bulletin of the American Mathematical Society **52** (1946), 1087–1091.

[ES83]   P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), no. 2, 181–192.

[ET36]   P. Erdős and P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261–264.

[FR02]   P. Frankl and V. Rödl, *Extremal problems on set systems*, Random Structures and Algorithms **20** (2002), no. 2, 131–164.

[Für]   Z. Füredi, oral communication.

[Für95]   _____, *Extremal hypergraphs and combinatorial geometry*, Proceedings of the International Congress of Mathematicians (Zürich, Switzerland 1994), Birkhäuser, Basel, 1995, pp. 1343–1352.

[Gow]   W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, manuscript.

[KS96]   J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics—Paul Erdős is eighty, vol. 2 (Keszthely, 1993) (D. Miklós, V. T. Sós, and T. Szőnyi, eds.), Bolyai Society Mathematical Studies, vol. 2, János Bolyai Mathematical Society, Budapest, 1996, pp. 295–352.

[KSSS02] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi, *The regularity lemma and its applications in graph theory*, Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Computer Science, vol. 2292, Springer, Berlin, 2002, pp. 84–112.

[Mat02]   J. Matoušek, *Lectures on discrete geometry*, Graduate Texts in Mathematics, vol. 212, Springer-Verlag, Berlin, 2002, xvi+481 pp.

[NR03]   B. Nagle and V. Rödl, *Regularity properties for triple systems*, Random Structures and Algorithms **23** (2003), no. 3, 264–332.

[NRSa]   B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, manuscript.

[NRSb]   _____, *Extremal hypergraph problems and the regularity method*, in preparation.

[Röd91]   V. Rödl, *Some developments in Ramsey theory*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), Math. Soc. Japan, Springer-Verlag, Tokyo, 1991, pp. 1455–1466.

[RSa]   V. Rödl and J. Skokan, *Counting subgraphs in quasi-random 4-uniform hypergraphs*, submitted.

[RSb]   _____, *Regularity lemma for uniform hypergraphs*, Random Structures and Algorithms, to appear.

[RS78]   I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam, 1978, pp. 939–945.

[RSTT]   V. Rödl, M. Schacht, E. Tengan, and N. Tokushige, *Density theorems and the regularity method*, manuscript.

[SEB73]   V. T. Sós, P. Erdős, and W. G. Brown, *On the existence of triangulated spheres in 3-graphs, and related problems*, Period. Math. Hungar. **3** (1973), no. 3-4, 221–228.

[Sko00]   J. Skokan, *Uniformity of set systems*, Ph.D. thesis, Emory University, 2000, (available at `http://www.mathcs.emory.edu/∼rodl/grads.html`).

[Sola]   J. Solymosi, *Note on a question of Erdős and Graham*, to appear.

[Solb]   _____, *Unavoidable substructures in extremal point-line arrangements*, manuscript.

[Sze75]   E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199–245.

[Sze78]   _____, *Regular partitions of graphs*, Problèmes Combinatoires et Théorie des Graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976) (Paris), Colloques Internationaux CNRS n. 260, 1978, pp. 399–401.

[Szé97]   L. Székely, oral communication, 1997.

[Tur41]   P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452, in Hungarian, with German summary.

Department of Mathematics and Computer Science, Emory University, Atlanta, GA, 30322, USA

E-mail address: `rodl@mathcs.emory.edu`

Department of Mathematics, MC-382, University of Illinois at Urbana–Champaign, 1409 W. Green Street, Urbana, IL, 61801, USA and Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–900 São Paulo, Brazil

E-mail address: `jozef@member.ams.org`