EFFICIENT TESTING OF HYPERGRAPHS (EXTENDED ABSTRACT)

Y. KOHAYAKAWA*, B. NAGLE**, AND V. RÖDL***

Abstract. We investigate a basic problem in combinatorial property testing, in the sense of Goldreich, Goldwasser, and Ron [9, 10], in the context of 3-uniform hypergraphs, or 3-graphs for short. As customary, a 3-graph F is simply a collection of 3-element sets. Let $\mathrm{Forb}_{\mathrm{ind}}(n, F)$ be the family of all 3-graphs on n vertices that contain no copy of F as an induced subhypergraph. We show that the property "$H \in \mathrm{Forb}_{\mathrm{ind}}(n, F)$" is testable, for any 3-graph F. In fact, this is a consequence of a new, basic combinatorial lemma, which extends to 3-graphs a result for graphs due to Alon, Fischer, Krivelevich, and Szegedy [2, 3].

Indeed, we prove that if more than ζn^3 ($\zeta > 0$) triples must be added or deleted from a 3-graph H on n vertices to destroy all induced copies of F, then H must contain $\geq cn^{|V(F)|}$ induced copies of F, as long as $n \geq n_0(\zeta, F)$. Our approach is inspired in [2, 3], but the main ingredients are hypergraph regularity lemmas and counting lemmas for 3-graphs.

1. Introduction

We consider combinatorial property testing, in the sense of Goldreich, Goldwasser, and Ron [9, 10] (see also Ron [14] for a recent survery). We address the problem of testing hypergraph properties; we in fact focus on testing induced subhypergraphs in 3-uniform hypergraphs. (For a recent result on hypergraph property testing, see Czumaj and Sohler [6], where k-colourability is proved to be testable.)

We start with some basic definitions. A *k-uniform hypergraph H*, *k-graph* for short, is a family of k-element sets. When $k = 2$, we speak of *graphs*, and when $k = 3$, we have *triple systems*. A *k-graph property* $\mathcal{P}$ is an infinite class of k-graphs closed under isomorphism. A k-graph H *satisfies* property $\mathcal{P}$ if $H \in \mathcal{P}$. A k-graph H is said to be *ζ-far* from property $\mathcal{P}$ if every k-graph $\widetilde{H} \in \mathcal{P}$ with $V(\widetilde{H}) = V(H)$ differs from H in at least $\zeta|V(H)|^k$ k-tuples of vertices (i.e., the symmetric difference $H \triangle \widetilde{H}$ has size at least $\zeta|V(H)|^k$).

A *ζ-test* for property $\mathcal{P}$ is a randomized algorithm which, given as input a k-graph H and $n = |\bigcup H|$, is allowed to make queries whether any given k-tuple of vertices belongs to H or not, and distinguishes, with high probability, between the case that H satisfies $\mathcal{P}$ and the case that H is ζ-far from $\mathcal{P}$. A property $\mathcal{P}$ is said to be *testable* if, for every $\zeta > 0$, there exists a function $f(\zeta)$ and an ζ-test for $\mathcal{P}$ which makes a total of $f(\zeta)$ queries for any input k-graph. Note that, in particular, the number of queries does not depend on the order of the input k-graph. A ζ-test is

*Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508-900 São Paulo, Brazil. *E-mail address*: `yoshi@ime.usp.br`. Partially supported by MCT/CNPq through ProNEx Programme (Proc. CNPq 664107/1997–4), by CNPq (Proc. 300334/93–1 and 468516/2000–0).

**School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332, USA, and Department of Mathematics, University of Nevada, Reno, Reno, Nevada, 89557, USA. *E-mail address*: `nagle@math.gatech.edu`.

***Department of Mathematics and Computer Science, Emory University, Atlanta, GA, 30322, USA. *E-mail address*: `rodl@mathcs.emory.edu`. Partially supported by NSF Grant 0071261.

Collaboration supported by a CNPq/NSF cooperative grant.

said to be a *one-sided* test if when H satisfies property $\mathcal{P}$, the test determines that this is the case with probability 1. A property $\mathcal{P}$ is said to be *strongly-testable* if for every $\zeta > 0$, there exists a one sided ζ-test for $\mathcal{P}$.

For a k-graph F, we let $\mathrm{Forb}_{\mathrm{ind}}(n, F)$ be the property of all k-graphs on n vertices not containing a copy of F as an induced subhypergraph. Recently, Alon, Fischer, Krivelevich and Szegedy [2, 3] studied the graph properties $\mathrm{Forb}_{\mathrm{ind}}(n, F)$ for arbitrary graphs F. Among other results, they proved that all graph properties of the type $\mathrm{Forb}_{\mathrm{ind}}(n, F)$ are strongly-testable. A statement central to their main result was the following combinatorial theorem.

Theorem 1. *For every $\zeta > 0$ and every graph F, there exists $c > 0$ so that if a graph G on $n > n_0(\zeta, F)$ vertices is ζ-far from $\mathrm{Forb}_{\mathrm{ind}}(n, F)$, then G must contain at least $cn^{|V(F)|}$ copies of F as an induced subgraph.*

Observe that Theorem 1 implies that $\mathrm{Forb}_{\mathrm{ind}}(n, F)$ is a strongly-testable property; if $f = |V(F)|$, then the maximum number of queries required is $O(f^2/c)$: we simply sample α/c random vertices from F, where $\alpha > 0$ is some large enough constant, and check $O(f^2/c)$ suitable adjacencies. This is a strong ζ-test: suppose a given graph G on n vertices contains no copy of F as an induced subgraph. Then, our test will certainly find no copy of F in G as an induced subgraph. Consequently, the test correctly decides with probability 1 that $G \in \mathrm{Forb}_{\mathrm{ind}}(n, F)$. On the other hand, if G is ζ-far from $\mathrm{Forb}_{\mathrm{ind}}(n, F)$, then, by Theorem 1, the graph G contains cn^f copies of F as an induced subgraph. This means that our randomized algorithm (which makes $O(f^2/c)$ queries) is able to locate a copy of F with high probability.

The goal of this paper is to extend Theorem 1 to 3-graphs. Since the connection between testability and the combinatorial property illustrated in Theorem 1 remains unchanged from graphs to hypergraphs, we choose to present our work in a purely combinatorial language.

Theorem 2. *For every $\zeta > 0$ and 3-graph F, there exists $c > 0$ so that if a 3-graph G on $n > n_0(\zeta, F)$ vertices is ζ-far from $\mathrm{Forb}_{\mathrm{ind}}(n, F)$, then G contains at least $cn^{|V(F)|}$ copies of F as an induced subhypergraph.*

We conjecture that Theorems 1 and 2 are true for general k-graphs.

Conjecture 3. *For every $\zeta > 0$ and k-graph F, there exists $c > 0$ so that if a k-graph G on $n > n_0(\zeta, F)$ vertices is ζ-far from $\mathrm{Forb}_{\mathrm{ind}}(n, F)$, then G contains at least $cn^{|V(F)|}$ copies of F as an induced subhypergraph.*

In Sections 1.1 and 1.2 below, we discuss Conjecture 3 and some related problems.

1.1. **On Conjecture 3: the Special Case $F^{(k)} = K_{k+1}^{(k)}$.** The validity of Conjecture 3 in the special case $F^{(k)} = K_{k+1}^{(k)}$, the complete k-graph on $k+1$ vertices, has an interesting connection to the following well known and deep problem concerning arithmetic progressions.

Let $r_k(n)$ be the maximum cardinality of a set of integers $A \subset \{1, \ldots, n\}$ which contains no arithmetic progressions of length k. A conjecture of Erdős and Turán from 1936 stated that $r_k(n) = o(n)$. Roth [15] proved $r_3(n) = o(n)$. In 1975, Szemerédi proved his celebrated result that $r_k(n) = o(n)$, confirming the Erdős–Turán conjecture. Furstenberg [8] and, more recently, Gowers [11] have given alternative proofs. The sharpest result to date for $k = 3$ is due to Bougain [5].

In [16], Ruzsa and Szemerédi solved an extremal combinatorial problem yielding an alternative proof to Roth's theorem [15]. In [7], the following extremal problem was considered. Let $F_1^{(k+1)}$ be the $(k+1)$-graph consisting of two edges intersecting in k points. Let $F_2^{(k+1)}$ be the $(k+1)$-graph with

$2k+2$ vertices $\{a_1, \ldots, a_{k+1}, b_1, \ldots, b_{k+1}\}$ and all $(k+1)$-tuples of the form $\{a_1, \ldots, a_{k+1}, b_i\} \setminus \{a_i\}$, $1 \leq i \leq k+1$. Let $\mathrm{ex}(n, \{F_1^{(k+1)}, F_2^{(k+1)}\})$ denote the maximum number of edges of any $(k+1)$-graph $H^{(k+1)}$ not containing $F_1^{(k+1)}$ or $F_2^{(k+1)}$ as a (not necessarily induced) subhypergraph. In [7], a constructive argument was given showing that

$$\mathrm{ex}(n, \{F_1^{(k+1)}, F_2^{(k+1)}\}) \geq r_{k+1}(n) n^{k-1}. \tag{1}$$

To verify Conjecture 3 for general k, even in the special case when $F^{(k)} = K_{k+1}^{(k)}$, will not be easy. Indeed, we claim that if Conjecture 3 is proved for general k and $F^{(k)} = K_{k+1}^{(k)}$, then one may quickly deduce Szemerédi's theorem $r_k(n) = o(n)$.

Proof of Claim. Indeed, fix $\zeta > 0$ and let $H^{(k+1)}$ be a $(k+1)$-graph on n points and ζn^k edges not containing $F_1^{(k+1)}$ as a subhypergraph. We shall prove that $H^{(k+1)}$ must contain $F_2^{(k+1)}$ if n is large enough. Consequently, for n sufficiently large, we have

$$\mathrm{ex}(n, \{F_1^{(k+1)}, F_2^{(k+1)}\}) < \zeta n^k. \tag{2}$$

Let

$$H^{(k)} = \{K \colon |K| = k,\ K \subset E \text{ for some } E \in H^{(k+1)}\}$$

be the set of all k-tuples contained in some $(k+1)$-tuple of $H^{(k+1)}$. Since $H^{(k+1)}$ contains no copy of $F_1^{(k+1)}$, every k-tuple $K \in H^{(k)}$ belongs to precisely one $(k+1)$-tuple $H \in H^{(k+1)}$. Consequently, $H^{(k)}$ is ζ-far from being $K_{k+1}^{(k)}$-free.

Assuming Conjecture 3 (for $F = K_{k+1}^{(k)}$), we get that $H^{(k)}$ contains

$$cn^{k+1} \gg \zeta n^k = |H^{(k+1)}|$$

copies of $K_{k+1}^{(k)}$.

Let $\{a_1, \ldots, a_{k+1}\}$ span a copy of $K_{k+1}^{(k)}$ in $H^{(k)}$ for which $\{a_1, \ldots, a_{k+1}\} \notin H^{(k+1)}$. Clearly, each $\{a_1, \ldots, a_{k+1}\} \setminus \{a_i\}$, $1 \leq i \leq k+1$, is an edge of $H^{(k)}$. Consequently, there exists b_i such that $\{a_1, \ldots, a_{k+1}, b_i\} \setminus \{a_i\}$ is an edge of $H^{(k+1)}$. Since $|H_1 \cap H_2| < k$ for all distinct $H_1, H_2 \in H^{(k+1)}$, these b_i, $1 \leq i \leq k+1$, have to be distinct. Consequently, $H^{(k+1)}$ contains a copy of $F_2^{(k+1)}$ as a subhypergraph. □

Owing to (1) and (2), we see $r_{k+1}(n) n^{k-1} \leq \mathrm{ex}(n, \{F_1^{(k+1)}, F_2^{(k+1)}\}) < \zeta n^k$. Consequently, $r_{k+1}(n) < \zeta n^k$ for large enough n.

1.2. **On the Non-induced Case.** Note that Conjecture 3 is formulated for classes of k-graphs not containing F as an *induced* subhypergraph (a feature of no importance when F is a clique). One may also consider the class $\mathrm{Forb}(n, F)$ of all k-graphs on n vertices which do not contain a copy of F as a (not necessarily induced) subhypergraph. We state the following analogue to Conjecture 3.

Conjecture 4. *For every $\zeta > 0$ and k-graph F, there exists $c > 0$ so that if a k-graph G on $n > n_0(\zeta, F)$ vertices is ζ-far from* $\mathrm{Forb}(n, F)$*, then G contains at least $cn^{|V(F)|}$ copies of F as a subhypergraph.*

For $k = 2$ and $k = 3$, Conjecture 4 is true. For $k = 2$, Conjecture 4 follows by a standard application of Szemerédi's regularity lemma. For $k = 3$, the result follows from results of [7] and [12]. Recently, Conjecture 4 was proved for $k = 4$ and $F = K_5^{(4)}$ (see [13]).

Since Conjecture 4 is true for $k = 2$, we may define the following (finite) function $C(\zeta, F) = C_2(\zeta, F)$. Recall that the quantification of Conjecture 4 is of the form "$(\forall \zeta, F)(\exists c, n_0)$". For given F and ζ, let $C(\zeta, F)$ be the supremum of all c for which the implication in Conjecture 4 holds (for the given value of ζ and the given graph F) for some large enough n_0. The question of how $C(\zeta, F)$ behaves for a fixed F as a function of ζ has been addressed recently by Alon [4], who proved the following theorem.

Theorem 5. *For a fixed graph F, the function $C(\zeta, F)$ is polynomial in ζ if and only if F is bipartite.*

Preliminary results suggest that Theorem 5 extends to k-graphs. Indeed, if F is a k-partite k-graph with partition classes of size $t_1, \dots, t_k$, then, for a constant $c_1 = c_1(F)$ depending only on F, we have

$$C(\zeta, F) \leq (k/\zeta)^{c_1 t_1 \dots t_k}.$$

On the other hand, there exist non k-partite F (e.g., $F = K_{k+1}^{(k)}$), for which there is a constant $c_2 = c_2(F)$ depending only on F, so that

$$C(\zeta, F) \geq \zeta^{c_2 \log 1/\zeta}.$$

2. Definitions and Notation

2.1. **Graph Concepts.** We begin with some basic notation. As is customary, if X is a set, we write $[X]^k$ for the set of k-element subsets of X. For a graph P and two disjoint sets $X, Y \subset V(P)$, we set $P[X, Y] = \big\{\{x, y\} \in P \colon x \in X,\ y \in Y\big\}$. We define the *density $d_P(X, Y)$ of P with respect to X and $Y \neq \emptyset$* by $d_P(X, Y) = |P[X, Y]|/|X||Y|$.

For a graph P, we let $\mathcal{K}_3(P)$ be the set of vertex sets of triangles in P. Thus, $\{x, y, z\} \in \mathcal{K}_3(P)$ if and only if x, y, and z are mutually adjacent in P.

A graph P with a fixed k-partition $V_1 \cup \dots \cup V_k$ is referred to as a *k-partite cylinder*. We write $P = \bigcup_{1 \leq i < j \leq k} P^{ij}$, where $P^{ij} = P[V_i, V_j]$, $1 \leq i < j \leq k$. For $B \in [k]^3$, we sometimes write $P(B)$ to denote the subgraph of P induced on the vertex set $\bigcup_{i \in B} V_i$. When $k = 3$, we call P a *triad*.

We proceed with the following definitions.

Definition 6 ((α, ε)-regularity). *For α and $\varepsilon > 0$ reals, we say that a bipartite graph P with vertex bipartition $X \cup Y$ is (α, ε)-regular if $\alpha(1 - \varepsilon) < d_P(X_0, Y_0) < \alpha(1 + \varepsilon)$ for every pair of subsets $X_0 \subseteq X$ and $Y_0 \subseteq Y$ with $|X_0| > \varepsilon|X|$ and $|Y_0| > \varepsilon|Y|$.*

Definition 7 ((ℓ, ε, k)-cylinder). *For an integer ℓ and a real $\varepsilon > 0$, we call a k-partite cylinder $P = \bigcup_{1 \leq i < j \leq k} P^{ij}$ an (ℓ, ε, k)-cylinder if each bipartite graph P^{ij}, $1 \leq i < j \leq k$, is $(1/\ell, \varepsilon)$-regular.*

Note that when a triad P is an $(\ell, \varepsilon, 3)$-cylinder with 3-partition satisfying $|V_1| = |V_2| = |V_3| = n$, the number $|\mathcal{K}_3(P)|$ of triangles in P is about n^3/ℓ^3.

2.2. **Hypergraph Concepts.** We refer to any k-partite 3-uniform hypergraph $\mathcal{H}$ with a fixed k-partition $V_1 \cup \dots \cup V_k$ as a *k-partite 3-cylinder*. For $B \in [k]^3$, we set $\mathcal{H}(B)$ to be the set of triples of $\mathcal{H}$ induced on the vertex set $\bigcup_{i \in B} V_i$.

Let a 3-uniform hypergraph $\mathcal{H}$ and a graph P be given so that $V(\mathcal{H}) = V(P)$. We say that P *underlies* $\mathcal{H}$ if $\mathcal{H} \subseteq \mathcal{K}_3(P)$. In the remainder of this paper, we only consider hypergraphs $\mathcal{H}$ together with graphs P that underly them. We continue with the following technical definitions.

Definition 8 (Density of $\vec{Q}$; density of a triad). *Let $\mathcal{H}$ be a 3-partite 3-cylinder with underlying 3-partite cylinder $P = P^{12} \cup P^{23} \cup P^{13}$. Let $\vec{Q} = (Q(1), \dots, Q(r))$ be an r-tuple of 3-partite* *cylinders* $Q(s) = Q^{12}(s) \cup Q^{23}(s) \cup Q^{13}(s)$ *satisfying that, for every* $s \in \{1, 2, \ldots, r\}$ *and for each* $\{i, j\}$, $1 \le i < j \le 3$, *we have* $Q^{ij}(s) \subseteq P^{ij}$. *Let* $\mathcal{K}_3(\overrightarrow{Q}) = \bigcup_{s=1}^{r} \mathcal{K}_3(Q(s))$. *We define the density* $d_{\mathcal{H}}(\overrightarrow{Q})$ *of* $\overrightarrow{Q}$ *as*

$$d_{\mathcal{H}}(\overrightarrow{Q}) = \begin{cases} \dfrac{|\mathcal{H} \cap \mathcal{K}_3(\overrightarrow{Q})|}{|\mathcal{K}_3(\overrightarrow{Q})|} & \text{if } |\mathcal{K}_3(\overrightarrow{Q})| > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

If $r = 1$, *we have the notion of the density* $d_{\mathcal{H}}(P)$ *of a (single) triad* P *with respect to* $\mathcal{H}$.

Definition 9 ((α, δ, r)-regularity)**.** *Let* $\mathcal{H}$ *be a* 3*-partite* 3*-cylinder with underlying* 3*-partite cylinder* $P = P^{12} \cup P^{23} \cup P^{13}$. *Let a positive integer* r *and a real* $\delta > 0$ *be given. We say that the* 3*-cylinder* $\mathcal{H}$ *is* (α, δ, r)*-regular with respect to* P *if for any* r*-tuple of* 3*-partite cylinders* $\overrightarrow{Q} = (Q(1), \ldots, Q(r))$ *as above, if* $\left|\mathcal{K}_3(\overrightarrow{Q})\right| = \left|\bigcup_{s=1}^{r} \mathcal{K}_3(Q(s))\right| > \delta |\mathcal{K}_3(P)|$, *then* $|d_{\mathcal{H}}(\overrightarrow{Q}) - \alpha| < \delta$.

We say $\mathcal{H}$ *is* (δ, r)*-regular with respect to* P *if it is* (α, δ, r)*-regular for some* α. *If the regularity condition fails to be satisfied for every* α, *we say that* $\mathcal{H}$ *is* (δ, r)*-irregular with respect to* P.

2.3. **A Hypergraph Regularity Lemma.** In this section, we state the a Hypergraph Regularity Lemma, due to Frankl and Rödl [7]. First, we state a number of supporting definitions.

Definition 10 ($(\ell, t, \gamma, \varepsilon)$-partition)**.** *Let* V *be a set with* $|V| = N$. *An* $(\ell, t, \gamma, \varepsilon)$*- partition* $\mathcal{P}$ *of* $[V]^2$ *is an (auxiliary) partition* $V = V_1 \cup \cdots \cup V_t$ *of* V, *together with a system of edge-disjoint bipartite graphs* $\{P_\alpha^{ij} : 1 \le i < j \le t,\ 0 \le \alpha \le \ell\}$, *such that*

- (i) $V = \bigcup_{1 \le i \le t} V_i$ *is a* t*-equitable partition, i.e., we have* $\lfloor N/t \rfloor \le |V_i| \le \lceil N/t \rceil$ *for all* $1 \le i \le t$,
- (ii) $\bigcup_{\alpha=0}^{\ell} P_\alpha^{ij} = K(V_i, V_j)$ *for all* i, j, $1 \le i < j \le t$, *where* $K(V_i, V_j)$ *denotes the complete bipartite graph with vertex bipartition* $V_i \cup V_j$,
- (iii) *for all but* $\gamma\binom{t}{2}\ell$ *indices* $1 \le i < j \le t$, $1 \le \alpha \le \ell$, *the graph* P_α^{ij} *is* (ℓ^{-1}, ε)*-regular.*

If we do not have or do not care about (iii), and, moreover, $P_0^{ij} = \emptyset$ for all $1 \le i < j \le t$, we say that $\mathcal{P}$ above is an (ℓ, t)-partition of $[V]^2$. For an (ℓ, t)-partition $\mathcal{P} = \{P_\alpha^{ij} : 1 \le i < j \le t,\ 0 \le \alpha \le \ell\}$ of $[V]^2$, the set of triads generated by $\mathcal{P}$ is

$$\mathrm{Triad}(\mathcal{P}) = \left\{P = P_\alpha^{ij} \cup P_\beta^{jk} \cup P_\gamma^{ik} : 1 \le i < j < k \le t,\ 0 \le \alpha, \beta, \gamma \le \ell\right\}.$$

Definition 11 ((δ, r)-regular partition)**.** *Let* $\mathcal{H}$ *be a* 3*-uniform hypergraph with vertex set* V *where* $|V| = N$. *We say that an* $(\ell, t, \gamma, \varepsilon)$*-partition* $\mathcal{P}$ *of* $[V]^2$ *is* (δ, r)*-regular for* $\mathcal{H}$ *if*

$$\sum \left\{|\mathcal{K}_3(P)| \colon P \in \mathrm{Triad}(\mathcal{P}) \text{ is } (\delta, r)\text{-irregular}\right\} < \delta N^3.$$

Theorem 12 (Hypergraph Regularity Lemma [7])**.** *For every* δ *and* γ *with* $0 < \gamma \le 2\delta^4$, *for all integers* t_0 *and* ℓ_0 *and for all integer-valued functions* $r(t, \ell)$ *and all functions* $\varepsilon(\ell) > 0$, *there exist* T_0, L_0, *and* N_0 *such that any* 3*-uniform hypergraph* $\mathcal{H} \subseteq [N]^3$, $N \ge N_0$, *admits a* $(\delta, r(t, \ell))$*-regular* $(\ell, t, \gamma, \varepsilon(\ell))$*-partition for some* t *and* ℓ *satisfying* $t_0 \le t \le T_0$ *and* $\ell_0 \le \ell \le L_0$.

2.4. **The Counting Lemma.** In this section, we present the Counting Lemma, Lemma 14 below. We begin by describing the context in which this lemma applies.

Setup 13. *Fix integers* f, ℓ, *and* r. *Let* δ *and* $\varepsilon > 0$ *be given, together with an indexed family* $\{\alpha_B \colon B \in [f]^3\}$ *of positive reals. We consider the following conditions on a given hypergraph* $\mathcal{H}$ *and underlying graph* P.

- (i) $\mathcal{H}$ *is an* f*-partite* 3*-cylinder with* f*-partition* $V_1 \cup \cdots \cup V_f$, *where* $|V_1| = \cdots = |V_f| = m$.

(ii) $P = \bigcup_{1 \leq i < j \leq f} P^{ij}$ *is an underlying* (ℓ, ε, f)*-cylinder of* $\mathcal{H}$.

(iii) *For all* $B \in [f]^3$, *the* 3*-partite* 3*-cylinder* $\mathcal{H}(B)$ *is* (α_B, δ, r)*-regular with respect to the triad* $P(B)$ *(cf. Definition 9).*

Finally, suppose $\mathcal{F}$ is a 3-uniform hypergraph with vertex set $V(\mathcal{F}) = [f]$. For $B \in [f]^3$, we let $\rho_B = \alpha_B$ if $B \in \mathcal{F}$, and we let $\rho_B = 1 - \alpha_B$ if $B \notin \mathcal{F}$. In the main lemma of this section, we are concerned with the number of induced, *transversal* copies of $\mathcal{F}$ in $\mathcal{H}$, by which we mean the number of functions $\iota \colon V(\mathcal{F}) = [f] \hookrightarrow V(\mathcal{H})$ with $\iota(i) \in V_i$ for all $1 \leq i \leq f$ that induces an isomorphism of $\mathcal{F}$ onto the image $\mathcal{H}[\operatorname{im} \iota] = \mathcal{H} \cap [\operatorname{im} \iota]^3$ of ι. We denote the set of such ι by $\mathcal{F}_{\mathrm{ind}}(\mathcal{H}; V_1, \ldots, V_f)$.

Lemma 14 (Counting Lemma [12]). *Let* $f \geq 3$ *be a fixed integer. For all constants* α *and* $\beta > 0$, *there exists* $\delta > 0$ *for which the following holds. For all integers* ℓ, *there exist an integer* r *and a real* $\varepsilon > 0$ *so that whenever an* f*-partite* 3*-cylinder* $\mathcal{H}$ *with an underlying cylinder* P *satisfy the conditions of Setup 13 with constants* f, δ, l, r *and* ε *and an indexed family* $\{\alpha_B \colon B \in [f]^3\}$ *of reals, with* $\alpha \leq \alpha_B \leq 1 - \alpha$ *for all* $B \in [f]^3$, *we have*

$$\frac{\Pi_{B \in [f]^3} \rho_B}{\ell^{\binom{f}{2}}} m^f (1 - \beta) \leq |\mathcal{F}_{\mathrm{ind}}(\mathcal{H}; V_1, \ldots, V_f)| \leq \frac{\Pi_{B \in [f]^3} \rho_B}{\ell^{\binom{f}{2}}} m^f (1 + \beta)$$

for any triple system $\mathcal{F}$ *on* $[f]$.

3. The Perfect Reduction Lemma

It is well-known that in the regularity lemma of Szemerédi, one cannot avoid the existence of irregular pairs [1]. It proved to be vital for the proof of Theorem 1 to develop a version of the regularity lemma that, within large subsets of a Szemerédi partition, admits no irregular pairs. Here, we need a similar version of Theorem 12 that admits no irregular triads.

3.1. Preliminary Definitions.

In this section, we consider the following setup.

Setup 15. *Let* $\mathcal{H}$ *be a* 3*-uniform hypergraph on* n *vertices. Let* $V(\mathcal{H}) = V_1 \cup \cdots \cup V_t$ *be a* t*-equitable partition, that is, for all* $1 \leq i \leq t$, *suppose we have* $\lfloor n/t \rfloor \leq |V_i| \leq \lceil n/t \rceil$. *Let* $\mathcal{P} = \left\{ P^{ij}_\alpha \colon 1 \leq i < j \leq t,\ 1 \leq \alpha \leq \ell \right\}$ *be an* (ℓ, t)*-partition of* $[V(\mathcal{H})]^2$.

Observe that $\alpha \geq 1$ in Setup 15, since in (ℓ, t)-partitions we have $P^{ij}_0 = \emptyset$.

Definition 16 ((δ_0, τ)-reduction). *Let* $\mathcal{H}$ *and* $\mathcal{P}$ *as in Setup 15 be given. For* δ_0, $\tau > 0$, *we say the graph* Q *is a* (δ_0, τ)*-reduction of* $\mathcal{P}$ *if the following hold:*

1. *Q has a t-partition $V(Q) = \bigcup_{1 \leq i \leq t} W_i$, where $W_i \subseteq V_i$, $1 \leq i \leq t$, and $|W_1| = \cdots = |W_t| \geq \tau n$.*
2. *For all but $\delta_0 \binom{t}{3} \ell^3$ indices $1 \leq i < j < k \leq t$, $1 \leq \alpha, \beta, \gamma \leq \ell$, the triads $P = P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma$ and $Q_P = Q \cap P$ satisfy $|d_{\mathcal{H}}(Q_P) - d_{\mathcal{H}}(P)| < \delta_0$.*

Definition 17 ($(\delta, r, \varepsilon, \ell')$-perfect reduction). *Let* $\mathcal{H}$ *and* $\mathcal{P}$ *as in Setup 15 be given. Let a graph* Q *be a* (δ_0, τ)*-reduction of* $\mathcal{P}$. *For reals* δ *and* $\varepsilon > 0$ *and for integers* r *and* ℓ', *we say that* Q *is a* $(\delta, r, \varepsilon, \ell')$*-perfect reduction of* $\mathcal{P}$ *if* Q *satisfies the following conditions:*

1. *For each $1 \leq i < j \leq t$, $1 \leq \alpha \leq \ell$, the graph $Q \cap P^{ij}_\alpha$ is $((\ell' \ell)^{-1}, \varepsilon)$-regular.*
2. *For each $1 \leq i < j < k \leq t$, $1 \leq \alpha, \beta, \gamma \leq \ell$, the triad $Q \cap \left(P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma \right)$ is (δ, r)-regular.*

We are now able to state the main lemma in this section.

Lemma 18 (Perfect Reduction Lemma). *For all integers t_0, for all $\delta_0 > 0$, for all functions $\delta(\ell)$, $\varepsilon(\ell)$, and $r(\ell)$, there exist $\tau > 0$ and integers T_0, L_0, L'_0 and N_0 for which the following holds: for any $\mathcal{H}$ on $n \ge N_0$ vertices, there exist integers $t_0 \le t \le T_0$, $1 \le \ell \le L_0$, $1 \le \ell' \le L'_0$, such that*

(a) *there exists an (ℓ, t)-partition $\mathcal{P} = \left\{P^{ij}_\alpha : 1 \le i < j \le t,\ 1 \le \alpha \le \ell\right\}$, as in Setup 15,*

(b) *there exists a (δ_0, τ)-reduction Q of $\mathcal{P}$ (see Definition 16) which is a $(\delta(\ell), r(\ell'\ell), \varepsilon(\ell'\ell), \ell')$-perfect reduction of $\mathcal{P}$ (see Definition 17).*

4. Sketch of the Proof of Theorem 2

4.1. **Definition of the Constants.** Theorem 2 asserts that for any triple system $\mathcal{F}$ and any $\zeta > 0$, there is a constant $c > 0$ for which we have $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})| \ge cn^{|V(\mathcal{F})|}$, whenever $\mathcal{H}$ is ζ-far from $\mathrm{Forb}_{\mathrm{ind}}(n, \mathcal{F})$. Thus, suppose we are given $\mathcal{F}$ and $\zeta > 0$. We let $f = |V(\mathcal{F})|$. All the constants below are defined in terms of f and ζ.

In fact, for simplicity, and also because we do not care about the exact value of the constants (and we do not know their best possible values), we only give the 'hierarchy' governing their sizes. Roughly speaking, in what follows, when we already have B and we say '$A \ll B$', we mean that A should be small enough with respect to B.

We start by letting $\alpha_0 \ll \zeta$. We apply Lemma 14, the Counting Lemma, with $f = |V(\mathcal{F})|$, $\alpha = \alpha_0$, and $\beta = 1/2$. Lemma 14 then gives us a constant $\delta_{\mathrm{CL}} = \delta_{\mathrm{CL}}(f, \alpha_0, 1/2) > 0$. Given that the quantification in Lemma 14 is "$(\forall f, \alpha, \beta)(\exists \delta)(\forall \ell)(\exists r, \varepsilon)$", that lemma gives us functions $r^{f,\alpha,\beta,\delta}(\ell)$ and $\varepsilon^{f,\alpha,\beta,\delta}(\ell)$. We let $r_{\mathrm{CL}}(\ell) = r^{f,\alpha_0,1/2,\delta_{\mathrm{CL}}}_{\mathrm{CL}}(\ell)$ and $\varepsilon_{\mathrm{CL}}(\ell) = \varepsilon^{f,\alpha_0,1/2,\delta_{\mathrm{CL}}}_{\mathrm{CL}}(\ell)$.

One part of the proof we sketch below is based on a lemma, the statement of which is implicit in Section 4.2.2: roughly speaking, this lemma says that there exists a real number

$$\tau' > 0 \tag{4}$$

and there exist integers

$$K_1 = K_1(f, \zeta) \quad \text{and} \quad K_2 = K_2(f, \zeta) \tag{5}$$

for which one may produce the setup described in (S1)–(S3) and (P4)–(P7) (see Section 4.2.2). Now, once we have the constants K_1 and K_2 as in (5), we may complete the definition of our constants.

We now invoke Lemma 18, the Perfect Reduction Lemma, with constants $t_0 \gg 1/\zeta$ and $\delta_0 \ll \zeta$. We let the functions $\delta(\ell)$, $\varepsilon(\ell)$, and $r(\ell)$ in the statement of that lemma be $\delta(\ell) = \delta_{\mathrm{CL}}/K_1^3K_2^3$, $\varepsilon(\ell) = \varepsilon_{\mathrm{CL}}(\ell)$, and $r(\ell) = r_{\mathrm{CL}}(\ell)$. Lemma 18 then gives us a real constant $\tau > 0$ and integers T_0, L_0, L'_0 and N_0.

We remark in passing that τ' in (4) may be taken to be τ/K_1. In fact, K_1 and K_2 in (5) arise from an application of a certain variant of Theorem 12 (see Nagle and Rödl [12]): roughly speaking, K_1 is an upper bound for the number of parts into which we split the vertex set of a given hypergraph and K_2 is an upper bound for the number of parts into which we split certain bipartite graphs.

Finally, we define the constant $c = c(f, \zeta) > 0$ by putting

$$c = \frac{\alpha_0^{\binom{f}{3}} (\tau')^f}{2(L_0 L'_0 K_2)^{\binom{f}{2}}}.$$

Claim 19. *The constant $c = c(f, \zeta) > 0$ defined above will do in Theorem 2.*

4.2. **Proof of Claim 19.** We split the proof into Steps I to IV. Suppose we are given a hypergraph $\mathcal{H}$ on $V = \bigcup \mathcal{H}$ as in the statement of the theorem.

4.2.1. *Step I.* We first apply Lemma 18, the Perfect Reduction Lemma, to obtain

  (a) an (ℓ, t)-partition $\mathcal{P}$ of $[V]^2$ on $V = \bigcup_{1 \le i \le t} V_i$,

  (b) a $(\delta(\ell), r(\ell\ell'), \varepsilon(\ell\ell'), \ell')$-perfect (δ, τ)-reduction Q of $\mathcal{P}$,

where $t_0 \le t \le T_0$, $1 \le \ell \le L_0$, and $1 \le \ell' \le L_0'$.

4.2.2. *Step II.* Applying a variant of Theorem 12 (see Nagle and Rödl [12]) and several further combinatorial arguments (Turán type arguments and Ramsey type arguments), we obtain the setup we shall now describe. The objects that constitute our setup are described in (S1)–(S3) below.

(S1) *For every $1 \le i \le t$, we have a partition $W_i' = \bigcup_{1 \le i' \le f} A_{ii'}$ of a subset W_i' of W_i (see Definition 16). Moreover, we suppose that the $V_i \supset W_i'$ are totally ordered in such a way that $A_{i1} < \cdots < A_{if}$ for all i.*

(S2) *For every $1 \le i < j \le t$, $1 \le i', j' \le f$, and $1 \le \alpha \le \ell$, we have a subgraph ${}^*Q_\alpha^{ij}(i', j')$ of the bipartite graph $Q_\alpha^{ij}(i', j') = Q_\alpha^{ij} \cap K(A_{ii'}, A_{jj'}) = Q \cap P_\alpha^{ij} \cap K(A_{ii'}, A_{jj'})$. Moreover, for every $1 \le i \le t$ and $1 \le i', i'' \le f$ with $i' \ne i''$, we have a subgraph $Q(i; i', i'')$ of $K(A_{ii'}, A_{ii''})$.*

(S3) *We have functions χ_1^*, χ_2^*, and χ_3^* with domains and codomains as follows:*

$$\chi_1^* \colon [t] \to \{\pm 1\}, \quad \chi_2^* \colon [t] \times [t] \setminus \Delta \to \{\pm 1\}^{I_2}, \tag{6}$$

where $\Delta = \{(i, i) \colon 1 \le i \le t\}$ and $I_2 = [\ell] \times [\ell]$, and

$$\chi_3^* \colon [t]^3 \to \{0, \pm 1\}^{I_3}, \tag{7}$$

where $I_3 = [\ell] \times [\ell] \times [\ell]$.

The various objects in (S1)–(S3) satisfy the properties (P4)–(P7) given below.

(P4) *We have $|A_{ii'}| = m$ for all $1 \le i \le t$, $1 \le i' \le f$, where $m \ge \tau' n$ (see (4)).*

(P5) *All ${}^*Q_\alpha^{ij}(i', j')$ $(1 \le i < j \le t,\ 1 \le i', j' \le f,\ 1 \le \alpha \le \ell)$ and all $Q(i; i', i'')$ $(1 \le i \le t,$ $1 \le i', i'' \le f,\ i' \ne i'')$ are $(1/\ell\ell' k_2, \varepsilon_2(\ell\ell' k_2))$-regular for some $k_2 \le K_2 = K_2(f, \zeta)$.*

(P6) *We have the following regularity conditions for the triads defined by the bipartite graphs in* (S2).

  (i) *For any $1 \le i \le t$ and $1 \le i', i'', i''' \le f$, set $Q^i(i', i'', i''') = Q(i; i', i'') \cup Q(i; i'', i''') \cup Q(i; i', i''')$. Then all the triads $Q^i(i', i'', i''')$ are $(\delta_{\mathrm{CL}}, r_{\mathrm{CL}}(\ell\ell' k_2))$-regular.*

  (ii) *For any $1 \le i < j \le t$, $1 \le i', i'', j' \le f$ with $i' \ne i''$, and any $1 \le \alpha, \beta \le \ell$, set $Q_{\alpha\beta}^{ij}(i', i'', j') = {}^*Q_\alpha^{ij}(i', j') \cup {}^*Q_\beta^{ij}(i'', j') \cup Q(i; i', i'')$. Then all the triads $Q_{\alpha\beta}^{ij}(i', i'', j')$ are $(\delta_{\mathrm{CL}}, r_{\mathrm{CL}}(\ell\ell' k_2))$-regular.*

  (iii) *For any $1 \le i < j < k \le t$, $1 \le i', j', k' \le f$, and any $1 \le \alpha, \beta, \gamma \le \ell$, set $Q_{\alpha\beta\gamma}^{ijk}(i', j', k') = {}^*Q_\alpha^{ij}(i', j') \cup {}^*Q_\beta^{jk}(j', k') \cup {}^*Q_\gamma^{ik}(i', k')$. Then all the triads $Q_{\alpha\beta\gamma}^{ijk}(i', j', k')$ are $(\mu', r_{\mathrm{CL}}(\ell\ell' k_2))$-regular, where $\mu' = K_1^3 K_2^3 \delta(\ell) = \delta_{\mathrm{CL}}$.*

(P7) *The triads defined in* (P6) *satisfy the 'density-coherence' properties given below.*

  (i) *For any $1 \le i \le t$, the following holds.*

    (a) *If $\chi_1^*(i) = -1$, then $d_{\mathcal{H}}(Q^i(i', i'', i''')) < 1/2$ for all $1 \le i' < i'' < i''' \le f$.*

    (b) *If $\chi_1^*(i) = 1$, then $d_{\mathcal{H}}(Q^i(i', i'', i''')) \ge 1/2$ for all $1 \le i' < i'' < i''' \le f$.*

(*ii*) *For any* $1 \le i, j \le t$ *with* $i \ne j$ *and any* $1 \le \alpha, \beta \le \ell$, *the following holds.*

(*a*) *If* $(\chi_2^*(i,j))(\alpha,\beta) = -1$, *then* $d_{\mathcal{H}}(Q^{ij}_{\alpha\beta}(i', i'', j')) < 1/2$ *for all* $1 \le i' < i'' \le f$ *and for all* $1 \le j' \le f$.

(*b*) *If* $(\chi_2^*(i,j))(\alpha,\beta) = 1$, *then* $d_{\mathcal{H}}(Q^{ij}_{\alpha\beta}(i', i'', j')) \ge 1/2$ *for all* $1 \le i' < i'' \le f$ *and for all* $1 \le j' \le f$.

(*iii*) *For any* $1 \le i < j < k \le t$ *and any* $1 \le \alpha, \beta, \gamma \le \ell$, *the following holds.*

(*a*) *If* $(\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = -1$, *then* $d_{\mathcal{H}}(Q^{ijk}_{\alpha\beta\gamma}(i', j', k')) < \alpha_0$ *for all* $1 \le i', j', k' \le f$.

(*b*) *If* $(\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = 0$, *then* $\alpha_0 \le d_{\mathcal{H}}(Q^{ijk}_{\alpha\beta\gamma}(i', j', k')) \le 1 - \alpha_0$ *for all* $1 \le i', j', k' \le f$.

(*c*) *If* $(\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = 1$, *then* $d_{\mathcal{H}}(Q^{ijk}_{\alpha\beta\gamma}(i', j', k')) > 1 - \alpha_0$ *for all* $1 \le i', j', k' \le f$.

4.2.3. *Step III.* Based on the functions χ_1^*, χ_2^*, and χ_3^* from (S3), we define a 'perturbation' of $\mathcal{H}$. Using the main hypothesis on $\mathcal{H}$, namely, that any small perturbation of $\mathcal{H}$ does not destroy all the induced copies of $\mathcal{F}$ present in $\mathcal{H}$, we deduce that $\mathcal{H}'$ contains at least one induced copy of $\mathcal{F}$.

Let us define the hypergraph $\mathcal{H}'$. We shall have $\mathcal{H}' \subset [V]^3$, where $V = \bigcup_{1 \le i \le t} V_i$. Let

$$\mathcal{H}'_{\mathrm{tr}} = \{E \in \mathcal{H}' \colon |E \cap V_i| \le 1 \ (1 \le i \le t)\} \tag{8}$$

be the subhypergraph of $\mathcal{H}'$ formed by the 'transversal' triples in $\mathcal{H}'$. Recall that we have the (ℓ, t)-partition $\mathcal{P}$ of $[V]^2$. We thus have the family of triads $\mathrm{Triad}(\mathcal{P})$ given by $\mathcal{P}$. Now, the family of triple systems $\mathcal{K}_3(P)$ ($P \in \mathrm{Triad}(\mathcal{P})$) partitions the family $\{E \in [V]^3 \colon |E \cap V_i| \le 1 \ (1 \le i \le t)\}$ of the transversal triples in $[V]^3$. Therefore, to define $\mathcal{H}'_{\mathrm{tr}}$, we may define $\mathcal{H}'_{\mathrm{tr}} \cap \mathcal{K}_3(P)$ independently for all $P \in \mathrm{Triad}(\mathcal{P})$. For each $P \in \mathrm{Triad}(\mathcal{P})$, if $P = P^{ijk}_{\alpha\beta\gamma} = P^{ij}_{\alpha} \cup P^{jk}_{\beta} \cup P^{ik}_{\gamma}$, we let

$$\mathcal{H}'_{\mathrm{tr}} \cap \mathcal{K}(P) = \begin{cases} \emptyset & \text{if } (\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = -1 \\ \mathcal{H} \cap \mathcal{K}_3(P) & \text{if } (\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = 0 \\ \mathcal{K}_3(P) & \text{if } (\chi_3^*(\{i,j,k\}))(\alpha,\beta,\gamma) = 1. \end{cases} \tag{9}$$

We now proceed to define $\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}$. We have the following natural partition of the non-transversal triples in $[V]^3$:

$$\{E \in [V]^3 \colon |E \cap V_i| \ge 2 \text{ some } i\} = \bigcup_{1 \le i \le t} [V_i]^3 \cup \bigcup_{i,j,\alpha,\beta} \mathcal{K}_3(i,j;\alpha,\beta),$$

where the last union is over all $1 \le i, j \le t$ with $i \ne j$ and $1 \le \alpha, \beta \le \ell$, and

$$\mathcal{K}_3(i,j;\alpha,\beta) = \big\{\{u,v,w\} \in [V_i \cup V_j]^3 \colon u, v \in V_i,\ u < v,\ w \in V_j,\ \{u,w\} \in P^{ij}_{\alpha},\ \{v,w\} \in P^{ij}_{\beta}\big\}.$$

To define $\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}$, we may therefore define $(\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}) \cap [V_i]^3$ $(1 \le i \le t)$ and $(\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}) \cap \mathcal{K}_3(i,j;\alpha,\beta)$ $(1 \le i, j \le t,\ i \ne j,\ 1 \le \alpha, \beta \le \ell)$ independently. Fix $1 \le i \le t$. We let

$$(\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}) \cap [V_i]^3 = \begin{cases} \emptyset & \text{if } \chi_1^*(i) = -1 \\ [V_i]^3 & \text{if } \chi_1^*(i) = 1. \end{cases} \tag{10}$$

Now fix $1 \le i, j \le t$ with $i \ne j$ and $1 \le \alpha, \beta \le \ell$. We let

$$(\mathcal{H}' \setminus \mathcal{H}'_{\mathrm{tr}}) \cap \mathcal{K}_3(i,j;\alpha,\beta) = \begin{cases} \emptyset & \text{if } (\chi_2^*(i,j))(\alpha,\beta) = -1 \\ \mathcal{K}_3(i,j;\alpha,\beta) & \text{if } (\chi_2^*(i,j))(\alpha,\beta) = 1. \end{cases} \tag{11}$$

A crucial property about $\mathcal{H}'$ defined above is given in Claim 20 below.

Claim 20. *The hypergraphs* $\mathcal{H}'$ *and* $\mathcal{H}$ *are* ζ*-close. Therefore,* $\mathcal{H}'$ *contains an induced copy of* $\mathcal{F}$.

4.2.4. *Step IV.* We now fix an induced embedding $\iota\colon \bigcup \mathcal{F} \hookrightarrow \bigcup \mathcal{H}$ of $\mathcal{F}$ into $\mathcal{H}$. Using this embedding, one may obtain a k-partite 3-cylinder $\mathcal{H}^* \subset \mathcal{H}$ with underlying 2-cylinder P to which one may apply Lemma 14, the Counting Lemma.

To simplify the notation, we may and shall assume that the image $\operatorname{im}\iota$ of ι meets only $V_1, \dots, V_g$. Let $h_i = |\operatorname{im}\iota \cap V_i|$ $(1 \leq i \leq g)$. One may then check that $\mathcal{H}^*$ and P may be obtained as described in Claim 21 below.

Claim 21. *There is a choice of indices* $1 \leq \alpha(i,i',j,j') \leq \ell$ $(1 \leq i < j \leq g,\ 1 \leq i', j' \leq f)$ *for which the following holds. Consider the vertex sets* $A_{ii'}$ $(1 \leq i \leq g,\ 1 \leq i' \leq h_i)$, *and consider the k-partite cylinder P on the partition* $\bigcup_{i,i'} A_{ii'}$, *where the union is over all* $1 \leq i \leq g$ *and* $1 \leq i' \leq h_i$, *given by the following bipartite graphs:* (*i*) *for each* $1 \leq i \leq g$ *and each* $1 \leq i' < i'' \leq h_i$, *take* $Q(i;i',i'')$ *and* (*ii*) *for each* $1 \leq i < j \leq g$ *and for each* $1 \leq i' \leq h_i$, $1 \leq j' \leq h_j$, *take* ${}^*Q^{ij}_{\alpha(i,i',j,j')}(i',j')$. *Let* $\mathcal{H}^* \subset \mathcal{H}$ *be the subhypergraph of* $\mathcal{H}$ *formed by the triples in* $\mathcal{H} \cap \mathcal{K}_3(P)$. *Then the hypotheses in Lemma 14 apply, and, consequently,* $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})| \geq cn^k$.

Sketch of Proof. Suppose $\operatorname{im}\iota \cap V_i = \{x_{i1}, \dots, x_{ih_i}\}$ $(1 \leq i \leq g)$, where $x_{i1} < \cdots < x_{ih_i}$ (recall that we suppose that the V_i are totally ordered; see (S1)). Fix $1 \leq i < j \leq g$ and $1 \leq i' \leq h_i$ and $1 \leq j' \leq h_j$. We have $\{x_{ii'}, y_{jj'}\} \in P^{ij}_\alpha$ for some $1 \leq \alpha \leq \ell$. We let $\alpha(i,i',j,j') = \alpha$, and assert that this choice for $\alpha(i,i',j,j')$ $(1 \leq i < j \leq g,\ 1 \leq i' \leq h_i,\ 1 \leq j' \leq h_j)$ will do in Claim 21. To verify our assertion, it suffices to use the properties listed in (P7); we omit the details. □

References

1. N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, *The algorithmic aspects of the regularity lemma*, Journal of Algorithms **16** (1994), no. 1, 80–109.
2. N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, *Efficient testing of large graphs (extended abstract)*, 40th Annual Symposium on Foundations of Computer Science (New York City, NY), IEEE Comput. Soc. Press, 1999, pp. 656–666.
3. ———, *Efficient testing of large graphs*, Combinatorica **20** (2000), no. 4, 451–476.
4. Noga Alon, *Testing subgraphs in large graphs*, Proceedings of the 42nd IEEE Annual Symposium on Foundations of Computer Science (FOCS 2001), 2001, pp. 434–439.
5. J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984. MR **2001h:**11132
6. A. Czumaj and C. Sohler, *Testing hypergraph coloring*, Proc. of ICALP, 2001, pp. 493–505.
7. P. Frankl and V. Rödl, *Extremal problems on set systems*, Random Structures and Algorithms (2002), to appear.
8. Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR 58 #16583
9. Oded Goldreich, Shafi Goldwasser, and Dana Ron, *Property testing and its connection to learning and approximation*, 37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996), IEEE Comput. Soc. Press, Los Alamitos, CA, 1996, pp. 339–348. MR 1 450 632
10. ———, *Property testing and its connection to learning and approximation*, Journal of the Association for Computing Machinery **45** (1998), no. 4, 653–750. MR 1 675 099
11. W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588. MR 1 844 079
12. Brendan Nagle and Vojtěch Rödl, *Regularity properties for triple systems*, submitted, 2000.
13. Vojtěch Rödl and Jozef Skokan, *Uniformity of set systems*, manuscript, 2001.
14. Dana Ron, *Property testing*, Handbook of randomized algorithms (P. M. Pardalos, S. Rajasekaran, J. Reif, and J. D. P. Rolim, eds.), Kluwer Academic Publishers, 2001, to appear.
15. K. F. Roth, *On certain sets of integers*, Journal of the London Mathematical Society **28** (1953), 104–109.
16. I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proceedings of the Fifth Hungarian Colloquium, Keszthely, 1976), Vol. II (Amsterdam–New York), Colloq. Math. Soc. János Bolyai, no. 18, North-Holland, 1978, pp. 939–945.