

# INDEPENDENT SETS IN HYPERGRAPHS

JÓZSEF BALOGH, ROBERT MORRIS, AND WOJCIECH SAMOTIJ

ABSTRACT. Many important theorems and conjectures in combinatorics, such as the theorem of Szemerédi on arithmetic progressions and the Erdős-Stone Theorem in extremal graph theory, can be phrased as statements about families of independent sets in certain uniform hypergraphs. In recent years, an important trend in the area has been to extend such classical results to the so-called ‘sparse random setting’. This line of research has recently culminated in the breakthroughs of Conlon and Gowers and of Schacht, who developed general tools for solving problems of this type. Although these two papers solved very similar sets of longstanding open problems, the methods used are very different from one another and have different strengths and weaknesses.

In this paper, we provide a third, completely different approach to proving extremal and structural results in sparse random sets that also yields their natural ‘counting’ counterparts. We give a structural characterization of the independent sets in a large class of uniform hypergraphs by showing that every independent set is almost contained in one of a small number of relatively sparse sets. We then derive many interesting results as fairly straightforward consequences of this abstract theorem. In particular, we prove the well-known conjecture of Kohayakawa, Łuczak, and Rödl, a probabilistic embedding lemma for sparse graphs, for all 2-balanced graphs. We also give alternative proofs of many of the results of Conlon and Gowers and Schacht, such as sparse random versions of Szemerédi’s theorem, the Erdős-Stone Theorem and the Erdős-Simonovits Stability Theorem, and obtain their natural ‘counting’ versions, which in some cases are considerably stronger. We also obtain new results, such as a sparse version of the Erdős-Frankl-Rödl Theorem on the number of  $H$ -free graphs and, as a consequence of the KLR conjecture, we extend a result of Rödl and Ruciński on Ramsey properties in sparse random graphs to the general, non-symmetric setting. Similar results have been discovered independently by Saxton and Thomason.

## 1. INTRODUCTION

A great many of the central questions in combinatorics fall into the following general framework: Given a finite set  $V$  and a collection  $\mathcal{H} \subseteq \mathcal{P}(V)$  of *forbidden structures*, what can be said about sets  $I \subseteq V$  that do not contain any member of  $\mathcal{H}$ ? For example, the celebrated theorem of Szemerédi [62] states that if  $V = \{1, \dots, n\}$  and  $\mathcal{H}$  is the collection of  $k$ -term arithmetic progressions in  $\{1, \dots, n\}$ , then every set  $I$  that contains no member of  $\mathcal{H}$  satisfies  $|I| = o(n)$ . The archetypal problem studied in extremal graph theory, dating back to the work of Turán [64] and Erdős and Stone [18], is the problem of characterizing such sets  $I$  when  $V$  is the edge set of the complete graph on  $n$  vertices and  $\mathcal{H}$  is the collection of

---

*Date:* May 1, 2012.

Research supported in part by: (JB) NSF CAREER Grant DMS-0745185, UIUC Campus Research Board Grant 11067, and OTKA Grant K76099; (RM) CNPq bolsa de Produtividade em Pesquisa; (WS) ERC Advanced Grant DMMCA and a Trinity College JRF.

copies of some fixed graph  $H$  in  $K_n$ . In this setting, a great deal is known, not only about the maximum size of  $I$  that contains no member of  $\mathcal{H}$ , but also what the largest such sets look like, how many such sets there are, and what the structure of a typical such set is.

A collection  $\mathcal{H} \subseteq \mathcal{P}(V)$  as above is usually referred to as a *hypergraph* on the vertex set  $V$  and any set  $I \subseteq V$  that contains no element (*edge*) of  $\mathcal{H}$  is called an *independent set*. Therefore, one might say that a large part of extremal combinatorics is concerned with studying independent sets in various specific hypergraphs. We might add here that in many natural settings, such as the two mentioned above, the hypergraphs considered are *uniform*, that is, all edges of  $\mathcal{H}$  have the same size.

Although it might at first seem somewhat artificial to study concrete questions in such abstract setting, the past few years have proved that taking such a general approach can be highly beneficial. The recently-proved general transference theorems of Conlon and Gowers [12] and Schacht [58], which imply, among other things, sparse random analogues of the classical theorems of Szemerédi and of Erdős and Stone, were stated in the language of hypergraphs. Roughly speaking, these transference theorems say that if the edges of a hypergraph  $\mathcal{H}$  are sufficiently uniformly distributed, then the independence number of  $\mathcal{H}$  is ‘well-behaved’ with respect to taking subhypergraphs induced by (sufficiently dense) random subsets of the vertex set. More precisely, given  $p \in [0, 1]$  and a finite set  $V$ , we shall write  $V_p$  to denote the *p-random subset* of  $V$ , that is, the random subset of  $V$  in which each element of  $V$  is included with probability  $p$ , independently of all other elements. We write  $\alpha(\mathcal{H})$  and  $v(\mathcal{H})$  to denote the size of the largest independent set and the number of vertices in a hypergraph  $\mathcal{H}$ , respectively. The results of Conlon and Gowers [12] and Schacht [58] imply, in particular, that if the distribution of the edges of some uniform hypergraph  $\mathcal{H}$  is sufficiently ‘balanced’, then with probability tending to 1 as  $v(\mathcal{H}) \rightarrow \infty$ ,

$$\alpha(\mathcal{H}[V(\mathcal{H})_p]) \leq p\alpha(\mathcal{H}) + o(pv(\mathcal{H})),$$

provided that  $p$  is sufficiently large.

In this work, we give an approximate structural characterization of the family of all independent sets in uniform hypergraphs whose edge distribution satisfies a certain natural boundedness condition. More precisely, we shall prove that the independent sets of each such hypergraph  $\mathcal{H}$  exhibit a certain clustering phenomenon. Our main result, Theorem 2.2 below, states that the family  $\mathcal{I}(\mathcal{H})$  of independent sets in  $\mathcal{H}$  admits a partition into relatively few classes such that all members of each class are essentially contained in a single subset of  $V(\mathcal{H})$  that is almost independent, that is, it contains only a tiny proportion of all the edges of  $\mathcal{H}$ . This somewhat abstract statement has surprisingly many deep and interesting consequences, some of which we list in the remainder of this section. We remark that Theorem 2.2 was partly inspired by the work of Kleitman and Winston [37], who implicitly considered a statement of this type in the setting of graphs (2-uniform hypergraphs) and subsequently used it to bound the number of  $n$ -vertex graphs without a 4-cycle. We also note that a result similar to Theorem 2.2 was independently proved by Saxton and Thomason [57], who also use it to derive many of the statements that we present in Sections 1.1–1.3.

**1.1. The number of sets with no  $k$ -term arithmetic progression.** The celebrated theorem of Szemerédi [62] says that for every  $k \in \mathbb{N}$ , the largest subset of  $\{1, \dots, n\}$  that contains no  $k$ -term arithmetic progression (AP) has  $o(n)$  elements. It immediately follows

that there are only  $2^{o(n)}$  subsets of  $\{1, \dots, n\}$  with no  $k$ -term AP. Our first result can be viewed as a sparse analogue of this statement.

**Theorem 1.1.** *For every positive  $\beta$  and every  $k \in \mathbb{N}$ , there exist constants  $C$  and  $n_0$  such that the following holds. For every  $n \in \mathbb{N}$  with  $n \geq n_0$ , if  $m \geq Cn^{1-1/(k-1)}$ , then there are at most*

$$\binom{\beta n}{m}$$

*$m$ -subsets of  $\{1, \dots, n\}$  that contain no  $k$ -term AP.*

We shall deduce Theorem 1.1 from our main theorem, Theorem 2.2, and a robust version of Szemerédi's theorem, see Section 4. The sparse random analogue of Szemerédi's theorem, proved by Schacht [58] and independently by Conlon and Gowers [12], follows as an easy corollary of Theorem 1.1. Following [12], we shall say that a set  $A \subseteq \mathbb{N}$  is  $(\delta, k)$ -Szemerédi if every subset  $B \subseteq A$  with at least  $\delta|A|$  elements contains a  $k$ -term AP. For the sake of brevity, let  $[n] = \{1, \dots, n\}$  and recall that  $[n]_p$  denotes the  $p$ -random subset of  $[n]$ .

**Corollary 1.2.** *For every  $\delta \in (0, 1)$  and every  $k \in \mathbb{N}$ , there exists a constant  $C$  such that the following holds. If  $p_n \geq Cn^{-1/(k-1)}$  for all sufficiently large  $n$ , then*

$$\lim_{n \rightarrow \infty} \mathbb{P}([n]_{p_n} \text{ is } (\delta, k)\text{-Szemerédi}) = 1.$$

We remark that Theorem 1.1 and Corollary 1.2 are both sharp up to the value of the constant  $C$ , see the discussion in Section 4, where both of these statements are proved.

Our main result has a variety of other applications in additive combinatorics, see for example [1, 2] where, jointly with Alon, we used a much simpler version of it to count sum-free sets of fixed size in various Abelian groups and the set  $[n]$ . In Section 4, we shall mention two other applications: a generalization of Theorem 1.1 to higher dimensions and a sparse counting version of a theorem of Sárközy [56] and (independently) Furstenberg [24] on square differences in the integers. In each case, the random version (which was proved in [12, 58]) follows as an easy corollary.

**1.2. Turán's problem in random graphs.** A famous theorem of Erdős and Stone [18] states that the maximum number of edges in an  $H$ -free graph on  $n$  vertices, the *Turán number for  $H$* , denoted  $\text{ex}(n, H)$ , satisfies

$$\text{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \binom{n}{2}, \quad (1)$$

where  $\chi(H)$  is the chromatic number of  $H$ . The analogue of this theorem for the Erdős-Rényi random graph  $G(n, p)$  was first studied by Babai, Simonovits, and Spencer [4], who proved that *asymptotically almost surely* (a.a.s. for short), i.e., with probability tending to 1 as  $n \rightarrow \infty$ , the largest triangle-free subgraph of  $G(n, 1/2)$  is bipartite, and by Frankl and Rödl [20], who proved that if  $p \geq n^{-1/2+\epsilon}$  then a.a.s. the largest triangle-free subgraph of  $G(n, p)$  has  $n^2/8 + o(n^2)$  edges. The systematic study of the Turán problem in  $G(n, p)$  was initiated by Haxell, Kohayakawa, and Łuczak [33, 34] and by Kohayakawa, Łuczak, and

\*  $k = \chi(H) - 1$   
 Partition  $K_n$  into  $k$  blocks of  
 size  $n/k$ : graph induced  
 by  $k$ -partite graph is  $H$ -free  
 if random edges  $\sim \binom{n}{2} (1 - \frac{1}{k})$ .

Rödl [41], who posed the following problem. For a fixed graph  $H$ , determine necessary and sufficient conditions on a sequence  $\mathbf{p} \in [0, 1]^N$  of probabilities such that, a.a.s.,

$$\text{ex}(G(n, \mathbf{p}_n), H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \binom{n}{2} p_n, \quad (2)$$

where  $\text{ex}(G, H)$  denotes the maximum number of edges in an  $H$ -free subgraph of  $G$ .

By considering a random  $(\chi(H) - 1)$ -partition of the vertex set of  $G(n, p)$ , it is straightforward to show that the inequality  $\text{ex}(G(n, p), H) \geq \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \binom{n}{2} p$  holds for every  $p \in [0, 1]$ . On the other hand, if the number of copies of some subgraph  $H' \subseteq H$  in  $G(n, p)$  is much smaller than the number of edges in  $G(n, p)$ , then the converse inequality cannot hold, since one can make any graph  $H$ -free by removing from it one edge from each copy of  $H'$ . This observation motivates the notion of 2-density of  $H$ , denoted by  $m_2(H)$ , which is defined by

$$m_2(H) = \max \left\{ \frac{e(H') - 1}{v(H') - 2} : H' \subseteq H \text{ with } v(H') \geq 3 \right\}. \quad (3)$$

It now follows easily that for every graph  $H$  with maximum degree at least 2 and every  $\delta \in (0, 1/(\chi(H) - 1))$ , there exists a positive constant  $c$  such that if  $p_n \leq cn^{-1/m_2(H)}$ , then a.a.s.

$$\text{ex}(G(n, p_n), H) > \left(1 - \frac{1}{\chi(H) - 1} + \delta\right) \binom{n}{2} p_n.$$

If  $n^v p_n \ll n^{2p}$  then  
 one can remove  $\epsilon$  fraction  
 of edges and become  $H$ -free

It was conjectured by Haxell, Kohayakawa, and Łuczak [33] and Kohayakawa, Łuczak, and Rödl [41] that the above simple argument, removing an arbitrary edge from each copy of  $H'$  in  $G(n, p)$ , is the main obstacle that prevents (2) from holding asymptotically almost surely. The conjecture, often referred to as Turán's theorem for random graphs, has attracted considerable attention in the past fifteen years. Numerous partial results and special cases had been established by various researchers [21, 26, 29, 33, 34, 41, 43, 61] before the conjecture was finally proved by Conlon and Gowers [12] (under the assumption that  $H$  is strictly 2-balanced<sup>1</sup>) and by Schacht [58].

**Theorem 1.3.** For every graph  $H$  with  $\Delta(H) \geq 2$  and every positive  $\delta$ , there exists a positive constant  $C$  such that if  $p_n \geq Cn^{-1/m_2(H)}$ , then a.a.s.

$$\text{ex}(G(n, p_n), H) \leq \left(1 - \frac{1}{\chi(H) - 1} + \delta\right) \binom{n}{2} p_n.$$

Our methods give yet another proof of Theorem 1.3 in the case when  $H$  is 2-balanced. Note that most natural graphs, such as cycles and cliques, are 2-balanced. In fact, we shall deduce from our main result, Theorem 2.2, a version of the general transference theorem of Schacht [58, Theorem 3.3], which easily implies Theorem 1.3 for such graphs  $H$ . Our version of Schacht's transference theorem, Theorem 5.2, is stated and proved in Section 5. We then, in Section 7, use it to derive a natural generalization of Theorem 1.3 to  $t$ -balanced  $t$ -uniform hypergraphs, Theorem 7.2, which was also first proved in [12] and [58].

<sup>1</sup>A graph  $H$  is 2-balanced if the maximum in (3) is achieved with  $H' = H$ , that is, if  $m_2(H) = \frac{e(H) - 1}{v(H) - 2}$ . It is strictly 2-balanced if  $m_2(H) > m_2(H')$  for every proper subgraph  $H' \subsetneq H$ .

Our methods also yield the following sparse random analogue of the famous stability theorem of Erdős and Simonovits [14, 60], originally proved by Conlon and Gowers [12] in the case when  $H$  is strictly 2-balanced and then extended to arbitrary  $H$  by Samotij [55], who adapted the argument of Schacht [58] for this purpose.

**Theorem 1.4.** *For every 2-balanced graph  $H$  with  $\Delta(H) \geq 2$  and every positive  $\delta$ , there exist positive constants  $C$  and  $\varepsilon$  such that if  $p_n \geq Cn^{-1/m_2(H)}$ , then a.a.s. the following holds. Every  $H$ -free subgraph of  $G(n, p_n)$  with at least*

$$\left(1 - \frac{1}{\chi(H) - 1} - \varepsilon\right) \binom{n}{2} p_n$$

*edges may be made  $(\chi(H) - 1)$ -partite by removing from it at most  $\delta n^2 p_n$  edges.*

Similarly as with Theorem 1.3, we shall in fact deduce Theorem 1.4 from a more general statement, Theorem 6.2, which is a version of the general transference theorem for stability results proved in [55]. Theorem 6.2 is stated and proved in Section 6; in Section 7, we use it to derive Theorem 1.4.

**1.3. The typical structure of  $H$ -free graphs.** Let  $H$  be an arbitrary non-empty graph. We say that a graph  $G$  is  $H$ -free if  $G$  does not contain  $H$  as a subgraph. For an integer  $n$ , denote by  $f_n(H)$  the number of labeled  $H$ -free graphs on the vertex set  $[n]$ . Since every subgraph of an  $H$ -free graph is also  $H$ -free, it follows that  $f_n(H) \geq 2^{\text{ex}(n, H)}$ . Erdős, Frankl, and Rödl [15] proved that this crude lower bound is in a sense tight, namely that

$$f_n(H) = 2^{\text{ex}(n, H) + o(n^2)}. \quad (4)$$

Our next result can be viewed as a ‘sparse version’ of (4). Such a statement was already considered by Łuczak [45], who derived it from the so-called KLR conjecture, which we discuss in the next subsection. For integers  $n$  and  $m$  with  $0 \leq m \leq \binom{n}{2}$ , let  $f_{n, m}(H)$  be the number of labeled  $H$ -free graphs on the vertex set  $[n]$  that have exactly  $m$  edges. The following theorem refines (4) to  $n$ -vertex graphs with  $m$  edges.

**Theorem 1.5.** *Let  $H$  be a 2-balanced graph and let  $\delta$  be a positive constant. There exists a constant  $C$  such that for every  $n \in \mathbb{N}$ , if  $m \geq Cn^{2-1/m_2(H)}$ , then*

$$\binom{\text{ex}(n, H)}{m} \leq f_{n, m}(H) \leq \binom{\text{ex}(n, H) + \delta n^2}{m}.$$

In fact, we shall deduce from our main result, Theorem 2.2, a ‘counting version’ of the general transference theorem of Schacht [58, Theorem 3.3], which easily implies Theorem 1.5. This ‘counting version’ of Schacht’s theorem is stated and proved in Section 5. We then use it to derive Theorem 1.5 in Section 8. We remark that (4) was refined in a different sense by Balogh, Bollobás, and Simonovits [5], who showed that  $f_n(H) = 2^{\text{ex}(n, H) + O(n^{2-c(H)})}$ , where  $c(H)$  is some positive constant, and also gave a very precise structural description of almost all  $H$ -free graphs. We would also like to point out that our proof of Theorem 1.5 does not use Szemerédi’s regularity lemma, unlike the proof given in [45] or the proofs of Erdős, Frankl, and Rödl [15] and Balogh, Bollobás, and Simonovits [5].

The result of Erdős, Frankl, and Rödl has, in some cases, a structural counterpart that significantly strengthens (4). For example, Erdős, Kleitman, and Rothschild [16] proved that

almost all triangle-free graphs are bipartite, that is, that with probability tending to 1 as  $n \rightarrow \infty$ , a graph selected uniformly at random from the family of all triangle-free graphs on the vertex set  $[n]$  is bipartite or, in other words (since clearly every bipartite graph is triangle-free),  $f_n(K_3)$  is asymptotic to the number of bipartite graphs on the vertex set  $[n]$ . Extending this result, Osthus, Prömel, and Taraz [48] proved that if  $m \geq Cn^{3/2}\sqrt{\log n}$  for some  $C > \sqrt{3}/4$ , then almost all  $n$ -vertex triangle-free graphs with  $m$  edges are bipartite. Our next result, which is a strengthening of Theorem 1.5, is an approximate version of this statement for an arbitrary 2-balanced graph  $H$ . Such a statement was also considered by Łuczak [45], who derived it from the KLR conjecture. Following [45], given a positive real  $\delta$  and an integer  $k$ , let us say that a graph  $G$  is  $(\delta, k)$ -partite if  $G$  can be made  $k$ -partite by removing from it at most  $\delta e(G)$  edges.

**Theorem 1.6.** *Let  $H$  be a 2-balanced graph with  $\chi(H) \geq 3$  and let  $\delta$  be a positive constant. There exists a constant  $C$  such that if  $m \geq Cn^{2-1/m_2(H)}$ , then almost all  $H$ -free graphs with  $n$  vertices and  $m$  edges are  $(\delta, \chi(H) - 1)$ -partite.*

Similarly as with Theorem 1.5, we shall in fact deduce Theorem 1.6 from a ‘counting version’ of the general transference theorem for stability results proved in [55]. Our version of it, Theorem 6.3, is stated and proved in Section 6. In Section 8, we use it to derive Theorem 1.6. Once again, our proof does not use the regularity lemma, unlike that in [45]. Finally, we would like to mention that, as observed by Łuczak [45], Theorem 1.6 has the following elegant corollary.

**Corollary 1.7.** *Let  $H$  be a 2-balanced graph with  $\chi(H) \geq 3$  and let  $\varepsilon$  be a positive constant. There exist constants  $C$  and  $n_0$  such that for every  $n$  with  $n \geq n_0$  and every  $m$  with  $Cn^{2-1/m_2(H)} \leq m \leq n^2/C$ ,*

$$\left(\frac{\chi(H) - 2}{\chi(H) - 1} - \varepsilon\right)^m \leq \mathbb{P}(G_{n,m} \not\supseteq H) \leq \left(\frac{\chi(H) - 2}{\chi(H) - 1} + \varepsilon\right)^m,$$

where  $G_{n,m}$  is a uniformly selected random  $n$ -vertex graph with  $m$  edges.

$\# H\text{-free} \geq \binom{n}{m} - k \binom{n}{m}^{1/2}$   
 $\# \leq k^n \binom{n}{m}$   
 $k = \chi(H)$

We remark that a great deal more is known about the structure of a typical  $H$ -free graph (drawn uniformly at random from the set of all  $n$ -vertex  $H$ -free graphs), see [6] and the references therein for more details.

**1.4. The KLR conjecture.** The celebrated Szemerédi’s regularity lemma [63], which is considered to be one of the most important and powerful tools in extremal graph theory, says that the vertex set of every graph may be divided into a bounded number of parts of approximately the same size in such a way that most of the bipartite subgraphs induced between pairs of parts of the partition satisfy a certain pseudo-randomness condition termed  $\varepsilon$ -regularity. The strength of the regularity lemma lies in the fact that it may be combined with the so-called *embedding lemma* to show that a graph contains particular subgraphs. The combination of the regularity and embedding lemmas allows one to prove many well-known theorems in extremal graph theory, such as the theorem of Erdős and Stone [18] and the stability theorem of Erdős and Simonovits [14, 60], both mentioned in Section 1.2.

For sparse graphs, that is,  $n$ -vertex graphs with  $o(n^2)$  edges, the original version of the regularity lemma is vacuous since if the vertex set of a sparse graph is partitioned into a

bounded number of parts, then all induced bipartite subgraphs thus obtained are trivially  $\varepsilon$ -regular, provided that  $n$  is sufficiently large. However, it was independently observed by Kohayakawa [38] and Rödl (unpublished) that the notion of  $\varepsilon$ -regularity may be extended in a meaningful way to graphs with density tending to zero. Moreover, with this more general notion of regularity, they were also able to prove an associated regularity lemma which applies to a large class of sparse graphs, including (a.a.s.) the random graph  $G(n, p)$ .

Given a  $p \in [0, 1]$  and a positive  $\varepsilon$ , we say that a bipartite graph between sets  $V_1$  and  $V_2$  is  $(\varepsilon, p)$ -regular if for every  $W_1 \subseteq V_1$  and  $W_2 \subseteq V_2$  with  $|W_1| \geq \varepsilon|V_1|$  and  $|W_2| \geq \varepsilon|V_2|$ , the density  $d(W_1, W_2)$  of edges between  $W_1$  and  $W_2$  satisfies

$$|d(W_1, W_2) - d(V_1, V_2)| \leq \varepsilon p.$$

A partition of the vertex set of a graph into  $r$  parts  $V_1, \dots, V_r$  is said to be  $(\varepsilon, p)$ -regular if  $||V_i| - |V_j|| \leq 1$  for all  $i$  and  $j$  and for all but at most  $\varepsilon r^2$  pairs  $(V_i, V_j)$ , the graph induced between  $V_i$  and  $V_j$  is  $(\varepsilon, p)$ -regular. The class of graphs to which the Kohayakawa-Rödl regularity lemma applies are the so-called upper-uniform graphs. Given positive  $\eta$  and  $K$ , we say that an  $n$ -vertex graph  $G$  is  $(\eta, p, K)$ -upper-uniform if for all  $W \subseteq V(G)$  with  $|W| \geq \eta n$ , the density of edges within  $W$  satisfies  $d(W) \leq Kp$ . This condition is satisfied by many natural classes of graphs, including all subgraphs of random graphs of density  $p$ . The sparse regularity lemma of Kohayakawa [38] and Rödl says the following.

**The sparse Szemerédi regularity lemma.** *For all positive  $\varepsilon$ ,  $K$ , and  $r_0$ , there exist a positive constant  $\eta$  and an integer  $R$  such that for every  $p \in [0, 1]$ , the following holds. Every  $(\varepsilon, p, K)$ -upper-uniform graph with at least  $r_0$  vertices admits an  $(\varepsilon, p)$ -regular partition of its vertex set into  $r$  parts, for some  $r \in \{r_0, \dots, R\}$ .*

We remark that a version of this theorem avoiding the need for the upper-uniformity assumption was recently proved by Scott [59].

The aforementioned embedding lemma roughly says that if we start with an arbitrary graph  $H$ , replace its vertices by large independent sets and its edges by  $\varepsilon$ -regular bipartite graphs with density bounded away from zero, then this blown-up graph will contain a copy of  $H$ . To make it more precise, let  $H$  be a graph on the vertex set  $\{1, \dots, v(H)\}$ , let  $\varepsilon$  and  $p$  be as above, and let  $n$  and  $m$  be integers satisfying  $0 \leq m \leq n^2$ . Let us denote by  $\mathcal{G}(H, n, m, p, \varepsilon)$  the collection of all graphs  $G$  constructed in the following way. The vertex set of  $G$  is a disjoint union  $V_1 \cup \dots \cup V_{v(H)}$  of sets of size  $n$ , one for each vertex of  $H$ . For each edge  $\{i, j\}$  of  $H$ , we add to  $G$  an  $(\varepsilon, p)$ -regular bipartite graph with  $m$  edges between the sets  $V_i$  and  $V_j$ . These are the only edges of  $G$ . With this notation in hand, we can state the embedding lemma. Given any graph  $G$  as above, we define *canonical copies of  $H$*  to be all copies of  $H$  in  $G$  in which (the image of) each vertex  $i \in V(H)$  lies in the set  $V_i \subseteq V(G)$ .

**The embedding lemma.** *For every graph  $H$  and every positive  $d$ , there exist a positive  $\varepsilon$  and an integer  $n_0$  such that for every  $n$  and  $m$  with  $n \geq n_0$  and  $m \geq dn^2$ , every  $G \in \mathcal{G}(H, n, m, 1, \varepsilon)$  contains a canonical copy of  $H$ .*

One might hope that a similar statement holds when one replaces 1 by an arbitrary  $p$  and the assumption  $m \geq dn^2$  by  $m \geq pdn^2$ , even if  $p$  is a decreasing function of  $n$ . However, for an arbitrary function  $p$ , this is too much to hope for. Indeed, consider the random ‘blow-up’ of  $H$ , that is, the random graph  $G$  obtained from  $H$  by replacing each vertex of  $H$  by an

independent set of size  $n$  and each edge of  $H$  by a random bipartite graph with  $pn^2$  edges. With high probability, the number of canonical copies of  $H$  in  $G$  will be about  $p^{e(H)}n^{v(H)}$  and hence if  $p^{e(H)}n^{v(H)} \ll pn^2$ , then one can remove all copies of  $H$  from  $G$  by deleting a tiny proportion of all edges. Since in the above argument one may replace  $H$  with an arbitrary subgraph  $H' \subseteq H$ , it follows easily that if  $p \leq cn^{-1/m_2(H)}$  for a sufficiently small positive constant  $c$ , then there are graphs in  $\mathcal{G}(H, n, pn^2, p, \varepsilon)$  that do not contain any canonical copies of  $H$ .

As in the case of Turán's theorem for random graphs, see Section 1.2, one might still hope that if  $p \geq Cn^{-1/m_2(H)}$  for some large constant  $C$ , then the natural sparse analogue of the embedding lemma discussed above holds. However, it was observed by Łuczak (see [30, 42]) that, somewhat surprisingly, for any graph  $H$  which contains a cycle and any function  $p$  satisfying  $p = o(1)$ , there are graphs in  $\mathcal{G}(H, n, pn^2, p, \varepsilon)$  with no canonical copy of  $H$ . Nevertheless, it still seemed likely that such atypical graphs comprise only a tiny proportion of  $\mathcal{G}(H, n, m, p, \varepsilon)$ . This was formalized in the following conjecture of Kohayakawa, Łuczak, and Rödl [41], usually referred to as the *KLR conjecture*. Given a graph  $H$ , integers  $m$  and  $n$ , a  $p \in [0, 1]$ , and a positive  $\varepsilon$ , let  $\mathcal{G}^*(H, n, m, p, \varepsilon)$  denote the collection of graphs in  $\mathcal{G}(H, n, m, p, \varepsilon)$  that contain no canonical copy of  $H$ .

**The KLR Conjecture.** *Let  $H$  be a fixed graph. Then, for any positive  $\beta$ , there exist positive  $C$ ,  $n_0$ , and  $\varepsilon$  such that for all  $n$  and  $m$  with  $n \geq n_0$  and  $m \geq Cn^{2-1/m_2(H)}$ ,*

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

The KLR conjecture has been one of the central open questions in extremal graph theory and has attracted substantial attention of many researchers over the past fifteen years. It has been verified in several special cases. It is easy to see that it holds for all graphs  $H$  which do not contain a cycle. The cases  $H = K_3$ ,  $K_4$ , and  $K_5$  were resolved in [40], [28], and [29], respectively. The case  $H = C_\ell$  has also been resolved, but here the history is somewhat more complex. A proof under some extra technical assumptions was given in [39]. Those extra assumptions were later removed in [27] and, independently, in [11]. We remark here that in parallel to this work, Conlon, Gowers, Samotij, and Schacht [13] have proved a sparse analogue of the *counting lemma* for subgraphs of the random graph  $G(n, p)$ , which may be viewed as a version of the KLR conjecture that is stronger in some aspects and weaker in other aspects. Our next result is a proof of the KLR conjecture for all 2-balanced graphs.

**Theorem 1.8.** *Let  $H$  be a 2-balanced graph. Then, for any positive  $\beta$ , there exist positive  $C$ ,  $n_0$ , and  $\varepsilon$  such that for all  $n$  and  $m$  with  $n \geq n_0$  and  $m \geq Cn^{2-1/m_2(H)}$ ,*

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

It is well-known that Theorem 1.8 easily implies Turán's theorem for random graphs, Theorem 1.3, and also its stability version, Theorem 1.4. In fact, this was the original motivation behind the KLR conjecture, see [41]. Moreover, it was proved by Łuczak [45] that Theorem 1.8 implies Theorems 1.5 and 1.6. The work of Conlon and Gowers [12] and Schacht [58] (see also [55]), as well as this work, have shown that one does not need to



appeal to the sparse regularity lemma and to the KLR conjecture in order to prove such extremal statements in random graphs. Nevertheless, there is still a plentitude of beautiful corollaries of the conjecture that cannot (yet) be proved by other means. For discussion and derivation of some of them, we refer the reader to [13]. Here, we present only one corollary of the KLR conjecture, the threshold for asymmetric Ramsey properties of random graphs, which does not follow from the version of the conjecture proved in [13]. The deduction of this result from the KLR conjecture is essentially due to Kohayakawa and Kreuter [39]. We prove Theorem 1.8 in Section 9.

**1.5. Ramsey properties of random graphs.** Let  $H$  be a fixed graph and let  $r$  be a positive integer. For an arbitrary graph  $G$ , we write  $G \rightarrow (H)_r$  if every  $r$ -coloring of the edges of  $G$  contains a monochromatic copy of  $H$ . It follows from the classical result of Ramsey [50] that  $K_n \rightarrow (H)_r$ , provided that  $n$  is sufficiently large. Ramsey properties of random graphs were first investigated by Frankl and Rödl [20] and since then much effort has been devoted to their study. Most notably, Rödl and Ruciński [51, 52] established the following general threshold result.

**Theorem 1.9.** *Let  $r$  be a positive integer and let  $H$  be an arbitrary fixed graph that is not a forest. There exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G(n, p_n) \rightarrow (H)_r) = \begin{cases} 1 & \text{if } p_n \geq Cn^{-1/m_2(H)}, \\ 0 & \text{if } p_n \leq cn^{-1/m_2(H)}. \end{cases}$$

In the above discussion, a copy of the same graph  $H$  is forbidden in each of the  $r$  color classes. A natural generalization of Theorem 1.9 would determine thresholds for so-called *asymmetric Ramsey properties*. For any graphs  $G, H_1, \dots, H_r$ , we write  $G \rightarrow (H_1, \dots, H_r)$  if for every coloring of the edges of  $G$  with colors  $1, \dots, r$ , there exists, for some  $i \in [r]$ , a copy of  $H_i$  all of whose edges have color  $i$ . In the context of asymmetric Ramsey properties of random graphs, the following generalization of the 2-density  $m_2(\cdot)$  was introduced in [39]. For two graphs  $H_1$  and  $H_2$ , define

$$m_2(H_2, H_1) = \max \left\{ \frac{e(H'_1)}{v(H'_1) - 2 + 1/m_2(H_2)} : H'_1 \subseteq H_1 \text{ with } v(H'_1) \geq 3 \right\}. \quad (5)$$

Kohayakawa and Kreuter [39] formulated the following conjecture and proved it in the case when all  $H_i$  are cycles.

**Conjecture 1.10.** *Let  $H_1, \dots, H_r$  be graphs with  $1 < m_2(H_r) \leq \dots \leq m_2(H_1)$ . Then there exist constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G(n, p_n) \rightarrow (H_1, \dots, H_r)) = \begin{cases} 1 & \text{if } p_n \geq Cn^{-1/m_2(H_1, H_2)}, \\ 0 & \text{if } p_n \leq cn^{-1/m_2(H_1, H_2)}. \end{cases}$$

More accurately, the above conjecture was stated in [39] only in the case  $r = 2$ , but the above generalization is quite natural, see [46]. There had been little progress on Conjecture 1.10 until quite recently, when the 0-statement was proved by Marciniszyn, Skokan, Spöhel, and Steger [46] in the case where all of the  $H_i$  are cliques, and the 1-statement in the case  $r = 2$  was established by Kohayakawa, Schacht, and Spöhel [44] under very mild extra assumptions on  $H_1$  and  $H_2$ . Using Theorem 1.8, the approach of Kohayakawa and

Kreuter [39], which employs the sparse regularity lemma, can be adapted (see [46, Theorem 31]) to yield a proof of the 1-statement in Conjecture 1.10 for the following class of graphs.

**Theorem 1.11.** *Let  $H_1, \dots, H_r$  be graphs such that  $H_1$  is strictly 2-balanced,  $H_2, \dots, H_r$  are 2-balanced, and  $1 < m_2(H_r) \leq \dots \leq m_2(H_1)$ . There exists a constant  $C$  such that if  $p_n \geq Cn^{-1/m_2(H_2, H_1)}$ , then a.a.s.*

$$G(n, p_n) \rightarrow (H_1, \dots, H_r).$$

For the deduction of Theorem 1.11 from Theorem 1.8, see [39] and [46, Section 4]. ← Check

**1.6. Outline of the paper.** The remainder of this paper is organized as follows. In Section 2, we state and discuss our main result, Theorem 2.2, which we then prove in Section 3. In Section 4, we discuss the applications of Theorem 2.2 in the context of subsets of  $[n]$  with no  $k$ -term arithmetic progressions. In particular, we prove Theorem 1.1 and use it to derive Corollary 1.2. In Section 5, we prove two versions of the general transference theorem of Schacht [58, Theorem 3.3] (obtained independently, in a slightly different form, by Conlon and Gowers [12]) – a ‘random’ version suited for extremal problems in sparse random discrete structures and its ‘counting’ counterpart that generalizes Theorem 1.1. In Section 6, we prove ‘random’ and ‘counting’ versions of the general stability result of Conlon and Gowers [12] in a form that is easily comparable with [55, Theorem 3.4]. In Section 7, we discuss several applications of Theorem 2.2 in the context of the Turán problem in sparse random graphs. In particular, using the results of Sections 5 and 6 we give new proofs of the sparse random analogues (stated above) of the classical theorems of Erdős and Stone, and Erdős and Simonovits, see Section 1.2. In Section 8, we discuss applications of Theorem 2.2 to the problem of describing the typical structure of a sparse graph without a forbidden subgraph. In particular, we prove sparse analogues of classical theorems of Erdős, Frankl, and Rödl and Erdős, Kleitman, and Rothschild, see Section 1.3. Finally, in Section 9, we use Theorem 2.2 to prove the KLR conjecture for all 2-balanced graphs.

## 2. THE MAIN THEOREM

In this section, we present the main result of this paper, Theorem 2.2, which gives a structural characterization of the collection of all independent sets in a large class of uniform hypergraphs. We start with an important definition. Recall that a family of sets  $\mathcal{F} \subseteq \mathcal{P}(V)$  is called *increasing* (or an *upset*) if it is closed under taking supersets, that is, if for every  $A, B \subseteq V$ ,  $A \in \mathcal{F}$  and  $A \subseteq B$  imply that  $B \in \mathcal{F}$ .

**Definition 2.1.** Let  $\mathcal{H}$  be a uniform hypergraph with vertex set  $V$ , let  $\mathcal{F}$  be an increasing family of subsets of  $V$  and let  $\varepsilon \in (0, 1]$ . We say that  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense if

$$e(\mathcal{H}[A]) \geq \varepsilon e(\mathcal{H})$$

for every  $A \in \mathcal{F}$ .

A moment of thought reveals that for an arbitrary hypergraph  $\mathcal{H}$  and  $\varepsilon \in (0, 1]$ , it is extremely simple to construct families  $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$  for which  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense. To this end, let

$$\mathcal{F}_\varepsilon = \{A \subseteq V(\mathcal{H}) : e(\mathcal{H}[A]) \geq \varepsilon e(\mathcal{H})\}$$

## Triangle Removal Lemma

If  $G$  is  $\epsilon$ -far from being  $\Delta$ -free then  $G \supseteq \geq \delta n^3 \Delta$ 's.

$$\delta = \delta(\epsilon)$$

## Stability Theorem

Let  $X(F) = r+1 \geq 3$ .  $G$  is  $H$ -free. For every  $\epsilon, \exists \delta$  s.t.,  
and  $e(G) \geq \frac{r-1}{2r} n^2 - \delta n^2$  then  $\exists n_1 + n_2 + \dots + n_r$   
s.t.  $G$  differs by  $\leq \epsilon n^2$  edges from  $K_r(n_1, \dots, n_r)$ .

$G \in \mathcal{F}_2 \Rightarrow G$  more than  $\epsilon n^2$  from a bipartite graph

$\Rightarrow$  more than  $\epsilon n^2$  from  $\Delta$ -free (S.T.)

$\Rightarrow$  has  $\geq \delta n^3 \Delta$ 's (T.R.L.)

and note that  $\mathcal{F}_\varepsilon$  is increasing and  $\mathcal{H}$  is  $(\mathcal{F}_\varepsilon, \varepsilon)$ -dense. In fact, the families  $\mathcal{F}$  for which  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense are precisely all increasing subfamilies of  $\mathcal{F}_\varepsilon$ .

In this work, we will be interested in upsets that admit a much more ‘constructive’ and simpler description than that of  $\mathcal{F}_\varepsilon$ . Many such families arise naturally in the study of extremal and structural problems in combinatorics. For example, consider the  $k$ -uniform hypergraph  $\mathcal{H}_1$  on the vertex set  $[n]$  whose edges are all  $k$ -term arithmetic progressions in  $[n]$  and let  $\mathcal{F}_1$  be the collection of all subsets of  $[n]$  with at least  $\delta n$  elements. Clearly,  $\mathcal{F}_1$  is an upset and it follows from the famous theorem of Szemerédi [62] that  $\mathcal{H}_1$  is  $(\mathcal{F}_1, \varepsilon)$ -dense for some positive  $\varepsilon$  depending only on  $\delta$  and  $k$ , see Section 4. Similarly, consider the 3-uniform hypergraph  $\mathcal{H}_2$  on the vertex set  $E(K_n)$  whose edges are edge sets of all copies of  $K_3$  in the complete graph  $K_n$  and let  $\mathcal{F}_2$  be the family of all  $n$ -vertex graphs (subgraphs of  $K_n$ ) with at least  $(1/2 - \varepsilon)\binom{n}{2}$  edges such that every 2-coloring of its vertices yields at least  $\delta n^2$  monochromatic edges. Again,  $\mathcal{F}_2$  is increasing and it follows from the stability theorem of Erdős and Simonovits [14, 60] and the triangle removal lemma of Ruzsa and Szemerédi [54] that  $\mathcal{H}_2$  is  $(\mathcal{F}_2, \varepsilon)$ -dense, provided that  $\varepsilon$  is sufficiently small as a function of  $\delta$ .

Our main result roughly says the following. If  $\mathcal{H}$  is a uniform hypergraph that is  $(\mathcal{F}, \varepsilon)$ -dense for some family  $\mathcal{F}$  and whose edge distribution satisfies certain natural boundedness conditions, then the collection  $\mathcal{I}(\mathcal{H})$  of all independent sets in  $\mathcal{H}$  admits a partition into relatively few classes such that all independent sets in one class are essentially contained in a single set  $A \notin \mathcal{F}$ . Before we state the result, we first need to quantify the above boundedness condition for the edge distribution of a hypergraph. Given a hypergraph  $\mathcal{H}$ , for each  $T \subseteq V(\mathcal{H})$ , we define

$$\deg_{\mathcal{H}}(T) = |\{e \in \mathcal{H} : T \subseteq e\}|,$$

and let

$$\Delta_\ell(\mathcal{H}) = \max \{ \deg_{\mathcal{H}}(T) : T \subseteq V(\mathcal{H}) \text{ and } |T| = \ell \}.$$

Recall that  $\mathcal{I}(\mathcal{H})$  denotes the family of all independent sets in  $\mathcal{H}$ . The following theorem is our main result.

**Theorem 2.2.** *For every  $k \in \mathbb{N}$  and all positive  $c, c'$  and  $\varepsilon$ , there exists a positive constant  $C$  such that the following holds. Let  $\mathcal{H}$  be a  $k$ -uniform hypergraph and let  $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$  be an increasing family of sets such that  $|A| \geq \varepsilon v(\mathcal{H})$  for all  $A \in \mathcal{F}$ . Suppose that  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense and  $p \in (0, 1)$  is such that  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$  and for every  $\ell \in [k-1]$ ,*

$$\Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\}.$$

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

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

Roughly speaking, if  $\mathcal{H}$  satisfies certain technical conditions, then each independent set  $I$  in  $\mathcal{H}$  can be labeled with a small subset  $g(I)$  in such a way that all sets labeled with some  $S \in \mathcal{S}$  are essentially contained in a single set  $f(S)$  that contains very few edges of  $\mathcal{H}$ . We remark that the constant  $C$  in the theorem has only a polynomial dependence on  $\varepsilon$ .

## L12 of [39]

Let graphs  $H_1, \dots, H_r$  be given.  
 $\exists c > 0$  and  $n_0$  s.t. if  $n \geq n_0$   
and  $K_n$  is  $r$ -colored then  
 $\exists 1 \leq i \leq r$  for which there are  
 $\geq c_1 n^{v(H_i)}$   $i$ -colored copies  
of  $H_i$ .

## L16 of [39]

Let  $H = (\{w_1, \dots, w_h\}, E)$  be  
a fixed graph. Let  $W_1, \dots, W_h$   
be pair-wise disjoint with  
 $|W_i| \geq \frac{n}{10 \log n}$ . Let  $\chi_W(G_p)$   
= # embeddings  $\phi$  of  $H$  in  
 $G_p$  where each  $\phi(w_i) \in W_i$

and if  $(w_i, w_j) \in E(H)$  then  
 $(\phi(w_i), \phi(w_j)) \in E(G_p)$ .

Then whp,  $\forall W_1, \dots, W_h$   
 we have

$$\chi_W \cong c_{H^p}^{e(H)} |W_1| \cdots |W_h|.$$

$$m_2(H_1) \leq \dots \leq m_2(H_r)$$

$H = H_r$ :  $\widetilde{H}(G)$  is spanning  
 subgraph of  $G$  with  
 edges from isolated  
 copies of  $H$

$H$  is isolated if it does  
 not share an edge with  
 another copy.

$$p_0 = \Pr(e \in \text{isolated copy of } H)$$

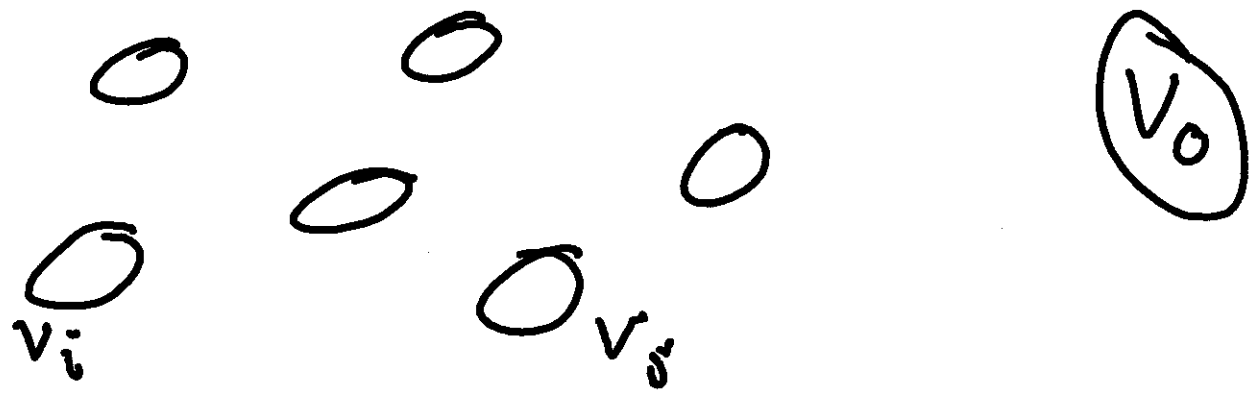
$$\approx c_2 n^{v-2} p^e = c_2 n^{-1/m_2}$$

$r$  color edges of  $G_p$ .

Assume there is no  $r$ -colored copy of  $H = H_r$ .

Define  $J_1, \dots, J_{r-1}$  by selecting one edge of color  $\neq r$  from each isolated copy of  $H$ . If edge has color  $i$ , then put it into  $J_i$ .

Using sparse regularity simultaneously partition each  $J_i$  into  $V_0, V_1, \dots, V_k$  for some small  $\epsilon$ .



$K_k$ : color  $(i,j)$  with  $r+1$  if  $(V_i, V_j)$  not  $\epsilon$ -regular for some  $b$ .

color  $(i,j)$  with  $1 \leq s < r$  if  $(V_i, V_j)$  has density  $\geq \gamma_0$  in some  $J_s$ .

Otherwise color  $(i,j)$  with 0

Possibilities:

A:  $\exists$   $i$  colored copy of  $H_i$  in

$K_k$ .  $K \neq R$  implies that whp this gets size  $\geq i$ -colored  $H_i$ .

B:  $\leq \epsilon \binom{k}{2}$  edges of color  $r+1$

$\Rightarrow \geq c_1 k^{v(H)}$  0-colored



# of edges in  $G$  corresponding to

this is

$$\geq \frac{1}{2} c_1 k^v \left(\frac{n}{k}\right)^v p^{e(H)}$$

$$= \frac{1}{2} c_1 n^{2-1/m_2}$$

On the other hand, # edges

$$\leq 2\gamma_0 \binom{k}{2} \left(\frac{n}{k}\right)^2 p_0$$

$$\leq \gamma_0 n^{2-1/m_2}.$$

Contradiction for  $\gamma_0$  small.

Unfortunately, however, in most of our applications  $\varepsilon$  will have a tower-type dependence on some other parameter.

Theorem 2.2 will be proved in Section 3. We end this section with a short informal discussion of its consequences. As we have already mentioned, Theorem 2.2 combined with some classical extremal results on discrete structures has strikingly strong implications. Let us briefly explain why this is so. Many classical extremal problems ask for an estimate on the number of independent sets (of certain size) in some auxiliary uniform hypergraph. If applicable, Theorem 2.2 implies that all such independent sets are almost contained in one of very few sets that are almost independent, that is, contain a small number of copies of some forbidden substructure. If we know a good characterization of sets that are almost independent in the above sense, which is often the case, we can easily obtain an upper bound on the number of independent sets. For example, consider the problem of counting subsets of  $[n]$  with no  $k$ -term AP and recall the definition of  $\mathcal{H}_1$  and  $\mathcal{F}_1$  from the beginning of this section. Theorem 2.2, applied to this pair, implies that the every subset of  $[n]$  with no  $k$ -term AP is essentially contained in one of at most  $\binom{n}{O(n^{1-1/(k-1)})}$  sets of size at most  $\delta n$  each, where  $\delta$  is an arbitrarily small positive constant. This easily implies that if  $m \gg n^{1-1/(k-1)}$ , then there are at most  $\binom{2\delta n}{m}$  sets of size  $m$  with no  $k$ -term AP. For details, we refer the reader to Section 4.

### 3. PROOF OF THE MAIN THEOREM

In this section, we shall prove Theorem 2.2. The main ingredient in the proof is the following proposition, which (roughly) says that Theorem 2.2 holds in the special case when  $\mathcal{F}$  is the family of all subsets of  $V(\mathcal{H})$  with at least  $(1 - \delta)v(\mathcal{H})$  elements. Theorem 2.2 follows by applying Proposition 3.1 a constant number of times.

**Proposition 3.1.** *For every integer  $k$  and positive  $c$  and  $c'$ , there exists a positive  $\delta$  such that the following holds. Let  $p \in (0, 1)$  and suppose that  $\mathcal{H}$  is a  $k$ -uniform hypergraph such that  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$  and for every  $\ell \in [k - 1]$ ,*

$$\Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\} \geq c' p^{\ell-k}$$

Then there exist a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leq (k-1)p v(\mathcal{H})}$  and functions  $f_0: \mathcal{S} \rightarrow \mathcal{P}(V(\mathcal{H}))$  and  $g_0: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{S}$  such that for every  $I \in \mathcal{I}(\mathcal{H})$ ,

$$g_0(I) \subseteq I \subseteq f_0(g_0(I)) \cup g_0(I) \quad \text{and} \quad |f_0(g_0(I))| \leq (1 - \delta)v(\mathcal{H}).$$

Moreover, if for some  $I, I' \in \mathcal{I}(\mathcal{H})$ ,  $g_0(I) \subseteq I'$  and  $g_0(I') \subseteq I$ , then  $g_0(I) = g_0(I')$ .

The final line of Proposition 3.1 states that the labeling function  $g_0$  exhibits a certain consistency. This property of  $g_0$ , which may look somewhat puzzling, will be crucial in the proof of Theorem 2.2.

In order to prove Proposition 3.1, given an independent set  $I \in \mathcal{I}(\mathcal{H})$ , we shall construct a sequence  $(B_{k-1}, \dots, B_q)$  of subsets of  $I$  with  $|B_{k-1}|, \dots, |B_q| \leq pv(\mathcal{H})$ , for some  $q \in [k - 1]$ , and use it to define a sequence  $(\mathcal{H}_{k-1}, \dots, \mathcal{H}_r)$ , where  $r \in \{q, q + 1\}$ , of hypergraphs such that the following holds for each  $i \in \{r, \dots, k - 1\}$ :

(a)  $\mathcal{H}_i$  is an  $i$ -uniform hypergraph on the vertex set  $V(\mathcal{H})$ ,

- (b)  $I$  is an independent set in  $\mathcal{H}_i$ , and
- (c)  $e(\mathcal{H}_i) \geq \Omega(p^{k-i}e(\mathcal{H}))$ .

We shall be able to do it in such a way that in the end, there will be a set  $A \subseteq V(\mathcal{H})$  of size at most  $(1 - \delta)v(\mathcal{H})$  such that the remaining elements of  $I$ , the set  $I \setminus S$ , must all lie inside  $A$ . If  $r = 1$ , then we will simply let  $A$  be the set of non-edges of the 1-uniform hypergraph  $\mathcal{H}_1$ ; in this case, the upper bound on  $|A|$  will follow from (c) and our assumption that  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$ . If  $r > 1$ , then we will obtain an appropriate  $A$  while trying (and failing) to construct the hypergraph  $\mathcal{H}_{r-1}$  using the hypergraph  $\mathcal{H}_r$  and the set  $B_r$ . Crucially, this set  $A$  will depend solely on  $S$ , that is, if for some pair  $I, I' \in \mathcal{I}(\mathcal{H})$  our procedure generates  $(S, A)$  and  $(S', A')$ , respectively, and  $S = S'$ , then also  $A = A'$ . This will allow us to set  $g_0(I) = S$  and  $f_0(S) = A$ .

**3.1. The Algorithm Method.** For the remainder of this section, let us fix  $k, c, c', p$  and  $\mathcal{H}$  as in the statement of Proposition 3.1. Without loss of generality, we may assume that  $c \geq 1$ . Let  $I$  be an independent set in  $\mathcal{H}$ . We shall describe a procedure of choosing the sets  $B_i \subseteq I$  and constructing the hypergraphs  $\mathcal{H}_i$  as above. This procedure, which we shall term the *Scythe Algorithm*, lies at the heart of the proof of Proposition 3.1.

The general strategy used in the Scythe Algorithm, that of selecting a small set  $S$  of high-degree vertices and using it to define a set  $A$  such that  $S \subseteq I \subseteq A \cup S$ , dates back to the work of Kleitman and Winston [37], who used it to bound the number of independent sets in graphs satisfying the following local density condition: all sufficiently large vertex sets induce subgraphs with many edges. Recently, Balogh and Samotij [9, 10] refined the ideas of Kleitman and Winston and obtained a bound on the number of independent sets in uniform hypergraphs satisfying a similar local density condition. Even more recently, Alon, Balogh, Morris and Samotij [1] used similar ideas to bound the number of independent sets in ‘almost linear’ 3-uniform hypergraphs satisfying a more general density condition termed  $(\alpha, \mathcal{B})$ -stability, see Definition 6.1. Here, we combine, generalize, and refine all of the above approaches and make them work in the general setting of  $(\mathcal{F}, \varepsilon)$ -dense uniform hypergraphs.

At each step of the Scythe Algorithm, we shall order the vertices of a certain subhypergraph of  $\mathcal{H}$  with respect to their degrees in that subhypergraph. For the sake of brevity and clarity of the presentation, let us make the following definition.

**Definition 3.2** (Max-degree order). Given a hypergraph  $\mathcal{G}$ , we define the *max-degree order* on  $V(\mathcal{G})$  as follows:

- (1) Fix an arbitrary total ordering of  $V(\mathcal{G})$ .
- (2) For each  $j \in \{1, \dots, v(\mathcal{G})\}$ , let  $u_j$  be the maximum-degree vertex in the hypergraph  $\mathcal{G}[V(\mathcal{G}) \setminus \{u_1, \dots, u_{j-1}\}]$ ; ties are broken by giving preference to vertices which come earlier in the order chosen in (1).
- (3) The max-degree order on  $V(\mathcal{G})$  is  $(u_1, \dots, u_{v(\mathcal{G})})$ .

Finally, we write  $W(u)$  to denote the initial segment of the max-degree order on  $V(\mathcal{G})$  that ends with  $u$ , i.e., for every  $j$ , we let  $W(u_j) = \{u_1, \dots, u_j\}$ .

We remark here that the only property of the max-degree order that will be important for us is that for every  $j \in \{1, \dots, v(\mathcal{G})\}$ , the degree of the vertex  $u_j$  in the hypergraph  $\mathcal{G}[V(\mathcal{G}) \setminus W(u_{j-1})]$  is at least as large as the average degree of this hypergraph.

We next define the numbers  $\Delta_\ell^i$ , where  $1 \leq \ell < i \leq k$ , which will play a crucial role in the description and the analysis of the algorithm.

**Definition 3.3.** For every  $\ell \in [k-1]$ , let  $\Delta_\ell^k = \Delta_\ell(\mathcal{H})$  and for all  $i \in [k-1]$  and  $\ell \in [i-1]$ , let

$$\Delta_\ell^i = \max \{2 \cdot \Delta_{\ell+1}^{i+1}, p \cdot \Delta_\ell^{i+1}\}. \quad (6)$$

We use the numbers  $\Delta_\ell^i$  to define the following families of sets with high degree.

**Definition 3.4.** Given an  $i \in [k]$ , an  $i$ -uniform hypergraph  $\mathcal{G}$  and an  $\ell \in [i-1]$ , let

$$M_\ell^i(\mathcal{G}) = \left\{ T \in \binom{V(\mathcal{G})}{\ell} : \deg_{\mathcal{G}}(T) \geq \frac{\Delta_\ell^i}{2} \right\}.$$

Moreover, let  $M_i^i(\mathcal{G}) = E(\mathcal{G})$ .

Let  $b = pv(\mathcal{H})$  and for each  $i \in [k]$ , let  $c_i = (ck2^{k+1})^{i-k}$ .

**Properties.** The key properties that we would like the constructed hypergraph  $\mathcal{H}_i$  to possess are:

- (P1)  $\mathcal{H}_i$  is  $i$ -uniform and  $V(\mathcal{H}_i) = V(\mathcal{H})$ ,
- (P2)  $I$  is an independent set in  $\mathcal{H}_i$ ,
- (P3)  $\Delta_\ell(\mathcal{H}_i) \leq \Delta_\ell^i$  for each  $\ell \in [i-1]$ ,
- (P4)  $e(\mathcal{H}_i) \geq c_i p^{k-i} e(\mathcal{H})$ .

Set  $\mathcal{H}_k = \mathcal{H}$  and note that (P1)–(P4) are vacuously satisfied for  $i = k$ . The main step of the Scythe Algorithm will be a procedure that, given  $\mathcal{H}_{i+1}$  and  $I$  satisfying (P1)–(P4), outputs a set  $B_i \subseteq I$  of cardinality  $b$ , a set  $A_i \subseteq V(\mathcal{H})$  with the property that  $I \setminus B_i \subseteq A_i$ , and a hypergraph  $\mathcal{H}_i$  satisfying (P1)–(P3). Moreover, if the constructed  $\mathcal{H}_i$  does not satisfy (P4), then we have  $|A_i| \leq (1 - c_i)v(\mathcal{H})$ . Crucially, these  $A_i$  and  $\mathcal{H}_i$  depend solely on  $B_i$  and  $\mathcal{H}_{i+1}$ , that is, if on two inputs  $(\mathcal{H}_{i+1}, I)$  and  $(\mathcal{H}_{i+1}, I')$ , the procedure outputs the same set  $B_i$ , it also outputs the same  $A_i$  and  $\mathcal{H}_i$ .

**The Scythe Algorithm.** Given an  $(i+1)$ -uniform hypergraph  $\mathcal{H}_{i+1}$  and an independent set  $I \in \mathcal{I}(\mathcal{H}_{i+1})$ , set  $\mathcal{A}_{i+1}^{(0)} = \mathcal{H}_{i+1}$  and let  $\mathcal{H}_i^{(0)}$  be the empty hypergraph on the vertex set  $V(\mathcal{H})$ . For  $j = 0, \dots, b-1$ , do the following:

- (1) If  $I \cap V(\mathcal{A}_{i+1}^{(j)}) = \emptyset$ , then set  $\mathcal{H}_i = \mathcal{H}_i^{(0)}$ ,  $A_i = \emptyset$ , and  $B_i = \{u_0, \dots, u_{j-1}\}$  and STOP.
- (2) Let  $u_j$  be the first vertex of  $I$  in the max-degree order on  $V(\mathcal{A}_{i+1}^{(j)})$ .
- (3) Let  $\mathcal{H}_i^{(j+1)}$  be the hypergraph on the vertex set  $V(\mathcal{H})$  defined by:

$$\mathcal{H}_i^{(j+1)} = \mathcal{H}_i^{(j)} \cup \left\{ D \in \binom{V(\mathcal{H})}{i} : D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j)} \right\}.$$

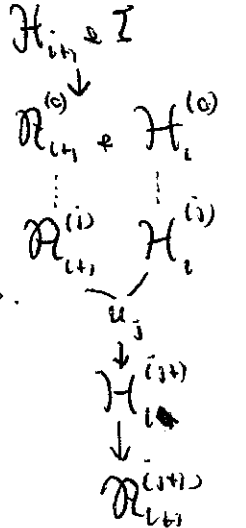
- (4) Let  $\mathcal{A}_{i+1}^{(j+1)}$  be the hypergraph on the vertex set  $\underbrace{V(\mathcal{A}_{i+1}^{(j)}) \setminus W(u_j)}_{\neq}$  defined by:

$$\mathcal{A}_{i+1}^{(j+1)} = \left\{ D \in \mathcal{A}_{i+1}^{(j)} : D \cap W(u_j) = \emptyset \text{ and } T \not\subseteq D \text{ for every } T \in \bigcup_{\ell=1}^i M_\ell^i(\mathcal{H}_i^{(j+1)}) \right\}.$$

Finally, set  $\mathcal{H}_i = \mathcal{H}_i^{(b)}$ ,  $A_i = V(\mathcal{A}_{i+1}^{(b)})$ , and  $B_i = \{u_0, \dots, u_{b-1}\} \subseteq I$ , by (2).

\* Only <sup>new</sup> members of  $I$  deleted is  $u_j$

$\mathcal{H}_k = \mathcal{H}$



We shall now establish various properties of the Scythe Algorithm. We begin by making some basic (but key) observations.

**Lemma 3.5.** *The following hold for every  $i \in [k - 1]$ :*

- (a)  $\mathcal{H}_i$  is  $i$ -uniform and  $V(\mathcal{H}_i) = V(\mathcal{H})$ .
- (b) If  $I \in \mathcal{I}(\mathcal{H}_{i+1})$ , then  $I \in \mathcal{I}(\mathcal{H}_i)$ .
- (c)  $B_i \subseteq I \subseteq A_i \cup B_i$ .
- (d) The hypergraph  $\mathcal{H}_i$  and the set  $A_i$  depend only on  $\mathcal{H}_{i+1}$  and the set  $B_i$ .

*Proof.* Property (a) is trivial. To see (b), simply observe that each edge of  $\mathcal{H}_i$  is of the form  $D \setminus \{u\}$  for some  $D \in \mathcal{H}_{i+1}$  and  $u \in I$ . Thus, if  $I$  contains an edge of  $\mathcal{H}_i$ , it must also contain an edge of  $\mathcal{H}_{i+1}$ . To see (c), observe that for each  $j$ ,  $u_j$  is the first vertex of  $I$  in the max-degree order on  $V(\mathcal{A}_{i+1}^{(j)})$  and hence  $W(u_j) \cap I = \{u_j\}$ . It follows that  $B_i \subseteq I$  and that  $I \setminus A_i = B_i$ . Note in particular that if  $A_i = \emptyset$ , then  $I \cap V(\mathcal{A}_{i+1}^{(j)}) = \emptyset$  for some  $j \in \{0, \dots, b\}$ , which implies that  $B_i = I$ . Finally, to prove (d), observe that all steps of the Scythe Algorithm are deterministic and that every element of  $I$  that we need to observe in order to define  $A_i$  and  $\mathcal{H}_i$  is placed in  $B_i$ . More precisely, note that while choosing the vertex  $u_j$ , we only need to know the first vertex of  $I$  in the max-degree order on  $V(\mathcal{A}_{i+1}^{(j)})$ ; the remaining vertices remain unobserved. Since we have  $W(u_j) \cap B_i = W(u_j) \cap I = \{u_j\}$ , this information can be recovered from  $B_i$ . Thus, at each step, the hypergraph  $\mathcal{H}_i^{(j+1)}$  can be recovered from  $\mathcal{H}_i^{(j)}$  and  $B_i$ , and the hypergraph  $\mathcal{A}_{i+1}^{(j+1)}$  can be recovered from  $\mathcal{A}_{i+1}^{(j)}$ ,  $\mathcal{H}_i^{(j+1)}$  and  $B_i$ . Hence, a trivial inductive argument proves that, if the algorithm does not stop in step (1), for each  $j \in \{0, \dots, b\}$ , the hypergraphs  $\mathcal{H}_i^{(j)}$  and  $\mathcal{A}_{i+1}^{(j)}$  are determined by  $\mathcal{H}_{i+1}$  and the set  $B_i$ , as required. Finally, the algorithm stops in step (1) if and only if  $|B_i| < b$ . If this happens, then  $\mathcal{H}_i$  and  $A_i$  are empty.  $\square$

We next show that the Scythe Algorithm exhibits a certain ‘consistency’ while generating its output. This property will be very important in the proof of Proposition 3.1.

**Lemma 3.6.** *Suppose that on inputs  $(\mathcal{H}_{i+1}, I)$  and  $(\mathcal{H}_{i+1}, I')$ , the Scythe Algorithm outputs  $(A_i, B_i, \mathcal{H}_i)$  and  $(A'_i, B'_i, \mathcal{H}'_i)$ , respectively. If  $B_i \subseteq I'$  and  $B'_i \subseteq I$ , then  $(A_i, B_i, \mathcal{H}_i) = (A'_i, B'_i, \mathcal{H}'_i)$ .*

*Proof.* By Lemma 3.5, it suffices to show that  $B_i = B'_i$ . Suppose that  $B_i \neq B'_i$ . Let us first consider the (degenerate) case when  $\min\{|B_i|, |B'_i|\} < b$ . Without loss of generality, we may assume that  $|B_i| < b$ . This means that, while running on  $(\mathcal{H}_{i+1}, I)$ , the Scythe Algorithm stopped in step (1). By Lemma 3.5, it follows that  $B_i = I$  and hence  $B'_i \subseteq B_i$ , which means that  $|B'_i| < b$  and therefore  $B'_i = I'$ . Hence,  $B_i = B'_i$ , as claimed. On the other hand, if  $|B_i| = |B'_i| = b$ , then there must exist some  $j$  such that  $u_j \neq u'_j$ . Let  $j$  be the smallest such index. Note that by the minimality of  $j$ , we have  $\mathcal{A}_{i+1}^{(j)} = (\mathcal{A}_{i+1}^{(j)})' = \mathcal{A}$ . Since  $u_j \neq u'_j$ , one of these vertices comes earlier in the max-degree order on  $V(\mathcal{A})$ ; without loss of generality, we may suppose that it is  $u_j$ . Since  $B_i \subseteq I'$ , it follows that  $u_j \in I'$  and hence the Algorithm, while running on the input  $(\mathcal{H}_{i+1}, I')$ , would not pick  $u'_j$  in step  $j$ , a contradiction. This shows that in fact  $B_i = B'_i$ , as required.  $\square$

The next lemma motivates the definition of  $M_i^i(\mathcal{G})$ ; it will be an important tool in the proof of Lemma 3.10, below.

**Lemma 3.7.** *For every  $D \in \mathcal{H}_i$ , there is a unique  $j \in \{0, \dots, b-1\}$  such that  $D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j)}$ . In other words, no edge is added to  $\mathcal{H}_i$  more than once.*

*Proof.* If an edge  $D$  is added to  $\mathcal{H}_i$  in step  $j$ , i.e., if  $D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j)}$ , then  $D \in M_i^i(\mathcal{H}_i^{(j+1)})$  and consequently all edges containing  $D$  are deleted from  $\mathcal{A}_{i+1}^{(j+1)}$ . It follows that  $D \cup \{u_{j'}\} \notin \mathcal{A}_{i+1}^{(j')}$  for every  $j' \in \{j+1, \dots, b-1\}$ .  $\square$

The next lemma shows that if  $\mathcal{H}_{i+1}$  satisfies (P3), then so does  $\mathcal{H}_i$ . The lemma follows easily from the definitions of  $\Delta_\ell^i$  and  $M_\ell^i(\mathcal{G})$ .

**Lemma 3.8.** *If  $\Delta_{\ell+1}(\mathcal{H}_{i+1}) \leq \Delta_{\ell+1}^{i+1}$  for some  $\ell \in [i-1]$ , then  $\Delta_\ell(\mathcal{H}_i) \leq \Delta_\ell^i$ .*

*Proof.* The crucial observation is that if

$$\deg_{\mathcal{H}_i^{(j)}}(T) \geq \frac{\Delta_\ell^i}{2}$$

for some  $T \in \binom{V(\mathcal{H})}{\ell}$  and  $j \in [b]$ , then all edges containing  $T$  are removed from  $\mathcal{A}_{i+1}^{(j)}$  and hence no more such edges are added to  $\mathcal{H}_i$ . It follows that  $\deg_{\mathcal{H}_i}(T) = \deg_{\mathcal{H}_i^{(j)}}(T)$ . Moreover, when we extend  $\mathcal{H}_i^{(j-1)}$  to  $\mathcal{H}_i^{(j)}$ , then we only add to it sets  $D$  such that  $D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j-1)} \subseteq \mathcal{H}_{i+1}$  and hence

$$\deg_{\mathcal{H}_i^{(j)}}(T) - \deg_{\mathcal{H}_i^{(j-1)}}(T) \leq \deg_{\mathcal{H}_{i+1}}(T \cup \{u_j\}) \leq \Delta_{|T|+1}(\mathcal{H}_{i+1}).$$

It follows that

$$\Delta_\ell(\mathcal{H}_i) \leq \frac{\Delta_\ell^i}{2} + \Delta_{\ell+1}(\mathcal{H}_{i+1}) \leq \frac{\Delta_\ell^i}{2} + \Delta_{\ell+1}^{i+1} \leq \Delta_\ell^i$$

where the last inequality follows from (6).  $\square$

Next, let us establish some simple properties of the numbers  $\Delta_\ell^i$ .

**Lemma 3.9.** *The following inequalities hold:*

- (a)  $\Delta_i^{i+1} \leq c2^k p^{-1}$  for every  $i \in [k-1]$  and
- (b)  $\Delta_1^i \leq c2^k p^{k-i} \frac{e(\mathcal{H})}{v(\mathcal{H})}$  for every  $i \in \{2, \dots, k\}$ .

*Proof.* To prove the lemma, simply note that, by the definition of  $\Delta_\ell^i$ , for every  $i \in [k]$  and every  $\ell \in [i-1]$ ,

$$\Delta_\ell^i = 2^d p^{k-i-d} \Delta_{d+\ell}(\mathcal{H}) \quad \text{for some } d \in \{0, \dots, k-i\}. \quad (7)$$

One easily proves (7) by induction on  $k-i$ . Intuitively,  $d$  in (7) is the number of times that the first term in the maximum in (6) is larger than the second term when following the recursive definition of  $\Delta_\ell^i$  back to  $\Delta_{d+\ell}^k$ .

Since  $\Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\}$ , as in the statement of Proposition 3.1, it follows from (7) that

$$\Delta_i^{i+1} = \max_{0 \leq d \leq k-i} \left\{ 2^d p^{k-(i+1)-d} \Delta_{d+i}(\mathcal{H}) \right\} \leq \max_{0 \leq d \leq k-i} \left\{ 2^d p^{k-i-1-d} \cdot c p^{d+i-k} \right\} \leq c \cdot 2^k p^{-1},$$

and

$$\Delta_1^i = \max_{0 \leq d \leq k-i} \left\{ 2^d p^{k-i-d} \Delta_{d+1}(\mathcal{H}) \right\} \leq \max_{0 \leq d \leq k-i} \left\{ 2^d p^{k-i-d} \cdot c p^d \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\} \leq c \cdot 2^k p^{k-i} \frac{e(\mathcal{H})}{v(\mathcal{H})},$$

as required.  $\square$

Finally, we show that if  $\mathcal{H}_{i+1}$  satisfies (P3) and (P4), then either  $\mathcal{H}_{i+1}$  also satisfies (P4) or we have  $|A_i| \leq (1 - c_i)v(\mathcal{H})$ . Recall that  $c_i = (ck2^{k+1})^{i-k}$ .

**Lemma 3.10.** *Let  $i \in [k - 1]$  and suppose that  $e(\mathcal{H}_{i+1}) \geq c_{i+1}p^{k-(i+1)}e(\mathcal{H})$  and that  $\Delta_\ell(\mathcal{H}_{i+1}) \leq \Delta_\ell^{i+1}$  for every  $\ell \in [i]$ . Then either*

$$e(\mathcal{H}_i) \geq \frac{p}{c \cdot 2^{k+1}k} e(\mathcal{H}_{i+1}) \geq c_i p^{k-i} e(\mathcal{H}) \quad \checkmark \quad (8)$$

or  $|A_i| \leq (1 - c_i)v(\mathcal{H})$ .  $\nwarrow$

*Proof.* If the Scythe Algorithm stops in step (1), then  $|A_i| = 0$  and there is nothing to prove. Hence, we may assume that steps (2)–(4) are executed  $b$  times. Since no edge is added to  $\mathcal{H}_i$  more than once, see Lemma 3.7, then for each  $j \in \{0, \dots, b - 1\}$ ,

$$e(\mathcal{H}_i^{(j+1)}) - e(\mathcal{H}_i^{(j)}) = \deg_{\mathcal{A}_{i+1}^{(j)}}(u_j). \quad (9)$$

By the definition of the max-deg order, the right-hand side of (9) is at least the average degree of the hypergraph  $\tilde{\mathcal{A}}_{i+1}^{(j)}$ , the subhypergraph of  $\mathcal{A}_{i+1}^{(j)}$  induced by the set  $(V(\mathcal{A}_{i+1}^{(j)}) \setminus W(u_j)) \cup \{u_j\}$ . Therefore, by the definition of  $\mathcal{A}_{i+1}^{(j+1)}$ , we have

$$e(\mathcal{H}_i^{(j+1)}) - e(\mathcal{H}_i^{(j)}) \geq \frac{(i+1)e(\tilde{\mathcal{A}}_{i+1}^{(j)})}{v(\tilde{\mathcal{A}}_{i+1}^{(j)})} \geq \frac{(i+1)e(\mathcal{A}_{i+1}^{(j+1)})}{v(\mathcal{H})}. \quad \leftarrow \text{Took out large degree edge}$$

Hence, if  $(i+1)e(\mathcal{A}_{i+1}^{(j+1)}) \geq e(\mathcal{H}_{i+1})$  for every  $j \in \{0, \dots, b - 1\}$ , then

$$e(\mathcal{H}_i) \geq \sum_{j=0}^{b-1} \frac{(i+1)e(\mathcal{A}_{i+1}^{(j+1)})}{v(\mathcal{H})} \geq b \cdot \frac{e(\mathcal{H}_{i+1})}{v(\mathcal{H})} = pe(\mathcal{H}_{i+1}),$$

as required. Thus, we may assume that for some  $j$ ,

$$e(\mathcal{A}_{i+1}^{(b)}) \leq e(\mathcal{A}_{i+1}^{(j+1)}) < \frac{e(\mathcal{H}_{i+1})}{i+1}. \quad (10)$$

Intuitively, (10) means that while running the Scythe Algorithm on  $\mathcal{H}_{i+1}$  and  $I$ , many edges are removed from  $\mathcal{A}_{i+1}$  (that is,  $\mathcal{H}_{i+1}$ ) in step (4). This may happen for one of the following two reasons: either many of the initial segments  $W(u_j)$  are long or one of the families  $M_\ell^i(\mathcal{H}_i)$  of sets with high degree in  $\mathcal{H}_i$  is large.

**Claim.** Either

$$\sum_{j=0}^{b-1} |W(u_j)| \geq \frac{1}{4\Delta_1^{i+1}} \cdot e(\mathcal{H}_{i+1})$$

or for some  $\ell \in [i]$ ,

$$|M_\ell^i(\mathcal{H}_i)| \geq \frac{1}{2(i+1)\Delta_\ell^{i+1}} \cdot e(\mathcal{H}_{i+1}).$$

*Proof of claim.* Recall that  $\mathcal{A}_{i+1}^{(0)} = \mathcal{H}_{i+1}$  and observe that for every  $j \in \{0, \dots, b-1\}$ ,

$$e(\mathcal{A}_{i+1}^{(j)}) - e(\mathcal{A}_{i+1}^{(j+1)}) \leq |W(u_j)| \cdot \Delta_1(\mathcal{H}_{i+1}) + \sum_{\ell=1}^i \left| M_\ell^i(\mathcal{H}_i^{(j+1)}) \setminus M_\ell^i(\mathcal{H}_i^{(j)}) \right| \cdot \Delta_\ell(\mathcal{H}_{i+1}). \quad (11)$$

Inequality (11) follows since in step (4) of the Scythe Algorithm, we remove from  $\mathcal{A}_{i+1}^{(j)}$  only the edges that contain either a vertex of  $W(u_j)$  or a member of  $M_\ell^i(\mathcal{H}_i^{(j+1)})$  for some  $\ell \in [i]$ . Thus, since  $\Delta_\ell(\mathcal{H}_{i+1}) \leq \Delta_\ell^{i+1}$  for every  $\ell \in [i]$ , summing (11) over all  $j$ , we get

$$e(\mathcal{H}_{i+1}) - e(\mathcal{A}_{i+1}^{(b)}) \leq \sum_{j=0}^{b-1} |W(u_j)| \cdot \Delta_1^{i+1} + \sum_{\ell=1}^i \left| M_\ell^i(\mathcal{H}_i^{(b)}) \right| \cdot \Delta_\ell^{i+1}.$$

Since we assumed that  $e(\mathcal{A}_{i+1}^{(b)}) < e(\mathcal{H}_{i+1})/(i+1)$ , see (10), and  $\mathcal{H}_i = \mathcal{H}_i^{(b)}$ , it follows that if

$$\sum_{j=0}^{b-1} |W(u_j)| \cdot \Delta_1^{i+1} < \frac{e(\mathcal{H}_{i+1})}{4} \leq \frac{i}{2(i+1)} \cdot e(\mathcal{H}_{i+1}), \quad (\text{one alternative})$$

then

$$\left| M_\ell^i(\mathcal{H}_i) \right| \cdot \Delta_\ell^{i+1} \geq \frac{1}{2(i+1)} \cdot e(\mathcal{H}_{i+1}) \quad \text{for some } \ell \in [i],$$

as claimed.  $\square$

Finally, let us deal with the two cases implied by the claim. In the remainder of the proof, we will show that if  $M_\ell^i(\mathcal{H}_i)$  is large for some  $\ell \in [i]$ , then  $e(\mathcal{H}_i)$  is large and if  $\sum_{j=0}^{b-1} |W(u_j)|$  is large, then  $|A_i|$  is small.

**Case 1:**  $\left| M_\ell^i(\mathcal{H}_i) \right| \geq \frac{1}{2(i+1)\Delta_\ell^{i+1}} \cdot e(\mathcal{H}_{i+1})$  for some  $\ell \in [i]$ .

If  $\ell < i$ , then  $\deg_{\mathcal{H}_i}(T) \geq \Delta_\ell^i/2$  for every  $T \in M_\ell^i(\mathcal{H}_i)$ , so by the handshaking lemma,

$$e(\mathcal{H}_i) = \binom{i}{\ell}^{-1} \sum_{T \in \binom{V(\mathcal{H}_i)}{\ell}} \deg_{\mathcal{H}_i}(T) \geq \frac{\left| M_\ell^i(\mathcal{H}_i) \right| \cdot \Delta_\ell^i}{2 \binom{i}{\ell}}. \quad (12)$$

Recalling that  $\Delta_\ell^i \geq p\Delta_\ell^{i+1}$ , see (6), we have

$$e(\mathcal{H}_i) \geq \frac{e(\mathcal{H}_{i+1})}{4(i+1) \binom{i}{\ell}} \cdot \frac{\Delta_\ell^i}{\Delta_\ell^{i+1}} \geq \frac{p}{2^{i+2}(i+1)} \cdot e(\mathcal{H}_{i+1}) \geq \frac{p}{2^{k+1}k} \cdot e(\mathcal{H}_{i+1}), \quad \binom{i}{\ell} \leq 2^i$$

as required. On the other hand, if  $\ell = i$ , then recalling that  $\Delta_i^{i+1} \leq c2^k p^{-1}$ , see Lemma 3.9, we have

$$e(\mathcal{H}_i) = \left| M_i^i(\mathcal{H}_i) \right| \geq \frac{e(\mathcal{H}_{i+1})}{2(i+1)\Delta_i^{i+1}} \geq \frac{p}{c2^{k+1}(i+1)} \cdot e(\mathcal{H}_{i+1}) \geq \frac{p}{c2^{k+1}k} \cdot e(\mathcal{H}_{i+1}),$$

as required.

**Case 2:**  $\sum_{j=0}^{b-1} |W(u_j)| \geq \frac{1}{4\Delta_1^{i+1}} \cdot e(\mathcal{H}_{i+1})$ .



We claim that in this case,  $|A_i| \leq (1 - c_i)v(\mathcal{H})$ . Indeed, we have

$$v(\mathcal{H}) - |A_i| = v(\mathcal{A}_{i+1}^{(0)}) - v(\mathcal{A}_{i+1}^{(b)}) = \sum_{j=0}^{b-1} |W(u_j)| \geq \frac{e(\mathcal{H}_{i+1})}{4\Delta_1^{i+1}}.$$

Recall that  $\Delta_1^{i+1} \leq c2^k p^{k-i-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}$  by Lemma 3.9. Thus,

$$v(\mathcal{H}) - |A_i| \geq \frac{p^{i+1-k}}{c2^{k+2}} \cdot \frac{v(\mathcal{H})}{e(\mathcal{H})} \cdot e(\mathcal{H}_{i+1}) \geq c_i v(\mathcal{H}),$$

since  $e(\mathcal{H}_{i+1}) \geq c_{i+1} p^{k-(i+1)} e(\mathcal{H})$  and  $c_{i+1}/(c2^{k+2}) \geq c_i$ .  $\square$

### 3.2. The proof of Proposition 3.1 and Theorem 2.2.

*Proof of Proposition 3.1.* Let  $k$  be an integer and let  $c$  and  $c'$  be two positive constants. Furthermore, let  $p \in (0, 1)$  and let  $\mathcal{H}$  be a  $k$ -uniform hypergraph that satisfy the assumptions of Proposition 3.1. Let  $\delta = (ck2^{k+1})^{1-k} c'$  and let  $b = pv(\mathcal{H})$ . We shall use the Scythe Algorithm, described in Section 3.1, to construct a family  $\mathcal{S}$  and functions  $f_0$  and  $g_0$  as in the statement of Proposition 3.1. We shall obtain them by running the following algorithm (with  $\mathcal{H}_k = \mathcal{H}$ ) on every independent set  $I \in \mathcal{I}(\mathcal{H})$ . We shall define  $f_0$  somewhat implicitly by defining a function  $f_0^*: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{P}(V(\mathcal{H}))$  that is constant on the set  $g_0^{-1}(S)$  for every  $S \in \mathcal{S}$ .

**Constructing  $g_0$  and  $f_0^*$ .** Given an  $I \in \mathcal{I}(\mathcal{H})$ , set  $i = k - 1$  and repeat the following:

- (1) Apply the Scythe Algorithm to  $\mathcal{H}_{i+1}$  and  $I$ . Suppose that it outputs  $\mathcal{H}_i$ ,  $A_i$  and  $B_i$ .
- (2) If  $|A_i| \leq (1 - \delta)v(\mathcal{H})$ , then set  $q = i$ ,  $r = i + 1$  and STOP.
- (3) If  $i > 1$ , then set  $i = i - 1$ . Otherwise, set  $q = r = 1$  and STOP.

Let  $I$  be an independent set and let us execute the above procedure (with  $\mathcal{H}_k = \mathcal{H}$ ) on  $I$ . We claim that for every  $i \in \{r, \dots, k\}$ , the hypergraph  $\mathcal{H}_i$  satisfies properties (P1)–(P4) defined in Section 3.1. This follows by induction on  $k - i$ . The base of the induction, the case  $i = k$ , follows vacuously from the definitions of  $c_k$  and  $\Delta_k^\ell$  for  $\ell \in [k - 1]$ . The inductive step follows from Lemmas 3.5, 3.8 and 3.10. To see this, note that since  $|A_i| > (1 - \delta)v(\mathcal{H}) \geq (1 - c_i)v(\mathcal{H})$  for all  $i \in \{r, \dots, k - 1\}$ , then (8) in Lemma 3.10 always holds.

Now, let us define  $g_0(I)$  and  $f_0^*(I)$ . Suppose first that  $r > 1$  and note that in this case, the algorithm stopped in step (2), which means that  $|A_q| \leq (1 - \delta)v(\mathcal{H})$ ; we set

$$g_0(I) = B_{k-1} \cup \dots \cup B_q \quad \text{and} \quad f_0^*(I) = A_q.$$

On the other hand, if  $r = 1$ , then we set

$$g_0(I) = B_{k-1} \cup \dots \cup B_1 \quad \text{and} \quad f_0^*(I) = \{v \in V(\mathcal{H}_1) : \{v\} \notin \mathcal{H}_1\}.$$

Finally, we let

$$\mathcal{S} = \{g_0(I) : I \in \mathcal{I}(\mathcal{H})\}.$$

We will define  $f_0$  by letting  $f_0(S) = f_0^*(I)$  for some  $I \in g_0^{-1}(S)$ . We first show that this definition will not depend on the choice of  $I$ . In fact, we shall prove a slightly stronger statement, which also establishes the consistency property of  $g_0$  stated in the final line of Proposition 3.1.

*S can be small enough*

**Claim.** Suppose that for some  $I, I' \in \mathcal{I}(\mathcal{H})$ ,  $g_0(I) \subseteq I'$  and  $g_0(I') \subseteq I$ . Then  $g_0(I) = g_0(I')$  and  $f_0^*(I) = f_0^*(I')$ .

*Proof of claim.* Suppose that while running the algorithm on some  $I$ , we obtain a sequence  $(B_{k-1}, \dots, B_q)$ . Since  $g_0(I)$  depends solely on  $(B_{k-1}, \dots, B_q)$  and, by Lemma 3.5, for each  $i$ , the hypergraph  $\mathcal{H}_i$  and the set  $A_i$  depend only on  $(B_{k-1}, \dots, B_i)$ , then also  $f_0^*(I)$  depends solely on  $(B_{k-1}, \dots, B_q)$ . Hence, it suffices to show that if, while running the algorithm on some  $I'$  with  $B_{k-1} \cup \dots \cup B_q \subseteq I'$ , we obtain a sequence  $(B'_{k-1}, \dots, B'_{q'})$  with  $B'_{k-1} \cup \dots \cup B'_{q'} \subseteq I$ , then  $(B'_{k-1}, \dots, B'_{q'}) = (B_{k-1}, \dots, B_q)$ . To this end, let us first observe that, under the above assumptions, for every  $i \in [k-1]$ , if  $\mathcal{H}_{i+1} = \mathcal{H}'_{i+1}$ , then  $B_i = B'_i$ . Indeed, note that  $B_i$  and  $B'_i$  are the outputs of the Scythe Algorithm executed on the inputs  $(\mathcal{H}_{i+1}, I)$  and  $(\mathcal{H}'_{i+1}, I')$ , respectively. Hence, if  $\mathcal{H}_{i+1} = \mathcal{H}'_{i+1}$ , then since

$$B_i \subseteq B_{k-1} \cup \dots \cup B_q \subseteq I' \quad \text{and} \quad B'_i \subseteq B'_{k-1} \cup \dots \cup B'_{q'} \subseteq I,$$

then Lemma 3.6 implies that  $B_i = B'_i$ . Since clearly  $\mathcal{H}_k = \mathcal{H}'_k = \mathcal{H}$  and, as noted before, for each  $i$ ,  $\mathcal{H}_{i+1}$  depends only on  $(B_{k-1}, \dots, B_{i+1})$ , it follows that  $B_i = B'_i$  for all  $i$ , as required.  $\square$

By the above claim, we can define  $f_0$  by letting, for every  $S \in \mathcal{S}$ ,  $f_0(S) = f_0^*(I)$  for any  $I \in g_0^{-1}(S)$ . Finally, let us show that the  $\mathcal{S}$ ,  $g_0$ , and  $f_0$ , which we have just defined, satisfy the required conditions, that is, for all  $I, I' \in \mathcal{I}(\mathcal{H})$ ,

- (i)  $|S| \leq (k-1)pv(\mathcal{H})$  for every  $S \in \mathcal{S}$ ,
- (ii)  $g_0(I) \subseteq I \subseteq f_0(g_0(I)) \cup g_0(I)$ ,
- (iii)  $|f_0(g_0(I))| \leq (1-\delta)v(\mathcal{H})$ ,
- (iv)  $g_0(I) \subseteq I'$  and  $g_0(I') \subseteq I$  imply that  $g_0(I) = g_0(I')$ .

To see (i), simply recall that  $|B_i| \leq pv(\mathcal{H})$  for every  $i \in [k-1]$ . To see (ii), note that  $B_i \subseteq I \subseteq A_i \cup B_i$  for every  $i \in \{q, \dots, k-1\}$ , by Lemma 3.5, that  $I$  is an independent set in  $\mathcal{H}_1$  (if  $r = 1$ ) and, crucially, that  $f_0(g_0(I)) = f_0^*(I)$ . To see (iii), note that if  $r > 1$ , then  $|A_q| \leq (1-\delta)v(\mathcal{H})$ , see step (2) of the algorithm; if  $r = 1$ , then observe that  $|\{v \in V(\mathcal{H}_1) : \{v\} \notin \mathcal{H}_1\}| \leq (1-\delta)v(\mathcal{H})$  since  $\mathcal{H}_1$  satisfies property (P4) and hence

$$e(\mathcal{H}_1) \geq c_1 p^{k-1} e(\mathcal{H}) \geq c_1 c' \cdot v(\mathcal{H}) \gg \delta v(\mathcal{H}),$$

where the second inequality follows from our assumption that  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$ . Finally, (iv) follows directly from the claim.  $\square$

*Proof of Theorem 2.2.* The theorem follows by applying Proposition 3.1 a bounded number of times. Given an integer  $k$  and positive reals  $c, c'$  and  $\varepsilon$ , let  $\delta = \delta_{3.1}(c/\varepsilon, \varepsilon c')$  and let

$$C = (k-1) \cdot \left( \frac{1}{\delta} \log \frac{1}{\varepsilon} + 1 \right).$$

Let  $V$  be a finite set and let  $\mathcal{F}$  be an increasing family of subsets of  $V$  such that  $|A| \geq \varepsilon|V|$  for every  $A \in \mathcal{F}$ . Let  $p \in (0, 1)$  and suppose that  $\mathcal{H}$  is a  $k$ -uniform hypergraph on the vertex set  $V$  that is  $(\mathcal{F}, \varepsilon)$ -dense and satisfies the assumptions of the theorem, that is,  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$  and

$$\Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\}$$

$|f_0(g_0(I))| \leq (1-\delta)v$

??

for every  $\ell \in [k-1]$ . We now show how to construct a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leq Cpv(\mathcal{H})}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{S}$  such that

$$g(I) \subseteq I \quad \text{and} \quad I \setminus g(I) \subseteq f(g(I)) \quad (13)$$

for every  $I \in \mathcal{I}(\mathcal{H})$ . Similarly as in the proof of Proposition 3.1, we shall define  $f$  via a function  $f^*: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{P}(V)$  that is constant on each set  $g^{-1}(S)$  with  $S \in \mathcal{S}$ .

Fix some  $I \in \mathcal{I}(\mathcal{H})$ . Using Proposition 3.1, we shall construct a sequence  $(A_j, S_j)_{j=1}^J$  of pairs of subsets of  $V$  such that for each  $j$ ,

$$S_1 \cup \dots \cup S_j \subseteq I \subseteq A_j \cup S_1 \cup \dots \cup S_j.$$

Moreover,  $A_j \in \overline{\mathcal{F}}$  while  $|S_1 \cup \dots \cup S_j| \leq Cpv(\mathcal{H})$ . Crucially, the set  $A_j$  will depend solely on  $S_1 \cup \dots \cup S_j$ . We will let  $g(I) = S_1 \cup \dots \cup S_J$  and  $f^*(I) = A_J$ .

**Construction.** Let  $S_0 = \emptyset$  and let  $A_0 = V$ . For  $j = 0, 1, \dots$ , do the following:

- (1) If  $A_j \in \mathcal{F}$ , then let  $I_j = I \cap A_j$  and apply Proposition 3.1 with  $c_{3.1} = c/\varepsilon$ ,  $c'_{3.1} = \varepsilon c'$  and  $p_{3.1} = p$  to the hypergraph  $\mathcal{H}[A_j]$  and the set  $I_j$  to obtain sets  $g_0(I_j)$  and  $f_0(g_0(I_j))$  such that  $g_0(I_j) \subseteq I_j$  and  $I_j \setminus g_0(I_j) \subseteq f_0(g_0(I_j))$ . Otherwise, if  $A_j \in \overline{\mathcal{F}}$ , then STOP.
- (2) Let  $S_{j+1} = g_0(I_j)$  and let  $A_{j+1} = f_0(g_0(I_j))$ .

Let us first show that the above procedure is well-defined, that is, that the assumptions of Proposition 3.1 are satisfied each time we are in (1). To this end, fix some  $A \subseteq V$  and note that if  $A \in \mathcal{F}$ , then, since  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense,

$$p^{k-1}e(\mathcal{H}[A]) \geq \varepsilon p^{k-1}e(\mathcal{H}) \geq \varepsilon c'v(\mathcal{H}) \geq \varepsilon c'v(\mathcal{H}[A])$$

and

$$\Delta_\ell(\mathcal{H}[A]) \leq \Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\} \leq \frac{c}{\varepsilon} \cdot \min \left\{ p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H}[A])}{v(\mathcal{H}[A])} \right\}.$$

Next, let us show that the above procedure terminates, therefore producing a finite sequence  $(A_j, S_j)$  with  $j \in [J]$ . To this end, let us simply note that by Proposition 3.1,  $|A_{j+1}| \leq (1-\delta)|A_j|$  for all  $j$ ,  $A_0 = V$  and  $|A| \geq \varepsilon|V|$  for every  $A \in \mathcal{F}$ . Moreover, since  $A_{J-1} \in \mathcal{F}$ , then

$$\varepsilon|V| \leq |A_{J-1}| \leq (1-\delta)^{J-1}|A_0| = \exp(-(J-1)\delta)|V|$$

and hence  $J \leq \frac{1}{\delta} \log \frac{1}{\varepsilon} + 1$ . It immediately follows that

$$|g(I)| \leq \sum_{j=1}^J |S_j| \leq \sum_{j=1}^J (k-1)pv(\mathcal{H}[A_j]) \leq J(k-1)pv(\mathcal{H}) \leq Cpv(\mathcal{H}).$$

Finally, let  $\mathcal{S} = \{g(I) : I \in \mathcal{I}(\mathcal{H})\}$ . It remains to show that for every  $S \in \mathcal{S}$ ,  $f^*$  is constant on  $g^{-1}(S)$ . Similarly as in the proof of Proposition 3.1, we shall prove a somewhat stronger statement.

**Claim.** Suppose that for some  $I, I' \in \mathcal{I}(\mathcal{H})$ ,  $g(I) \subseteq I'$  and  $g(I') \subseteq I$ . Then  $g(I) = g(I')$  and  $f^*(I) = f^*(I')$ .

*Proof of claim.* Suppose that while running the above procedure on some  $I$ , we generate a sequence  $(A_j, S_j)_{j=1}^J$ . Since for each  $j$ ,  $A_{j+1}$  depends solely on  $A_j$  and  $S_{j+1}$ , where  $A_0 = V$ , then both  $g(I)$  and  $f^*(I)$  depend solely on  $(S_1, \dots, S_J)$ . Hence, it suffices to show that if, while running the above procedure on some  $I'$  with  $S_1 \cup \dots \cup S_J \subseteq I'$ , we generate a sequence  $(A'_j, S'_j)_{j=1}^{J'}$  with  $S'_1 \cup \dots \cup S'_{J'} \subseteq I$ , then  $(S_1, \dots, S_J) = (S'_1, \dots, S'_{J'})$ . To this end, it suffices to note that if  $A_j = A'_j$ , then, since

$$S_{j+1} \subseteq S_1 \cup \dots \cup S_J \subseteq I' \quad \text{and} \quad S'_{j+1} \subseteq S'_1 \cup \dots \cup S'_{J'} \subseteq I,$$

by the consistency property of  $g_0$  stated in the final line of Proposition 3.1,  $S_{j+1} = S'_{j+1}$ . Since  $A_0 = A'_0 = V$  and for each  $j$ ,  $A_j$  depends only on  $(S_1, \dots, S_j)$ , it follows that  $S_j = S'_j$  for all  $j$ , as required.  $\square$

Finally, for every  $S \in \mathcal{S}$ , we let  $f(S) = f^*(I)$  for some  $I \in g^{-1}(S)$ . This completes the proof of Theorem 2.2.

$$g(\mathcal{D}) = \mathcal{S} \subseteq \mathcal{I} \quad g(\mathcal{I}') = \mathcal{S} \subseteq \mathcal{I}' \Rightarrow f^*(\mathcal{D}) = f^*(\mathcal{I}') \quad \square$$

#### 4. SZEMERÉDI'S THEOREM FOR SPARSE SETS

In this section, we prove Theorem 1.1 and derive from it Corollary 1.2. Before we get to the proofs, let us first remark that Theorem 1.1 and Corollary 1.2 are both sharp up to the value of the constant  $C$  in the lower bounds for  $p$  and  $m$ . More precisely, let us make the following two observations.

- (1) For every  $\beta \in (0, 1)$ , there is a positive  $c$  such that if  $m \leq cn^{1-1/(k-1)}$ , then the number of  $m$ -subsets of  $[n]$  that contain no  $k$ -term AP is at least  $(1 - \beta)^m \binom{n}{m}$ . To see this, let  $\varepsilon = \beta^2$  and observe that if  $c$  is sufficiently small and  $m \leq cn^{1-1/(k-1)}$ , then the expected number of  $k$ -term APs in a random  $(1 + \varepsilon)m$ -subset of  $[n]$  is smaller than  $\varepsilon m/2$  and hence by Markov's inequality, at least half of all  $(1 + \varepsilon)m$ -subsets of  $[n]$  contain a subset of size  $m$  with no  $k$ -term AP. Hence

$E(\#) \leq (1+\varepsilon)m \cdot n \cdot \binom{(1+\varepsilon)m}{n}^{k-1}$   
 $P(\# \geq \varepsilon m) \leq \frac{1}{2}$   
 Remove one from each progression

$$\#\{m\text{-subsets of } [n] \text{ with no } k\text{-term AP}\} \geq \frac{\binom{n}{(1+\varepsilon)m}}{2 \binom{n}{\varepsilon m}} \geq (1 - \sqrt{\varepsilon})^m \binom{n}{m}.$$

- (2) There is a positive constant  $c$  such that if  $p_n \leq cn^{-1/(k-1)}$ , then

$$\mathbb{P}([n]_{p_n} \text{ is } (\delta, k)\text{-Szemerédi}) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

$X = \text{random set}$   
 $n^2 p^k \ll np \Rightarrow \#(k\text{-AP in } X) \ll |X|$   
 $\Rightarrow \text{large subset with no } k\text{-AP.}$

For a (simple) proof of this statement, we refer the reader to [58].

We shall in fact prove the following somewhat stronger version of Corollary 1.2, originally proved by Schacht [58] (the approach of Conlon and Gowers [12] yields a somewhat weaker probability estimate).

**Corollary 4.1.** *For every  $k \in \mathbb{N}$  and every  $\delta \in (0, 1)$ , there exists a constant  $C$  such that for all sufficiently large  $n$ , if  $p \geq Cn^{-1/(k-1)}$ , then*

$$\mathbb{P}([n]_p \text{ is } (\delta, k)\text{-Szemerédi}) \geq 1 - 2 \exp(-pn/8).$$

In the proofs of Theorem 1.1 and Corollary 4.1, and frequently in later sections, we shall need various estimates on binomial coefficients, which we list here for future reference. Let  $a, b$ , and  $c$  be integers satisfying  $a \geq b \geq c \geq 0$ . Then the following inequalities hold:

## L4.2

Fix  $\delta$ . Let  $A \subseteq [n]$  satisfy  $|A| \geq \delta n$ .

Let  $L$  be such that every  $\frac{1}{2}\delta L$ -subset of  $[L]$  has a  $k$ -AP.

Now consider all AP's

$$\bar{\Phi}(u, d) = \{u, u+d, u+2d, \dots, u+(L-1)d\}$$

where  $d \leq \frac{\delta}{L^2} n$ .

$\bar{\Phi} = \bar{\Phi}(u, d)$  is good if  $|A \cap \bar{\Phi}| \geq \frac{1}{2}\delta L$ .

Fix  $d$  and choose  $u$  at random from  $[1, n - kLd]$

$$E(|\bar{\Phi}(u, d) \cap A|) = \sum_{i=0}^{k-1} P_i[u+id \in A] \geq \frac{L(\delta n - 2kLd)}{n - kLd}$$

Let  $\phi(d) = \#$  of good  $\bar{\Phi}(\cdot, d)$ . Then

$$E(|\bar{\Phi}(u, d) \cap A|) \leq \frac{\phi(d)}{n - 2kLd} \cdot L + \frac{1}{2}\delta L$$

$\implies \phi(d) \geq \frac{\delta n}{3}$  for large  $n$ .

$\bar{\Phi}$  good  $\implies \bar{\Phi} \supseteq$  a  $k$ -AP.

$$\text{Now } \sum_d \phi(d) \geq \frac{\delta n}{3} \cdot \frac{\delta n}{L^2}.$$

A  $k$ -AP progression with step  $d^*$  can be in at most

$L$   $\bar{\Phi}(u, d^*)$  and at most  $L^2$   $\bar{\Phi}(u, d)$  where  $d|d^*$  since

we must have  $Ld \geq d^*$  i.e.  $d^* = \alpha d$  where  $\alpha \in [L]$ .

$$\implies \# \text{ } k\text{-AP} \geq \frac{\delta^2}{6L^4} n^2.$$

$$\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b, \quad (14) \quad \binom{b}{c} \leq \left(\frac{b}{a}\right)^c \binom{a}{c}, \quad (15)$$

$$\binom{a}{b-c} \leq \left(\frac{b}{a-b}\right)^c \binom{a}{b}, \quad (16) \quad \binom{a}{c} \leq \left(\frac{a-c}{b-c}\right)^c \binom{b}{c}. \quad (17)$$

We remark that each inequality above follows easily from the definition of  $\binom{a}{b}$ .

*Proof of Corollary 4.1.* Fix  $k \in \mathbb{N}$  and  $\delta \in (0, 1)$ , let  $\beta = \delta/(2e) \cdot e^{-1/\delta}$ , and set  $C = 2C_{1.1}(\beta, k)/\delta$ . Assume that  $p \geq Cn^{-1/(k-1)}$ , let  $m = \delta pn/2$ , and let  $X_m$  denote the number of  $k$ -term-AP-free  $m$ -subsets of  $[n]_p$ . By Theorem 1.1 and (14), we have

$$\mathbb{P}(X_m > 0) \leq \mathbb{E}[X_m] \leq \binom{\beta n}{m} p^m \leq \left(\frac{\beta e p n}{m}\right)^m = \left(\frac{2\beta e}{\delta}\right)^m = e^{-m/\delta}. \quad (18)$$

Let  $\mathcal{A}$  denote the event that  $[n]_p$  is *not*  $(\delta, k)$ -Szemerédi, i.e., that there exists a  $k$ -term-AP-free subset of  $[n]_p$  with  $\delta|[n]_p|$  elements. By (18) and Chernoff's inequality (see, e.g., [3, Appendix A]), it follows that

$$\mathbb{P}(\mathcal{A}) \leq \mathbb{P}\left(\underbrace{\mathcal{A} \wedge |[n]_p| \geq \frac{pn}{2}}_{\delta|[n]_p| \geq m}\right) + \mathbb{P}\left(|[n]_p| < \frac{pn}{2}\right) \leq \mathbb{P}(X_m > 0) + e^{-pn/8} \leq 2e^{-pn/8},$$

as required.  $\square$

Finally, let us show how to deduce Theorem 1.1 from Theorem 2.2. Our proof will use the following robust version of Szemerédi's theorem, which can be proved by a simple averaging argument, originally observed by Varnavides [65].

**Lemma 4.2.** *For every positive  $\delta$  and  $k \in [n]$ , there exists a positive  $\varepsilon$  such that the following holds for all sufficiently large  $n$ . Every subset of  $[n]$  with at least  $\delta n$  elements contains at least  $\varepsilon n^2$   $k$ -term APs.*

*Proof of Theorem 1.1.* Given  $k \in \mathbb{N}$  and positive  $\beta$ , let  $\delta = \min\{\beta/2, 1/10\}$  and let  $n \in \mathbb{N}$  be sufficiently large. Let  $\mathcal{H}$  be the  $k$ -uniform hypergraph of  $k$ -term APs in  $[n]$ , i.e., the hypergraph on the vertex set  $[n]$  whose edges are all  $k$ -term APs in  $[n]$ , let  $\mathcal{F}$  denote the family of subsets of  $[n]$  with at least  $\delta n$  elements, and let  $\varepsilon = \varepsilon_{4.2}(\delta, k)$ . By Lemma 4.2, the hypergraph  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense, provided that  $n$  is sufficiently large. Let  $p = n^{-1/(k-1)}$ , let  $c' = 1/k^2$ , and let  $c = 2k^2$ . Since  $e(\mathcal{H}) \geq c'n^2$ , it follows that  $p^{k-1}e(\mathcal{H}) \geq c'v(\mathcal{H})$ . Moreover,

$$\Delta_1(\mathcal{H}) \leq k \cdot \frac{n}{k-1} \leq cc'n \leq c \cdot \min\left\{p^{1-k}, p^{1-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}\right\}$$

and for every  $\ell \in \{2, \dots, k-1\}$ ,

$$\Delta_\ell(\mathcal{H}) \leq \binom{k}{2} \leq cc'n^{1/(k-1)} \leq c \cdot \min\left\{p^{\ell-k}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}\right\}.$$

Let  $C' = C_{2.2}(k, \varepsilon, c, c')$  and let  $C = C'/\delta$  and assume that  $m \geq Cn^{1-1/(k-1)}$ . Note that if  $m > \delta n/2$ , then  $\mathcal{I}(\mathcal{H}, m) = 0$  by Szemerédi's theorem, so we may assume that  $m \leq \delta n/2$ . Since  $C'pn \leq \delta m$ , then by Theorem 2.2, there exists a family  $\mathcal{S} \subseteq \binom{[n]}{\leq C'pn} \subseteq \binom{[n]}{\leq \delta m}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{S}$ , such that for every  $I \in \mathcal{I}(\mathcal{H})$ ,

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

$$\delta m \geq C' n^{1-\frac{1}{k-1}} = C' n_p$$

Therefore, using (14) and (16), the number of independent sets of size  $m$  in  $\mathcal{H}$  can be estimated as follows:

$$\begin{aligned} |\mathcal{I}(\mathcal{H}, m)| &= \sum_{S \in \mathcal{S}} |\{I \in \mathcal{I}(\mathcal{H}, m) : g(I) = S\}| \leq \sum_{S \in \mathcal{S}} \binom{|f(S)|}{m - |S|} \\ &\leq \sum_{k \leq \delta m} \binom{n}{k} \binom{\delta n}{m - k} \leq \sum_{k \leq \delta m} \left(\frac{en}{k}\right)^k \underbrace{\left(\frac{m}{\delta n - m}\right)^k}_{\beta = 2\delta} \binom{\delta n}{m}. \end{aligned} \quad (16)$$

$a = \delta n$   
 $b = m$   
 $c = k$

Since  $m \leq \delta n/2$  and the function  $x \mapsto (y/x)^x$  is increasing on  $(0, y/e)$ , it follows that

$$|\mathcal{I}(\mathcal{H}, m)| \leq \sum_{k \leq \delta m} \left(\frac{2em}{\delta k}\right)^k \binom{\delta n}{m} \leq m \left(\frac{2e}{\delta^2}\right)^{\delta m} \binom{\delta n}{m} \leq \binom{\beta n}{m}, \quad \beta = 2\delta$$

where the final inequality follows since  $\binom{\delta n}{m} \leq 2^{-m} \binom{2\delta n}{m}$ , by (15), and since  $2^{1/\delta} > 2e/\delta^2$  if  $\delta \leq 1/10$ . This proves Theorem 1.1.  $\square$

Finally, from the same proof, combined with an analogue of Lemma 4.2 due to Furstenberg and Katznelson [25], we obtain the following generalization of Theorem 1.1, which strengthens Theorem 2.3 of [58]. Given a set  $F \subseteq \mathbb{N}^\ell$ , we call a set of the form  $a + bF = \{a + bx : x \in F\}$ , with  $a \in \mathbb{N}^\ell$  and  $b \in \mathbb{Z} \setminus \{0\}$ , a *homothetic copy* of  $F$ .

**Theorem 4.3.** *For every positive  $\beta$ , every  $\ell \in \mathbb{N}$ , and every finite configuration  $F \subseteq \mathbb{N}^\ell$ , there exist constants  $C$  and  $n_0$  such that the following holds. For every  $n \in \mathbb{N}$  with  $n \geq n_0$ , if  $m \geq Cn^{\ell-1}/(|F|-1)$ , then there are at most*

$$\binom{\beta n^\ell}{m}$$

*$m$ -subsets of  $[n]^\ell$  that contain no homothetic copy of  $F$ .*

Finally, we mention one more straightforward application of Theorem 2.2, which is a sparse version of a result of Sárközy [56] and Furstenberg [24] on square differences. A robust version of it was proved by Hamel and Łaba [32, Theorem 3.1] using a Varnavides-type averaging argument. The following theorem improves Theorem 1.2 of [32].

**Theorem 4.4.** *For every positive  $\beta$ , there exist constants  $C$  and  $n_0$  such that the following holds. For every  $n \in \mathbb{N}$  with  $n \geq n_0$ , if  $m \geq C\sqrt{n}$ , then there are at most*

$$\binom{\beta n}{m}$$

*$m$ -subsets of  $[n]$  that contain no pair  $\{x, y\}$  such that  $x - y$  is a perfect square.*

## 5. EXTREMAL RESULTS FOR SPARSE SETS

In this section, we shall deduce from Theorem 2.2 two versions of the general transference theorem of Schacht [58, Theorem 3.3]. We remind the reader that a statement very similar to Schacht's theorem was proved independently by Conlon and Gowers [12]. For the benefit of the readers who are familiar with [58], we shall state it using the terminology used there.

**Definition 5.1.** Let  $\mathcal{H} = (\mathcal{H}_n)_{n \in \mathbb{N}}$  be a sequence of  $k$ -uniform hypergraphs and let  $\alpha \in [0, 1)$ . We say that  $\mathcal{H}$  is  $\alpha$ -dense if the following is true: For every positive  $\delta$ , there exist positive  $\varepsilon$  and  $n_0$  such that for every  $n$  with  $n \geq n_0$  and every  $U \subseteq V(\mathcal{H}_n)$  with  $|U| \geq (\alpha + \delta)v(\mathcal{H}_n)$ , we have

$$e(\mathcal{H}_n[U]) \geq \varepsilon e(\mathcal{H}_n). \quad (\ast 25)$$

Let us remark here that Definition 2.1 is a generalization of Definition 5.1. Indeed, if  $\mathcal{F}_\delta$  denotes the collection of all subsets of  $V(\mathcal{H}_n)$  with at least  $(\alpha + \delta)v(\mathcal{H}_n)$  elements, then a sequence  $\mathcal{H}$  of hypergraphs is  $\alpha$ -dense if and only if for every positive  $\delta$ , there exists a positive  $\varepsilon$  such that for all sufficiently large  $n$ , the hypergraph  $\mathcal{H}_n$  is  $(\mathcal{F}_\delta, \varepsilon)$ -dense.

We start with the ‘random’ version of our extremal result, which is a slight weakening of Theorem 3.3 of [58], see the discussion below.

**Theorem 5.2.** Let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs, let  $\alpha \in [0, 1)$ , and let  $c$  and  $c'$  be positive constants. Suppose that  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$  is a sequence of probabilities such that for all sufficiently large  $n \in \mathbb{N}$ , we have  $p_n^{k-1}e(\mathcal{H}_n) \geq c'v(\mathcal{H}_n)$  and for every  $\ell \in [k-1]$ ,

$$\Delta_\ell(\mathcal{H}_n) \leq c \cdot \min \left\{ p_n^{\ell-k}, p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \right\}. \quad (19)$$

If  $\mathcal{H}$  is  $\alpha$ -dense, then the following holds. For every positive  $\delta$ , there exists a constant  $C$  such that if  $q_n \geq Cp_n$  and  $q_nv(\mathcal{H}_n) \rightarrow \infty$  as  $n \rightarrow \infty$ , then a.a.s.

$$\alpha(\mathcal{H}_n[V(\mathcal{H}_n)_{q_n}]) \leq (\alpha + \delta)q_nv(\mathcal{H}_n).$$

—— random vertex subset

We note that the probability bounds implicit in the ‘asymptotically almost surely’ statement that we obtain are, as in [58], optimal, that is, they decay exponentially in  $p_nv(\mathcal{H}_n)$ .

**Remark 5.3.** We remark that the only difference between Theorem 5.2 and Theorem 3.3 of [58] are the assumptions on the hypergraph sequence  $\mathcal{H}$ , which are somewhat more restrictive here. In fact, it turns out that the conditions  $p_n^{k-1}e(\mathcal{H}_n) \geq c'v(\mathcal{H}_n)$  and  $\Delta_\ell(\mathcal{H}_n) \leq c \cdot p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}$  for every  $\ell \in [k-1]$  are essentially equivalent to the condition that  $\mathcal{H}$  is  $(K, \mathbf{p})$ -bounded (see below), whereas the condition  $\Delta_\ell(\mathcal{H}_n) \leq cp_n^{\ell-k}$  for every  $\ell \in [k-1]$ , which is essential in the theorem above, is not needed in [58].

To be more precise, let us first recall from [58] that a sequence of  $k$ -uniform hypergraphs  $\mathcal{H}$  is said to be  $(K, \mathbf{p})$ -bounded if

$$\mu_i(\mathcal{H}_n, q) = \mathbb{E} \left[ \sum_{v \in V} \deg_i(v, V_q)^2 \right] \leq Kq^{2i} \frac{e(\mathcal{H}_n)^2}{v(\mathcal{H}_n)},$$

for every  $i \in [k-1]$ , every  $q$  with  $q \geq p_n$ , and every sufficiently large  $n$ , where  $V$  abbreviates  $V(\mathcal{H}_n)$  and  $\deg_i(v, V_q)$  denotes the number of edges of  $\mathcal{H}_n$  which contain  $v$  and at least  $i$  other vertices of  $V_q$ . We claim that if (a)  $p_n^{k-1}e(\mathcal{H}_n) \geq c'v(\mathcal{H}_n)$  and (b)  $\Delta_\ell(\mathcal{H}_n) \leq c \cdot p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}$  for every  $\ell \in [k-1]$ , then  $\mathcal{H}$  is  $(K, \mathbf{p})$ -bounded for some constant  $K$  that depends only on



$c$ ,  $c'$ , and  $k$ . Indeed, we have

$$\begin{aligned} \mu_i(\mathcal{H}_n, q) &\leq v(\mathcal{H}_n) \cdot \Delta_1(\mathcal{H}_n) \cdot \sum_{\ell=0}^i \Delta_{\ell+1}(\mathcal{H}_n) \binom{k-1}{i}^2 q^{2i-\ell} \\ &\leq c^2(i+1) \binom{k-1}{i}^2 q^{2i} \frac{e(\mathcal{H}_n)^2}{v(\mathcal{H}_n)} + \frac{c}{c'} q^{2(k-1)} \frac{e(\mathcal{H}_n)^2}{v(\mathcal{H}_n)} \leq K q^{2i} \frac{e(\mathcal{H}_n)^2}{v(\mathcal{H}_n)}, \end{aligned}$$

for every  $i \in [k-1]$  and  $q$  with  $q \geq p_n$ , where we used (b) to bound both  $\Delta_1(\mathcal{H}_n)$  and  $\Delta_{\ell+1}(\mathcal{H}_n)$  except in the case  $\ell = i = k-1$ , where we used (a) and the fact that  $\Delta_k(\mathcal{H}_n) = 1$ .

Conversely, suppose that  $\mathcal{H}$  is  $(K, \mathbf{p})$ -bounded and observe that  $\mu_{k-1}(\mathcal{H}_n, p_n) \geq p_n^{k-1} e(\mathcal{H}_n)$ . Thus, setting  $c' = 1/K$ , it follows that  $p_n^{k-1} e(\mathcal{H}_n) \geq c' v(\mathcal{H}_n)$ . Moreover, we claim that for every positive  $\varepsilon$ , there is a constant  $c$  that depends only on  $K$ ,  $k$ , and  $\varepsilon$ , such that for all sufficiently large  $n$ , the hypergraph  $\mathcal{H}_n$  contains a subhypergraph  $\mathcal{H}'_n \subseteq \mathcal{H}_n$  satisfying  $e(\mathcal{H}'_n) \geq (1-\varepsilon)e(\mathcal{H}_n)$  and

$$\Delta_\ell(\mathcal{H}'_n) \leq c \cdot p_n^{\ell-1} \frac{e(\mathcal{H}'_n)}{v(\mathcal{H}'_n)} \quad \text{for every } \ell \in [k-1]. \quad (20)$$

Indeed, fix a large constant  $c$  and suppose that  $\mathcal{H}_n$  does not contain a subhypergraph with at least  $(1-\varepsilon)e(\mathcal{H}_n)$  edges that satisfies (20). In this case, let us greedily construct a hypergraph  $\mathcal{H}''_n$  as follows. Start with  $\mathcal{H}'_n = \mathcal{H}_n$  and  $\mathcal{H}''_n$  empty. Whenever there is an  $\ell$ -set  $T \subseteq V(\mathcal{H}_n)$ , for some  $\ell \in [k-1]$ , whose degree in  $\mathcal{H}'_n$  exceeds the right-hand side of (20), then move an arbitrary edge containing  $T$  from  $\mathcal{H}'_n$  to  $\mathcal{H}''_n$ . By our assumption, when the process terminates,  $\mathcal{H}''_n$  will contain more than  $\varepsilon e(\mathcal{H}_n)$  edges. Now, if  $c$  is sufficiently large (as a function of  $\varepsilon$ ,  $k$ , and  $K$ ), then for some  $i \in [k-1]$  we have

$$\mu_i(\mathcal{H}_n, p_n) \geq \frac{e(\mathcal{H}''_n)}{k-1} \cdot c \cdot p_n^{\ell-1} \frac{e(\mathcal{H}'_n)}{v(\mathcal{H}'_n)} \cdot p_n^{2i-(\ell-1)} > K p_n^{2i} \frac{e(\mathcal{H}_n)^2}{v(\mathcal{H}_n)},$$

which is a contradiction.

Finally, note that, trivially, if  $\mathcal{H}_n$  is  $(\mathcal{F}, 2\varepsilon)$ -dense for some family  $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}_n))$ , then every  $\mathcal{H}'_n$  with  $e(\mathcal{H}'_n) \geq (1-\varepsilon)e(\mathcal{H}_n)$  is  $(\mathcal{F}, \varepsilon)$ -dense. It follows from the above discussion that, up to the value of the involved constants, the only difference between Theorem 5.2 and Theorem 3.3 of [58] is in the additional assumption that  $\Delta_\ell(\mathcal{H}_n) \leq c p_n^{\ell-k}$  for every  $\ell \in [k-1]$ .

Our methods also yield the following ‘counting’ analogue of Theorem 5.2, a generalization of Theorem 1.1 that does not follow from the methods of Schacht [58] or Conlon and Gowers [12] and, in the case  $\alpha = 0$ , can be thought of as a strengthening of Theorem 5.2, see Corollary 1.2.

**Theorem 5.4.** *Let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs, let  $\alpha \in [0, 1)$ , and let  $c$  and  $c'$  be positive constants. Suppose that  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$  is a sequence of probabilities such that for all sufficiently large  $n \in \mathbb{N}$ , we have  $p_n^{k-1} e(\mathcal{H}_n) \geq c' v(\mathcal{H}_n)$  and for every  $\ell \in [k-1]$ ,*

$$\Delta_\ell(\mathcal{H}_n) \leq c \cdot \min \left\{ p_n^{\ell-k}, p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \right\}.$$

If  $\mathcal{H}$  is  $\alpha$ -dense, then the following holds. For every positive  $\delta$ , there exists a constant  $C$  such that for all sufficiently large  $n$ , if  $m \geq Cp_nv(\mathcal{H}_n)$ , then

$$|\mathcal{I}(\mathcal{H}_n, m)| \leq \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m}$$

— independent in  $\mathcal{H}_n$  of size  $m$

*Proof of Theorem 5.2.* Let  $\alpha \in [0, 1)$ , let  $k \in \mathbb{N}$ , let  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$ , and let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs as in the statement of Theorem 5.2. Furthermore, suppose that  $\mathcal{H}$  is  $\alpha$ -dense and fix some positive  $\delta$ ; without loss of generality, we may assume that  $\delta$  is sufficiently small. Let  $n \in \mathbb{N}$  be sufficiently large, let  $\delta' = \delta/3$ , and let  $\mathcal{F}$  denote the family of all subsets of  $V(\mathcal{H}_n)$  with at least  $(\alpha + \delta')v(\mathcal{H}_n)$  elements. Since  $\mathcal{H}_n$  is  $\alpha$ -dense, it follows that  $\mathcal{H}_n$  is  $(\mathcal{F}, \varepsilon)$ -dense for some small positive  $\varepsilon$  that does not depend on  $n$ . Let  $C' = C_{2.2}(k, \varepsilon, c, c')$ . By Theorem 2.2, there exist a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H}_n)}{\leq C'p_nv(\mathcal{H}_n)}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}_n) \rightarrow \mathcal{S}$  such that

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

for every  $I \in \mathcal{I}(\mathcal{H}_n)$ . Let  $C = C'/\delta^3$  and assume that  $q_n \geq Cp_n$ . Let  $m = (\alpha + \delta)q_nv(\mathcal{H}_n)$  and, for the sake of brevity, let us write  $V = V(\mathcal{H}_n)$  and  $q = q_n$ . Observe that

$$\begin{aligned} \mathbb{P}(\alpha(\mathcal{H}_n[V_q]) \geq m) &= \mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}(\mathcal{H}_n, m)) \\ &\leq \sum_{S \in \mathcal{S}} \mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}(\mathcal{H}_n, m) \text{ such that } g(I) = S). \end{aligned} \quad (21)$$

Fix an  $S \in \mathcal{S}$  and let  $\mathcal{I}'_S = \{I \in \mathcal{I}(\mathcal{H}_n, m) : g(I) = S\}$ . We estimate the summand in the right-hand side of (21) as follows:

$$\mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}'_S) \leq \sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq V_q) \cdot \mathbb{P}(|V_q \cap f(S)| \geq m - |S|). \quad (22)$$

To see the above inequality, simply note that for every  $I \in \mathcal{I}'_S$ , we have  $I \setminus S \subseteq f(S)$ .

Now, since  $m = (\alpha + 3\delta')q_nv(\mathcal{H}_n)$  and  $S \in \mathcal{S}$ , then

$$|S| \leq C'p_nv(\mathcal{H}_n) \leq \delta^3qv(\mathcal{H}_n) \leq \delta'qv(\mathcal{H}_n)$$

$$C' \leq \delta^3 q \quad /$$

and hence  $m - |S| \geq (\alpha + 2\delta')qv(\mathcal{H}_n)$ . On the other hand, since  $|f(S)| \leq (\alpha + \delta')v(\mathcal{H}_n)$  by the definition of  $\mathcal{F}$ , then

$$\mathbb{E}[|V_q \cap f(S)|] \leq (\alpha + \delta')qv(\mathcal{H}_n).$$

Hence, by Chernoff's inequality, we have

$$\mathbb{P}(|V_q \cap f(S)| \geq m - |S|) \leq \exp\left(-\frac{(\delta')^2qv(\mathcal{H}_n)}{4}\right) = \exp\left(-\frac{\delta^2qv(\mathcal{H}_n)}{36}\right). \quad (23)$$

Finally, note that since  $|S| \leq \delta^3qv(\mathcal{H}_n)$  for every  $S \in \mathcal{S}$ , and using (14),

$$\sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq V_q) \leq \sum_{s=0}^{\delta^3qv(\mathcal{H}_n)} \binom{v(\mathcal{H}_n)}{s} q^s \leq v(\mathcal{H}_n) \cdot \left(\frac{e}{\delta^3}\right)^{\delta^3qv(\mathcal{H}_n)}. \quad (24)$$

$\alpha =$  size of  
tail  
set

Putting (21), (22), (23), and (24) together, we obtain

$$\mathbb{P}(\alpha(\mathcal{H}_n[V_q]) \geq m) \leq \sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq V_q) \exp\left(-\frac{\delta^2 q v(\mathcal{H}_n)}{36}\right) \leq \exp(-\delta^3 q v(\mathcal{H}_n)),$$

as required.  $\square$

*Proof of Theorem 5.4.* Let  $\alpha \in [0, 1)$ , let  $k \in \mathbb{N}$ , let  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$ , and let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs as in the statement of Theorem 5.2. Furthermore, suppose that  $\mathcal{H}$  is  $\alpha$ -dense and fix some positive  $\delta$ . Let  $n$  be sufficiently large, let  $\delta' = \delta/2$ , and let  $\mathcal{F}$  denote the family of all subsets of  $V(\mathcal{H}_n)$  with at least  $(\alpha + \delta')v(\mathcal{H}_n)$  elements. Since  $\mathcal{H}_n$  is  $\alpha$ -dense, it follows that  $\mathcal{H}_n$  is  $(\mathcal{F}, \varepsilon)$ -dense for some small positive  $\varepsilon$  that does not depend on  $n$ . Let  $C' = C_{2.2}(k, \varepsilon, c, c')$ . By Theorem 2.2, there exist a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H}_n)}{\leq C' p_n v(\mathcal{H}_n)}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}_n) \rightarrow \mathcal{S}$  such that

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

for every  $I \in \mathcal{I}(\mathcal{H}_n)$ . Let  $C = C'/\delta^2$  and assume that  $m \geq C p_n v(\mathcal{H}_n)$ . Fix an  $S \in \mathcal{S}$ , let  $\mathcal{I}_S = \{I \in \mathcal{I}(\mathcal{H}_n, m) : g(I) = S\}$ , and note for future reference that

$$|S| \leq C' p_n v(\mathcal{H}_n) \leq \delta^2 m. \tag{25}$$

Since  $f(S) \in \overline{\mathcal{F}}$ , we have  $|f(S)| < (\alpha + \delta')v(\mathcal{H}_n)$ . Therefore,

$$|\mathcal{I}_S| \leq \binom{|f(S)|}{m - |S|} \leq \binom{(\alpha + \delta')v(\mathcal{H}_n)}{m - |S|}. \quad \downarrow \mathcal{I} = (\mathcal{I}_S) \cup \mathcal{S}$$

To see the above inequality, simply note that for every  $I \in \mathcal{I}_S$ , we have  $I \setminus S \subseteq f(S)$ .

Now, observe that if  $m \geq (\alpha + \delta')v(\mathcal{H}_n)$ , then every  $m$ -subset of  $V(\mathcal{H}_n)$  belongs to  $\mathcal{F}$  and hence there is no independent set of size  $m$ . Therefore, from now on we may assume that  $m < (\alpha + \delta')v(\mathcal{H}_n) = (\alpha + \delta/2)v(\mathcal{H}_n)$ . It follows, using (15) and (16), that

$$|\mathcal{I}_S| \leq \binom{(\alpha + \delta')v(\mathcal{H}_n)}{m - |S|} \leq \left(\frac{2m}{\delta v(\mathcal{H}_n)}\right)^{|S|} (1 - \delta')^{m - |S|} \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m}. \tag{26}$$

Setting  $s = |S|$  and recalling that  $s \leq \delta^2 m$  by (25), we obtain

$$\binom{v(\mathcal{H}_n)}{s} \cdot |\mathcal{I}_S| \leq \left(\frac{2em}{\delta s}\right)^s (1 - \delta')^{m/2} \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m} \leq e^{-\delta^2 m} \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m},$$

provided that  $\delta$  is sufficiently small. It follows that

$$|\mathcal{I}(\mathcal{H}_n, m)| = \sum_{S \in \mathcal{S}} |\mathcal{I}_S| \leq \sum_{s=0}^{\delta^2 m} \binom{v(\mathcal{H}_n)}{s} \max\{|\mathcal{I}_S| : |S| = s\} \leq \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m},$$

as claimed.

*Handwritten notes and diagrams:*

- \* (MS) implies large vertex sets contain edge of  $\mathcal{H}_n$
- ~~$\binom{(\alpha + \delta')v(\mathcal{H}_n)}{m - s} \leq \left(\frac{2em}{\delta s}\right)^s (1 - \delta')^{m/2} \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m}$~~
- $\binom{(\alpha + \delta')v(\mathcal{H}_n)}{m - s} \leq \left(\frac{\alpha + \delta'}{\alpha + \delta}\right)^{m - s} \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m - s}$
- $\binom{(\alpha + \delta)v(\mathcal{H}_n)}{m - s} \leq \left(\frac{m}{\frac{\alpha + \delta'}{2} m}\right)^s \binom{(\alpha + \delta)v(\mathcal{H}_n)}{m}$

6. STABILITY RESULTS FOR SPARSE SETS

In this section, we shall deduce from Theorem 2.2 two versions of the general transference theorem for stability results proved by Conlon and Gowers [12]. Similarly as in Section 5, we shall state our results using the terminology used by Schacht [58]. We remark here that in parallel to this work, Schacht's method was adapted to yield sparse random analogues of stability statements by Samotij [55]. The main result of this section is most easily compared with Theorem 3.4 of [55]. We begin by recalling the following definition from [1].

**Definition 6.1.** Let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs, let  $\alpha$  be a positive real and let  $\mathcal{B}$  be a sequence of sets with  $B_n \subseteq \mathcal{P}(V(\mathcal{H}_n))$ . We say that  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable if for every positive  $\delta$ , there exist positive  $\varepsilon$  and  $n_0$  such that the following holds. For every  $n$  with  $n \geq n_0$  and every  $U \subseteq V(\mathcal{H}_n)$  with  $|U| \geq (\alpha - \varepsilon)v(\mathcal{H}_n)$ , we have either  $e(\mathcal{H}_n[U]) \geq \varepsilon e(\mathcal{H}_n)$  or  $|U \setminus B| \leq \delta v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ .

Roughly speaking, a sequence  $\mathcal{H}$  of hypergraphs is  $(\alpha, \mathcal{B})$ -stable if for every  $A \subseteq V(\mathcal{H}_n)$  that is almost as large as  $\alpha v(\mathcal{H}_n)$ , the set  $A$  is either very 'close' to some extremal set  $B \in \mathcal{B}_n$  or it contains 'many' (a positive fraction of all) edges of  $\mathcal{H}_n$ . Note that in many natural settings, such a property does hold, for example, as a consequence of the Erdős-Simonovits stability theorem [14, 60] and the removal lemma for graphs.

We again start with the 'random' version of our stability result, which is a slight weakening of Theorem 3.4 of [55], see the discussion below Theorem 5.2.

**Theorem 6.2.** *Let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs, let  $\alpha \in (0, 1)$ , and let  $c$  and  $c'$  be positive constants. Let  $\mathbf{p}$  be a sequence of probabilities such that  $p_n^{k-1}e(\mathcal{H}_n) \geq c'v(\mathcal{H}_n)$  and, for every  $\ell \in [k - 1]$ ,*

$$\Delta_\ell(\mathcal{H}_n) \leq c \cdot \min \left\{ p_n^{\ell-k}, p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \right\}$$

and let  $\mathcal{B}$  be a sequence of sets with  $B_n \subseteq \mathcal{P}(V(\mathcal{H}_n))$ .

*If  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable, then the following holds. For every positive  $\delta$ , there exist  $\varepsilon$  and  $C$  such that if  $q_n \geq Cp_n$  and  $q_nv(\mathcal{H}_n) \rightarrow \infty$  as  $n \rightarrow \infty$ , then a.a.s. every independent set  $I \subseteq V(\mathcal{H}_n)_{q_n}$  with  $|I| \geq (\alpha - \varepsilon)q_nv(\mathcal{H}_n)$  satisfies  $|I \setminus B| < \delta q_nv(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ .*

The following theorem, a 'counting' analogue of Theorem 6.2, is our main stability result. A simple version of it, applicable to 3-uniform hypergraphs with  $\Delta_2(\mathcal{H}_n) = O(1)$ , was proved in [1] and used in [1, 2] to count sum-free subsets in Abelian groups and in the set  $[n]$ .

**Theorem 6.3.** *Let  $\mathcal{H}$  be a sequence of  $k$ -uniform hypergraphs, let  $\alpha \in (0, 1)$ , and let  $c$  and  $c'$  be positive constants. Let  $\mathbf{p}$  be a sequence of probabilities such that  $p_n^{k-1}e(\mathcal{H}_n) \geq c'v(\mathcal{H}_n)$  and, for every  $\ell \in [k - 1]$ ,*

$$\Delta_\ell(\mathcal{H}_n) \leq c \cdot \min \left\{ p_n^{\ell-k}, p_n^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \right\}$$

and let  $\mathcal{B}$  be a sequence of sets with  $B_n \subseteq \mathcal{P}(V(\mathcal{H}_n))$ .

If  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable, then the following holds. For every positive  $\delta$ , there exist  $\varepsilon$  and  $C$  such that if  $m \geq Cp_n v(\mathcal{H}_n)$ , then there are at most

$$(1 - \varepsilon)^m \binom{\alpha v(\mathcal{H}_n)}{m}$$

independent sets  $I \in \mathcal{I}(\mathcal{H}_n, m)$  such that  $|I \setminus B| \geq \delta m$  for every  $B \in \mathcal{B}_n$ .

*Proof of Theorem 6.2.* The proof is similar to the proof of Theorem 5.2. Let  $k \in \mathbb{N}$ ,  $\alpha \in (0, 1)$ ,  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$ , and  $\mathcal{H}$  and  $\mathcal{B}$  be as in the statement of Theorem 6.2. Furthermore, suppose that  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable and fix some small positive  $\delta$ . Let  $\varepsilon$  be a small positive constant, let  $n$  be sufficiently large, let  $\delta' = \delta/3$  and  $\varepsilon' = 3\varepsilon$ , and set

$$\mathcal{F} = \{A \subseteq V(\mathcal{H}_n) : |A| \geq (\alpha - \varepsilon')v(\mathcal{H}_n) \text{ and } |A \setminus B| \geq \delta'v(\mathcal{H}_n) \text{ for every } B \in \mathcal{B}_n\}.$$

Since  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable, it follows that  $\mathcal{H}_n$  is  $(\mathcal{F}, \varepsilon)$ -dense, provided that  $\varepsilon$  is sufficiently small. Let  $C' = C_{2.2}(k, \varepsilon, c, c')$ . By Theorem 2.2, there exist a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H}_n)}{\leq C' p_n v(\mathcal{H}_n)}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}_n) \rightarrow \mathcal{S}$  such that

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

for every  $I \in \mathcal{I}(\mathcal{H}_n)$ . Let  $C = C'/\varepsilon^3$  and assume that  $q_n \geq Cp_n$ . Let  $m = (\alpha - \varepsilon)q_n v(\mathcal{H}_n)$  and, for the sake of brevity, let us write  $V = V(\mathcal{H}_n)$  and  $q = q_n$ . Let

$$\mathcal{I}' = \{I \in \mathcal{I}(\mathcal{H}_n) : |I| \geq m \text{ and } |I \setminus B| \geq \delta q v(\mathcal{H}_n) \text{ for every } B \in \mathcal{B}_n\}$$

and let  $\mathcal{A}$  denote the event that  $\mathcal{H}_n[V_q]$  contains an independent set  $I \in \mathcal{I}'$ . We are required to prove that  $\mathbb{P}(\mathcal{A})$  tends to 0 as  $n \rightarrow \infty$ .

Observe first that

$$\mathbb{P}(\mathcal{A}) \leq \sum_{S \in \mathcal{S}} \mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}' \text{ such that } g(I) = S). \quad (27)$$

Fix an  $S \in \mathcal{S}$ , let  $\mathcal{I}'_S = \{I \in \mathcal{I}' : g(I) = S\}$ , and note for future reference that

$$|S| \leq C' p_n v(\mathcal{H}_n) \leq \varepsilon^3 q v(\mathcal{H}_n). \quad \underline{\text{Thm 2.2}} \quad (28)$$

We claim that

$$\mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}'_S) \leq \mathbb{P}(S \subseteq V_q) \cdot \exp\left(-\frac{\varepsilon^2 q v(\mathcal{H}_n)}{4}\right). \quad (29)$$

In order to prove (29), recall that since  $f(S) \in \overline{\mathcal{F}}$ , we either have  $|f(S)| < (\alpha - \varepsilon')v(\mathcal{H}_n)$  or  $|f(S) \setminus B| < \delta'v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ . We therefore consider two cases.

**Case 1:**  $|f(S)| < (\alpha - \varepsilon')v(\mathcal{H}_n)$ .

We bound the left-hand side of (29) as follows:

$$\mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}'_S) \leq \mathbb{P}(S \subseteq V_q) \cdot \mathbb{P}(|V_q \cap f(S)| \geq m - |S|). \quad (30)$$

In order to justify the above inequality, note that for every  $I \in \mathcal{I}'_S$ , we have  $I \setminus S \subseteq f(S)$ . Recall that  $\varepsilon' = 3\varepsilon$ . Since  $m - |S| \geq (\alpha - 2\varepsilon)q v(\mathcal{H}_n)$ , by (28), and

$$\mathbb{E}[|V_q \cap f(S)|] \leq (\alpha - \varepsilon')q v(\mathcal{H}_n) = (\alpha - 3\varepsilon)q v(\mathcal{H}_n),$$

then by Chernoff's inequality we have

$$\mathbb{P}(|V_q \cap f(S)| \geq m - |S|) \leq \exp\left(-\frac{\varepsilon^2 qv(\mathcal{H}_n)}{4}\right). \quad (31)$$

Combining (30) and (31), we obtain (29), as required.

**Case 2:**  $|f(S) \setminus B| < \delta'v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ .

We estimate the left-hand side of (29) as follows:

$$|I \setminus B| \geq \delta qv$$

$$\mathbb{P}(I \subseteq V_q \text{ for some } I \in \mathcal{I}'_S) \leq \mathbb{P}(S \subseteq V_q) \cdot \mathbb{P}(|V_q \cap (f(S) \setminus B)| \geq \delta qv(\mathcal{H}_n) - |S|).$$

This follows from the definition of  $\mathcal{I}'$  and the fact that  $I \setminus S \subseteq f(S)$  for every  $I \in \mathcal{I}'_S$ . Since  $|f(S) \setminus B| < \delta'v(\mathcal{H}_n)$ , we have

$$\mathbb{E}[|V_q \cap (f(S) \setminus B)|] < \delta'qv(\mathcal{H}_n),$$

whereas  $\delta qv(\mathcal{H}_n) - |S| \geq 2\delta'qv(\mathcal{H}_n)$  by (28) and since  $\delta = 3\delta'$ . By Chernoff's inequality, it follows that

$$\mathbb{P}(|V_q \cap (f(S) \setminus B)| \geq 3\delta'qv(\mathcal{H}_n) - |S|) \leq \exp\left(-\frac{(\delta')^2 qv(\mathcal{H}_n)}{4}\right) \leq \exp\left(-\frac{\varepsilon^2 qv(\mathcal{H}_n)}{4}\right)$$

since  $\varepsilon$  was chosen sufficiently small. Thus (29) follows in this case as well.

Finally, note that, since  $|S| \leq \varepsilon^3 qv(\mathcal{H}_n)$  for every  $S \in \mathcal{S}$ , as in (24), we have

$$\sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq V_q) \leq \sum_{s=0}^{\varepsilon^3 qv(\mathcal{H}_n)} \binom{v(\mathcal{H}_n)}{s} q^s \leq v(\mathcal{H}_n) \cdot \left(\frac{e}{\varepsilon^3}\right)^{\varepsilon^3 qv(\mathcal{H}_n)}. \quad (32)$$

Putting (27), (29), and (32) together, we obtain

$$\mathbb{P}(\mathcal{A}) \leq \sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq V_q) \exp\left(-\frac{\varepsilon^2 qv(\mathcal{H}_n)}{4}\right) \leq \exp(-\varepsilon^3 qv(\mathcal{H}_n)),$$

as required.  $\square$

*Proof of Theorem 6.3.* Let  $k \in \mathbb{N}$ ,  $\alpha \in (0, 1)$ ,  $\mathbf{p} \in [0, 1]^{\mathbb{N}}$ , and  $\mathcal{H}$  and  $\mathcal{B}$  be as in the statement of Theorem 6.3. Furthermore, suppose that  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable and fix some positive  $\delta$ . Let  $\delta'$  be a sufficiently small positive constant (depending only on  $\alpha$  and  $\delta$ ), let  $\varepsilon$  be a small positive constant, and let  $n$  be sufficiently large. Let  $\varepsilon' = 2\varepsilon$ , and set

$$\mathcal{F} = \{A \subseteq V(\mathcal{H}_n) : |A| \geq (\alpha - \varepsilon')v(\mathcal{H}_n) \text{ and } |A \setminus B| \geq \delta'v(\mathcal{H}_n) \text{ for every } B \in \mathcal{B}_n\}.$$

Since  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable, it follows that  $\mathcal{H}_n$  is  $(\mathcal{F}, \varepsilon)$ -dense, provided that  $\varepsilon$  is sufficiently small (as a function of  $\delta'$ ). Let  $C' = C_{2.2}(k, \varepsilon, c, c')$ . By Theorem 2.2, there exist a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H}_n)}{\leq C' p_n v(\mathcal{H}_n)}$  and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}_n) \rightarrow \mathcal{S}$  such that

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

for every  $I \in \mathcal{I}(\mathcal{H}_n)$ . Let  $C = C'/\varepsilon^2$ , assume that  $m \geq Cp_n v(\mathcal{H}_n)$ , and set

$$\mathcal{I}' = \{I \in \mathcal{I}(\mathcal{H}_n, m) : |I \setminus B| \geq \delta m \text{ for every } B \in \mathcal{B}_n\}.$$

*Handwritten notes:*  
 $\mathcal{I}' \subseteq \mathcal{I}$   
 $\mathcal{S} \subseteq \mathcal{I}'$

Our task is to bound the size of  $\mathcal{I}'$  from above. To this end, fix an  $S \in \mathcal{S}$  and let  $\mathcal{I}'_S = \{I \in \mathcal{I}' : g(I) = S\}$ . Note for future reference that

$$|S| \leq C' p_n v(\mathcal{H}_n) \leq \varepsilon^2 m. \quad (33)$$

Since  $f(S) \in \overline{\mathcal{F}}$ , we either have  $|f(S)| < (\alpha - \varepsilon')v(\mathcal{H}_n)$  or  $|f(S) \setminus B| < \delta'v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ . We therefore consider two cases.

**Case 1:**  $|f(S)| < (\alpha - \varepsilon')v(\mathcal{H}_n)$ .

We claim that in this case

$$\binom{v(\mathcal{H}_n)}{|S|} \cdot |\mathcal{I}'_S| \leq \frac{(1 - \varepsilon)^m}{2m} \binom{\alpha v(\mathcal{H}_n)}{m}. \quad (34)$$

To prove (34), we first estimate the size of  $\mathcal{I}'_S$  as follows:

$$|\mathcal{I}'_S| \leq \binom{|f(S)|}{m - |S|} \leq \binom{(\alpha - \varepsilon')v(\mathcal{H}_n)}{m - |S|}.$$

The above inequality follows since  $I \setminus S \subseteq f(S)$  for every  $I \in \mathcal{I}'_S$ . Observe that if  $m \geq (\alpha - \varepsilon')v(\mathcal{H}_n)$ , then  $\mathcal{I}' \subseteq \mathcal{F}$  and hence  $\mathcal{I}' = \emptyset$ , since  $\mathcal{H}_n$  is  $(\mathcal{F}, \varepsilon)$ -dense. Therefore, from now on let us assume that  $m < (\alpha - \varepsilon')v(\mathcal{H}_n) = (\alpha - 2\varepsilon)v(\mathcal{H}_n)$ . It follows from (15) and (16), as in (26), that

$$\binom{(\alpha - \varepsilon')v(\mathcal{H}_n)}{m - |S|} \leq \binom{m}{2\varepsilon v(\mathcal{H}_n)}^{|S|} \left(\frac{\alpha - \varepsilon'}{\alpha}\right)^{m - |S|} \binom{\alpha v(\mathcal{H}_n)}{m} \leq \binom{m}{(\alpha - \varepsilon')v - m}^{|S|} \binom{(\alpha - \varepsilon')v}{m}$$

and hence, since  $|S| \leq \varepsilon^2 m$  and  $\varepsilon' = 2\varepsilon$ ,

$$\binom{v(\mathcal{H}_n)}{|S|} |\mathcal{I}'_S| \leq \left(\frac{e}{|S|} \cdot \frac{m}{2\varepsilon}\right)^{|S|} (1 - \varepsilon')^{m - |S|} \binom{\alpha v(\mathcal{H}_n)}{m} \leq \frac{(1 - \varepsilon)^m}{2m} \binom{\alpha v(\mathcal{H}_n)}{m},$$

as claimed.

**Case 2:**  $|f(S) \setminus B| < \delta'v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ .

We claim that in this case

$$\binom{v(\mathcal{H}_n)}{|S|} \cdot |\mathcal{I}'_S| \leq \delta^m \binom{\alpha v(\mathcal{H}_n)}{m}. \quad (35)$$

To prove (35), we first estimate the size of  $\mathcal{I}'_S$  as follows:

$$|\mathcal{I}'_S| \leq \binom{|f(S) \setminus B|}{\delta m - |S|} \binom{|f(S)|}{m - \delta m} \leq \binom{\delta'v(\mathcal{H}_n)}{\delta m - |S|} \binom{v(\mathcal{H}_n)}{m - \delta m}. \quad (36)$$

To see the first inequality, recall that every  $I \in \mathcal{I}'_S$  contains at least  $\delta m - |S|$  elements of  $f(S) \setminus B$  for every  $B \in \mathcal{B}_n$ . Recall that  $|S| \leq \varepsilon^2 m$  and note that therefore, if  $m \geq (\alpha/2)v(\mathcal{H}_n)$ , then  $\delta m - |S| \geq \delta'v(\mathcal{H}_n)$  and hence  $\mathcal{I}'_S = \emptyset$ . Thus, we may assume that  $m < (\alpha/2)v(\mathcal{H}_n)$ . It follows, using (16) and (17), that

$$\binom{v(\mathcal{H}_n)}{m - \delta m} \leq \binom{m}{v(\mathcal{H}_n) - m}^{\delta m} \binom{v(\mathcal{H}_n)}{m} \leq \binom{2m}{v(\mathcal{H}_n)}^{\delta m} \left(\frac{2}{\alpha}\right)^m \binom{\alpha v(\mathcal{H}_n)}{m}. \quad (37)$$

Hence, by (33), (36), and (37), using (14), we have

$$\begin{aligned} \binom{v(\mathcal{H}_n)}{|S|} |\mathcal{I}'_S| &\leq \left( \frac{ev(\mathcal{H}_n)}{|S|} \right)^{|S|} \left( \frac{2e\delta'v(\mathcal{H}_n)}{\delta m} \right)^{\delta m - |S|} \left( \frac{2m}{v(\mathcal{H}_n)} \right)^{\delta m} \left( \frac{2}{\alpha} \right)^m \binom{\alpha v(\mathcal{H}_n)}{m} \\ &\Rightarrow \left( \frac{1}{|S|} \cdot \frac{\delta m}{2\delta'} \right)^{|S|} \left( \frac{4e\delta'}{\delta} \right)^{\delta m} \left( \frac{2}{\alpha} \right)^m \binom{\alpha v(\mathcal{H}_n)}{m} \leq \delta^m \binom{\alpha v(\mathcal{H}_n)}{m}, \end{aligned}$$

as claimed, since  $|S| \leq \varepsilon^2 m$  and  $\delta'$  and  $\varepsilon$  were chosen to be sufficiently small. Indeed, note that (for this calculation, and assuming that  $\delta$  is sufficiently small)  $\delta' = \delta^3 \cdot (\delta\alpha/2e)^{1/\delta}$  and  $\varepsilon < \delta'$  suffice.

Finally, by (34) and (35), we obtain

$$|\mathcal{I}'| = \sum_{S \in \mathcal{S}} |\mathcal{I}'_S| \leq \sum_{s=0}^{\varepsilon^2 m} \binom{v(\mathcal{H}_n)}{s} \max\{|\mathcal{I}'_S| : |S| = s\} \leq (1 - \varepsilon)^m \binom{\alpha v(\mathcal{H}_n)}{m},$$

as claimed.  $\square$

## 7. TURÁN'S PROBLEM IN RANDOM GRAPHS

In this section, we shall deduce from Theorems 5.2 and 6.2 the sparse random analogues of the classical theorems of Erdős and Stone [18] and Turán [64] and of Erdős and Simonovits [14, 60], Theorems 1.3 and 1.4. In fact, we will prove a natural generalization of Theorem 1.3 to  $t$ -balanced  $t$ -uniform hypergraphs, Theorem 7.2 below, which was already proved by Conlon and Gowers [12] and Schacht [58]. We first recall the following generalization of the notion of 2-density of a graph to  $t$ -uniform hypergraphs.

**Definition 7.1.** Let  $H$  be a  $t$ -uniform hypergraph with at least  $t + 1$  vertices. We define the  $t$ -density of  $H$ , denoted by  $m_t(H)$ , by

$$m_t(H) = \max \left\{ \frac{e(H') - 1}{v(H') - t} : H' \subseteq H \text{ with } v(H') \geq t + 1 \right\}.$$

Moreover, we say that  $H$  is  $t$ -balanced if  $m_t(H') \leq m_t(H)$  for all  $H' \subseteq H$ .

We also recall that the *Turán density* of a  $t$ -uniform hypergraph  $H$ , denoted  $\pi(H)$ , is defined by

$$\pi(H) = \lim_{n \rightarrow \infty} \frac{\text{ex}(K_n^t, H)}{\binom{n}{t}}, \quad (38)$$

where, as usual,  $\text{ex}(K_n^t, H)$  is the Turán number for  $H$ , that is, the maximum number of edges in an  $H$ -free  $t$ -uniform hypergraph with  $n$  vertices.

**Theorem 7.2.** For every  $t$ -balanced  $t$ -uniform hypergraph  $H$  with  $\Delta(H) \geq 2$  and every positive  $\delta$ , there exists a positive constant  $C$  such that if  $q_n \geq Cn^{-1/m_t(H)}$ , then

$$\mathbb{P} \left( \text{ex}(\underbrace{G^t(n, q_n)}_{\text{Random subgraph } \mathcal{G} \left( \binom{[n]}{t} \right)}, H) \leq (\pi(H) + \delta) q_n \binom{n}{t} \right) \rightarrow 1$$

as  $n \rightarrow \infty$ .

Random subgraph  $\mathcal{G} \left( \binom{[n]}{t} \right)$  where each edge is included with probability  $q_n$ .



Once again, we emphasize that we actually obtain essentially optimal bounds on the probability in the above statement, i.e., bounds of the form  $1 - \exp(-bq_n n^t)$  for some positive constant  $b$  that depends only on  $H$  and  $\delta$ .

Similarly as in [58] and [55], Theorems 7.2 and 1.4, and hence also Theorem 1.3, will follow easily from our general transference results, Theorems 5.2 and 6.2, and the classical supersaturation results of Erdős and Simonovits [17] (for Theorem 7.2) and the stability theorem of Erdős and Simonovits [14, 60] together with the so-called graph removal lemma (for Theorem 1.4). We only need to check that the hypergraph of copies of  $H$  in the complete hypergraph  $K_n^t$ , to which we would like to apply our transference theorems, satisfies the assumptions of Theorems 5.2 and 6.2. Since we are going to use this fact several times in this and later sections, we state it as a separate proposition. Let  $H$  be an arbitrary  $t$ -uniform hypergraph. The *hypergraph of copies of  $H$  in  $K_n^t$*  is the  $e(H)$ -uniform hypergraph on the vertex set  $E(K_n^t)$  whose edges are the edges sets of all copies of  $H$  in  $K_n^t$ .

**Proposition 7.3.** *Let  $n$  and  $t$  be integers with  $t \geq 2$  and let  $H$  be a  $t$ -balanced  $t$ -uniform hypergraph. Set  $k = e(H)$  and let  $\mathcal{H}$  be the  $k$ -uniform hypergraph of copies of  $H$  in  $K_n^t$ . There exist positive constants  $c$  and  $c'$  such that, letting  $p = n^{-1/m_t(H)}$ , the following holds:*

$e(\mathcal{H})$

- (a)  $p^{e(H)-1} e(\mathcal{H}) \geq c' v(\mathcal{H})$ .  $e(H) = k$   
 (b)  $\Delta_\ell(\mathcal{H}) \leq c \cdot \min \left\{ p^{\ell-1}, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right\}$  for every  $\ell \in [k-1]$ .

$$v(\mathcal{H}) = \binom{\binom{n}{t}}{k}$$

*Proof.* To prove (a), note that  $v(\mathcal{H}) = \binom{n}{t} = \Theta(n^t)$  and that  $e(\mathcal{H}) = \frac{(v(H))!}{|\text{Aut}(H)|} \cdot \binom{n}{v(H)} = \Theta(n^{v(H)})$ . Since  $H$  is  $t$ -balanced, then

$$p = n^{-1/m_t(H)} = n^{-\frac{v(H)-t}{e(H)-1}}$$

and thus

$$p^{e(H)-1} \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} \geq c' \cdot n^{-(v(H)-t)} \cdot \frac{n^{v(H)}}{n^t} = c' \quad (39)$$

for some positive constant  $c'$ , as required.

To prove (b), observe first that, for each  $\ell \in [k-1]$ ,

$$\Delta_\ell(\mathcal{H}) \leq c'' \cdot \max \left\{ n^{v(H)-v(H')} : H' \subseteq H \text{ with } e(H') = \ell \right\}$$

for some positive constant  $c''$ . Note that for every  $H' \subseteq H$ , we have  $m_t(H') \leq m_t(H)$  by our assumption that  $H$  is  $t$ -balanced and  $e(H') - 1 \leq m_t(H') \cdot (v(H') - t)$  by the definition of  $t$ -density. Moreover,  $p^{e(H)} n^{v(H)} = p n^t$ , by the choice of  $p$ . Thus,

$$\begin{aligned} \Delta_\ell(\mathcal{H}) \cdot p^{e(H)-\ell} &\leq c'' \cdot \max_{H' \subseteq H: e(H')=\ell} \left( \frac{p^{e(H)} n^{v(H)}}{p^{e(H')} n^{v(H')}} \right) = c'' \cdot \max_{H' \subseteq H: e(H')=\ell} \left( \frac{p^{1-e(H')}}{n^{v(H')-t}} \right) \\ &\leq c'' \cdot \max_{H' \subseteq H: e(H')=\ell} \left( \frac{(n^{1/m_t(H)})^{m_t(H') \cdot (v(H')-t)}}{n^{v(H')-t}} \right) \leq c'', \end{aligned} \quad (40)$$

$$m_t(H') \leq m_t(H)$$

$$v(H') \geq t$$

[17]: If  $L$  is a  $t$ -uniform hypergraph and  
 $H$  is a  $t$ -uniform hypergraph with  $n$  vertices  
and  $e(H) \geq ex(n, L) + c \cdot n^t$   
then  $H$  contains  $\underline{c_2 n^t}$  copies of  $L$ .

$V(H_n) = \binom{[n]}{t}$  and  $E(H_n)$  corr. to copies of  $H$   
 $\uparrow$   
 $H$  [11]  $\Rightarrow$   $\alpha$ -dense.

---

Thm: E2S

$\forall \delta > 0, c(H) > 0 \exists \epsilon$  s.t. every  $n$ -vertex  $H$ -free graph  
with  $\geq ex(K_n, H) - \epsilon n^2$  edges can be made  $(X(H)-1)$ -partite  
by removing  $\leq \delta n^2$  edges.

Thm 2.3

$\forall H, \delta > 0, \exists \epsilon$  such that any graph on  $n$  vertices with  
 $\leq \epsilon n^{\nu(H)}$  copies of  $H$  can be made  $H$ -free by removing  $\leq \delta n^2$   
edges.

since  $m_t(H') \leq m_t(H)$  and  $v(H') \geq t$  for every non-empty  $H' \subseteq H$ . Moreover, since  $e(\mathcal{H})/v(\mathcal{H}) \geq c' \cdot n^{v(H)-t}$ , by (a), then

$$\begin{aligned} \Delta_\ell(\mathcal{H}) \cdot \left( p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right)^{-1} &\leq c'' \cdot \frac{v(\mathcal{H})}{e(\mathcal{H})} \cdot \max_{H' \subseteq H: e(H')=\ell} \left( \frac{n^{v(H)}}{p^{e(H')-1} n^{v(H')}} \right) \\ &\leq \frac{c''}{c'} \cdot \max_{H' \subseteq H: e(H')=\ell} \left( \frac{p^{1-e(H')}}{n^{v(H')-t}} \right) \leq \frac{c''}{c'}, \end{aligned}$$

where the last inequality follows as in (40). Finally, we let  $c = \max\{c''/c', c''\}$ .  $\square$

*Proof of Theorem 7.2.* Let  $H$  be a  $t$ -balanced  $t$ -uniform hypergraph, let  $k = e(H)$ , and let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the sequence of  $k$ -uniform hypergraphs of copies of  $H$  in  $K_n^t$ . Let  $\alpha = \pi(H)$ , let  $\delta$  be a positive constant, and let  $p_n = n^{-1/m_t(H)}$ . It follows easily from the supersaturation theorem of Erdős and Simonovits [17] that  $\mathcal{H}$  is  $\alpha$ -dense, see [58]. Let  $C = C_{5.2}(\mathcal{H}, \delta)$  and assume that  $q_n \geq Cp_n = Cn^{-1/m_t(H)}$ . Note that the assumption that  $H$  contains a vertex of degree at least 2 implies that  $m_t(H) > 1/t$  and hence  $q_n v(\mathcal{H}_n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Together with Proposition 7.3, this implies that  $\mathcal{H}$  satisfies the assumptions of Theorem 5.2 and hence with probability tending to 1 as  $n \rightarrow \infty$ ,

$$\text{ex}(G^t(n, q_n), H) = \alpha(\mathcal{H}_n[E(G^t(n, q_n))]) \leq (\pi(H) + \delta)q_n \binom{n}{t},$$

as required.  $\square$

In the proof of Theorems 1.4 and 1.6, we shall need the following proposition, which is a fairly straightforward consequence of the Erdős-Simonovits stability theorem [14, 60] and the graph removal lemma [15]. A proof of this statement can be found in [55]. We remark that a new proof of the graph removal lemma, which avoids the use of the Szemerédi regularity lemma, was given recently by Fox [19].

**Proposition 7.4.** *Let  $H$  be an arbitrary graph. For every positive  $\delta$ , there exists a positive  $\varepsilon$  such that the following holds for every  $n \in \mathbb{N}$ . If  $G$  is an  $n$ -vertex graph with*

$$e(G) \geq \left( 1 - \frac{1}{\chi(H) - 1} - \varepsilon \right) \binom{n}{2},$$

*then either  $G$  may be made  $(\chi(H) - 1)$ -partite by removing from it at most  $\delta n^2$  edges or  $G$  contains at least  $\varepsilon n^{v(H)}$  copies of  $H$ .*

*Proof of Theorem 1.4.* Let  $H$  be a 2-balanced graph, let  $k = e(H)$ , and let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the sequence of  $k$ -uniform hypergraphs of copies of  $H$  in  $K_n$ . Let  $\alpha = \pi(H) = \left( 1 - \frac{1}{\chi(H)-1} \right)$ , let  $\delta$  be a positive constant, and let  $p_n = n^{-1/m_2(H)}$ . Moreover, let  $\mathcal{B}_n$  be the family of all complete  $(\chi(H) - 1)$ -partite subgraphs of  $K_n$ . By Proposition 7.4,  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable. Let  $C = C_{6.2}(\mathcal{H}, \delta)$ , let  $\varepsilon = \varepsilon_{6.2}(\mathcal{H}, \delta)$ , and assume that  $q_n \geq Cp_n = Cn^{-1/m_2(H)}$ . Note that the assumption that  $H$  contains a vertex of degree at least 2 implies that  $m_2(H) > 1/2$  and hence  $q_n v(\mathcal{H}_n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Together with Proposition 7.3, the discussion above implies that  $\mathcal{H}$  satisfies the assumptions of Theorem 6.2 and hence with probability tending to 1 as  $n \rightarrow \infty$ , every  $H$ -free subgraph of  $G' \subseteq G(n, q_n)$  satisfies  $|G' \setminus B| \leq \delta q_n v(\mathcal{H}_n)$  for some  $B \in \mathcal{B}_n$ . In other words, with probability tending to 1 as  $n \rightarrow \infty$ , every  $H$ -free subgraph

$V(\mathcal{H}_n) = \binom{[n]}{2}$   
 $\mathcal{B} \in \mathcal{B}_n \Rightarrow$   
 $\mathcal{B} \subseteq V(\mathcal{H}_n)$

of  $G(n, q_n)$  can be made  $(\chi(H) - 1)$ -partite by removing from it at most  $\delta q_n \binom{n}{2}$  edges, as required.  $\square$

## 8. THE TYPICAL STRUCTURE OF $H$ -FREE GRAPHS

In this section, we shall deduce from Theorems 5.4 and 6.3 the sparse analogue of the theorem of Erdős, Frankl, and Rödl [15], Theorem 1.5, and an approximate sparse analogue of the result of Erdős, Kleitman, and Rothschild [16], Theorem 1.6. We stress once again that neither proof employs Szemerédi's regularity lemma. In order to prove Theorem 1.5, we are actually going to prove the following natural generalization of it to  $t$ -balanced  $t$ -uniform hypergraphs. Generalizing the definition stated in Section 1.3, given integers  $n$  and  $m$  with  $0 \leq m \leq \binom{n}{t}$  and a  $t$ -uniform hypergraph  $H$ , let us denote by  $f_{n,m}(H)$  the number of  $H$ -free  $t$ -uniform hypergraphs on the vertex set  $[n]$  that have exactly  $m$  edges.

**Theorem 8.1.** *Let  $H$  be a  $t$ -balanced  $t$ -uniform hypergraph and let  $\delta$  be a positive constant. There exists a constant  $C$  such that for every  $n$ , if  $m \geq Cn^{t-1/m_t(H)}$ , then*

$$\binom{\text{ex}(n, H)}{m} \leq f_{n,m}(H) \leq \binom{\text{ex}(n, H) + \delta n^t}{m}.$$

We remark that Theorem 8.1 refines a result of Nagle, Rödl, and Schacht [47], who, using the hypergraph regularity lemma, generalized (4) to  $t$ -uniform hypergraphs.

*Proof of Theorem 8.1.* Let  $H$  be a  $t$ -balanced  $t$ -uniform hypergraph, let  $k = e(H)$ , and let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the sequence of  $k$ -uniform hypergraphs of copies of  $H$  in  $K_n^t$ . Let  $\alpha = \pi(H)$ , see (38), let  $\delta$  be a positive constant, and let  $p_n = n^{-1/m_t(H)}$ . It follows easily from the supersaturation theorem of Erdős and Simonovits [17] that  $\mathcal{H}$  is  $\alpha$ -dense, see [58]. Let  $C = C_{5.4}(\mathcal{H}, \delta)$  and assume that  $m \geq Cn^{t-1/m_t(H)} \geq Cp_n v(\mathcal{H}_n)$ . Note that Proposition 7.3 implies that  $\mathcal{H}$  satisfies the assumptions of Theorem 5.4 and hence

$$f_{n,m}(H) = |I(\mathcal{H}_n, m)| \leq \binom{(\pi(H) + \delta) \binom{n}{t}}{m} \leq \binom{\text{ex}(n, H) + \delta n^t}{m},$$

as required. The claimed lower bound on  $f_{n,m}(H)$  is trivial.  $\square$

*Proof of Theorem 1.6.* Let  $H$  be a 2-balanced graph, let  $k = e(H)$ , and let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the sequence of  $k$ -uniform hypergraphs of copies of  $H$  in  $K_n$ . Let  $\alpha = \pi(H) = \left(1 - \frac{1}{\chi(H)-1}\right)$ , let  $\delta$  be a positive constant, and let  $p_n = n^{-1/m_2(H)}$ . Moreover, let  $\mathcal{B}_n$  be the family of all complete  $(\chi(H) - 1)$ -partite subgraphs of  $K_n$ . By Proposition 7.4,  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -stable. Let  $C = C_{6.3}(\mathcal{H}, \delta)$ , let  $\varepsilon = \varepsilon_{6.3}(\mathcal{H}, \delta)$ , and assume that  $m \geq Cn^{2-1/m_2(H)} \geq Cp_n v(\mathcal{H}_n)$ . Together with Proposition 7.3, this implies that  $\mathcal{H}$  satisfies the assumptions of Theorem 6.3 and hence, letting  $f_{n,m}^\delta(H)$  denote the number of  $H$ -free graphs on the vertex set  $[n]$  that have exactly  $m$  edges and that are not  $(\delta, \chi(H) - 1)$ -partite,

$$f_{n,m}^\delta(H) \leq (1 - \varepsilon)^m \binom{\pi(H) \binom{n}{2}}{m}.$$

Finally, note that (trivially),

$$f_{n,m}(H) \geq \binom{\pi(H) \binom{n}{2}}{m}$$

and hence  $f_{n,m}^\delta(H) = o(f_{n,m}(H))$ , as claimed.  $\square$

For  $t$ -uniform hypergraphs there is no general stability theorem known; however, such results have been proved for a few specific hypergraphs (see [22, 23, 35, 36]), and in each case we obtain a corresponding result for sparse hypergraphs. For example, following [7], let  $F_5$  denote the ‘3-uniform triangle’, i.e., the hypergraph with edge set isomorphic to  $\{123, 124, 345\}$ , and say that a 3-uniform hypergraph is *triangle-free* if it contains no copy of  $F_5$ . The following theorem follows easily, as above, from Theorem 6.3 combined with the hypergraph removal lemma of Gowers [31] and Rödl and Skokan [53] and the stability theorem for 3-uniform triangle-free hypergraphs, which was proved by Keevash and Mubayi [35].

**Theorem 8.2.** *For every positive  $\delta$ , there exists a constant  $C$  such that the following holds. If  $m \geq Cn^2$ , then almost every triangle-free 3-uniform hypergraph with  $n$  vertices and  $m$  edges can be made tripartite by removing from it at most  $\delta m$  edges.*

*Proof.* Let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the sequence of 3-uniform hypergraphs of copies of  $F_5$  in  $K_n^3$ , set  $\alpha = 2/9$ , and let  $\mathcal{B}_n$  denote the collection of all complete tripartite subhypergraphs of  $K_n^3$ . By the hypergraph removal lemma [53, Theorem 1.3], combined with the stability theorem for triangle-free 3-uniform hypergraphs [35, Theorem 1.6], it follows that  $\mathcal{H}$  is  $(\alpha, \mathcal{B})$ -dense. Note that  $F_5$  is 3-balanced (though not strictly balanced); it follows by Proposition 7.3 that  $\mathcal{H}$  satisfies the conditions of Theorem 6.3 with  $p_n = n^{-1}$ . Hence the number of triangle-free 3-uniform hypergraphs with  $n$  vertices and  $m$  edges that cannot be made tripartite by removing at most  $\delta m$  edges is at most

$$(1 - \varepsilon)^m \binom{\pi(F_5) \binom{n}{3}}{m},$$

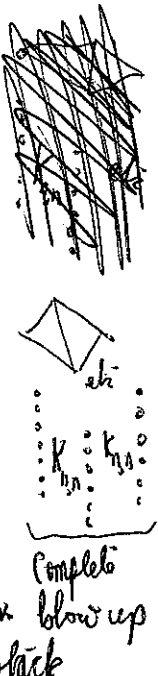
which easily implies the theorem.  $\square$

Finally, we remark that Theorem 8.2 can be seen as an approximate sparse analogue of a result of Balogh and Mubayi [7], who used the hypergraph regularity lemma and [35, Theorem 1.6] to show that almost all triangle-free 3-uniform hypergraphs are tripartite. For similar results for other forbidden hypergraphs, see [8] and [49].

### 9. THE KLR CONJECTURE

In this section, we shall deduce from Theorem 2.2 the KLR conjecture for 2-balanced graphs, Theorem 1.8. As in the preceding sections, the proof will be a fairly straightforward application of Theorem 2.2 to an appropriately defined hypergraph  $\mathcal{H}$  and family  $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$ . Let  $H$  be an arbitrary 2-balanced graph and let  $\mathcal{H}$  be the  $e(H)$ -uniform hypergraph of canonical copies of  $H$  in the complete blow-up of  $H$ . Defining an appropriate family  $\mathcal{F}$  and showing that  $\mathcal{H}$  is  $(\mathcal{F}, \varepsilon)$ -dense will require some work.

Given a graph  $H$  and integers  $n_1, \dots, n_{v(H)}$ , let us denote by  $\mathcal{G}(H; n_1, \dots, n_{v(H)})$  the collection of all graphs  $G$  constructed in the following way. The vertex set of  $G$  is a disjoint union  $V_1 \cup \dots \cup V_{v(H)}$  of sets of sizes  $n_1, \dots, n_{v(H)}$ , respectively, one for each vertex of  $H$ . The only edges of  $G$  lie between those pairs of sets  $(V_i, V_j)$  such that  $\{i, j\}$  is an edge of  $H$ . Recall the definition of  $\mathcal{G}(H, n, m, p, \varepsilon)$  from Section 1.4 and observe that  $\mathcal{G}(H, n, m, p, \varepsilon) \subseteq \mathcal{G}(H; n, \dots, n)$  for all  $m, p$ , and  $\varepsilon$ .



The following lemma, which is a robust version of the embedding lemma, stated in Section 1.4, suggests the right choice of  $\mathcal{F}$ . We remark that the lemma is well-known, but for completeness we provide a short proof of it in the form we shall use.

**Lemma 9.1.** *Let  $H$  be an arbitrary graph and let  $\delta: (0, 1] \rightarrow (0, 1)$  be an arbitrary function. There exist positive constants  $\alpha_0$ ,  $\xi$ , and  $N$  such that for every collection of integers  $n_1, \dots, n_{v(H)}$  satisfying  $n_1, \dots, n_{v(H)} \geq N$  and every graph  $G \in \mathcal{G}(H; n_1, \dots, n_{v(H)})$ , one of the following holds:*

- (a)  $G$  contains at least  $\xi n_1 \dots n_{v(H)}$  canonical copies of  $H$ .
- (b) There exist a positive constant  $\alpha$  with  $\alpha \geq \alpha_0$ , an edge  $\{i, j\} \in E(H)$ , and sets  $A_i \subseteq V_i$ ,  $A_j \subseteq V_j$  such that  $|A_i| \geq \alpha n_i$ ,  $|A_j| \geq \alpha n_j$ , and  $d_G(A_i, A_j) < \delta(\alpha)$ .

*Proof.* We prove the lemma by induction on the number of vertices of  $H$ . If  $v(H) = 1$ , then (a) holds vacuously with  $\xi = 1$  for every choice of  $G$ . Let us then assume that  $v(H) \geq 2$ , let  $v_1$  be the first vertex of  $H$  (i.e., the vertex corresponding to the set  $V_1$  from the definition of the family of blow-ups of  $H$ ), set  $\tilde{H} = H - v_1$ , and let  $\alpha_1 = \frac{1}{v(H)}$ . Given a function  $\delta$ , define  $\tilde{\delta}$  by letting  $\tilde{\delta}(x) = \delta(\delta(\alpha_1) \cdot x)$  for each  $x \in [0, 1]$  and let  $\tilde{\alpha}_0$ ,  $\tilde{\xi}$ , and  $\tilde{N}$  be the constants obtained by invoking the inductive assumption with  $H$  replaced by  $\tilde{H}$  and  $\delta$  replaced by  $\tilde{\delta}$ . Furthermore, let

$$\alpha_0 = \min \{ \alpha_1, \tilde{\alpha}_0 \cdot \delta(\alpha_1) \}, \quad N = \frac{\tilde{N}}{\delta(\alpha_1)}, \quad \text{and} \quad \xi = \alpha_1 \cdot \delta(\alpha_1)^{v(H)-1} \cdot \tilde{\xi}.$$

Now, let  $n_1, \dots, n_{v(H)}$  be integers satisfying  $n_1, \dots, n_{v(H)} \geq N$  and let  $G$  be an arbitrary graph from  $\mathcal{G}(H; n_1, \dots, n_{v(H)})$ . Suppose first that for some  $v_j \in N_H(v_1)$ , the set  $W_{1,j}$  defined by

$$W_{1,j} = \{ w \in V_1 : \deg_G(w, V_j) < \delta(\alpha_1) n_j \}$$

contains at least  $\alpha_1 n_1$  vertices. In this case, it is not hard to see that (b) is satisfied with  $\alpha = \alpha_1$ ,  $i = 1$ ,  $A_1 = W_{1,j}$ , and  $A_j = V_j$ . Hence, since  $\alpha_1 = \frac{1}{v(H)}$ , we may assume that the set  $W_1$  defined by

$$W_1 = V_1 \setminus \bigcup_{j \in N_H(v_1)} W_{1,j} = \{ w \in V_1 : \deg_G(w, V_j) \geq \delta(\alpha_1) n_j \text{ for all } v_j \in N_H(v_1) \} \quad |W_1| \geq (1 - (v-1)\alpha_1) n_1$$

contains at least  $\alpha_1 n_1$  vertices. For each  $w \in W_1$ , let  $G_w$  be the subgraph of  $G$  induced by the set  $V_2(w) \cup \dots \cup V_{v(H)}(w)$ , where for each  $j \in \{2, \dots, v(H)\}$ ,

$$V_j(w) = \begin{cases} V_j \cap N_G(w) & \text{if } v_j \in N_H(v_1), \\ V_j & \text{if } v_j \notin N_H(v_1). \end{cases}$$

Note that the assumption that  $w \in W_1$  implies that  $|V_j(w)| \geq \delta(\alpha_1) n_j \geq \tilde{N}$  for each  $j$ . Hence, we may apply the induction hypothesis to each graph  $G_w$ .

Suppose first that for some  $w \in W_1$ , we obtain an  $\tilde{\alpha}$  with  $\tilde{\alpha} \geq \tilde{\alpha}_0$ , vertices  $i, j \in V(\tilde{H})$ , and a pair  $A_i \subseteq V_i(w) \subseteq V_i$ ,  $A_j \subseteq V_j(w) \subseteq V_j$  such that

$$|A_\ell| \geq \tilde{\alpha} |V_\ell(w)| \geq \tilde{\alpha} \delta(\alpha_1) n_\ell \quad \text{for both } \ell \in \{i, j\}$$

and  $d_G(A_i, A_j) = d_{G_w}(A_i, A_j) < \tilde{\delta}(\tilde{\alpha}) = \delta(\tilde{\alpha}\delta(\alpha_1))$ . Then we are done, since (b) is satisfied with  $\alpha = \tilde{\alpha}\delta(\alpha_1)$ . Otherwise, for each  $w \in W_1$ , the number of canonical copies of  $\tilde{H}$  in the graph  $G_w$  is at least  $\tilde{\xi}|V_2(w)| \dots |V_{v(H)}(w)|$ , which is at least  $\tilde{\xi}\delta(\alpha_1)^{v(H)-1}n_2 \dots n_{v(H)}$ . Since  $w$  extends each of those copies to a canonical copy of  $H$  in  $G$ , it follows that in this case, the number of canonical copies of  $H$  in  $G$  is at least  $|W_1| \cdot \tilde{\xi}\delta(\alpha_1)^{v(H)-1}n_2 \dots n_{v(H)}$ , which is at least  $\xi n_1 \dots n_{v(H)}$ , as required.  $\square$

Our next lemma is more straightforward. It allows us to count  $(\varepsilon, p)$ -regular subgraphs of a graph that has a 'hole', as in Lemma 9.1(b). Recall that  $\mathcal{G}(K_2, n, m, p, \varepsilon)$  denotes the collection of all  $(\varepsilon, p)$ -regular bipartite graphs with  $m$  edges and  $n$  vertices in each part. Given such  $G$ , let  $V_1(G)$  and  $V_2(G)$  denote the two parts. For each  $\beta \in (0, 1)$ , define a function  $\delta: (0, 1] \rightarrow (0, 1)$  by setting

$$\delta(x) = \frac{1}{4e} \left( \frac{\beta}{2} \right)^{2/x^2} \quad (41)$$

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

**Lemma 9.2.** *For every positive  $\alpha_0$  and  $\beta$ , there exists a positive constant  $\varepsilon$  such that the following holds. Let  $\tilde{G} \subseteq K_{n,n}$  be such that there exist subsets  $A \subset V_1(\tilde{G})$  and  $B \subset V_2(\tilde{G})$  with*

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

for some  $\alpha \in [\alpha_0, 1]$ . Then, for every  $m$  with  $0 \leq m \leq n^2$ , there are at most

$$\beta^m \binom{n^2}{m}$$

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

*Proof.* We begin by noting that, by choosing random subsets of  $A$  and  $B$  if necessary, we may assume that  $|A| = |B| = \alpha n$ . Set  $\varepsilon = \min\{\alpha_0, 1/2\}$ , write  $\mathcal{G}^*$  for the family of all subgraphs of  $\tilde{G}$  that belong to  $\mathcal{G}(K_2, n, m, m/n^2, \varepsilon)$ , and let  $G \in \mathcal{G}^*$ . In particular,  $G$  is  $(\varepsilon, p)$ -regular, where  $p = m/n^2$  and since  $\varepsilon \leq \alpha$ , it follows that the pair  $(A, B)$  must have density at least  $(1 - \varepsilon)p$  in  $G$ . Thus, writing  $e_{\tilde{G}}(A, B)$  for the number of edges of  $\tilde{G}$  that lie between the sets  $A$  and  $B$ , since  $d_{\tilde{G}}(A, B) < \delta(\alpha)$ , then the number of choices for  $G$  can be estimated as follows:

$$|\mathcal{G}^*| \leq \sum_{\ell \geq (1-\varepsilon)p|A||B|} \binom{e_{\tilde{G}}(A, B)}{\ell} \binom{e(\tilde{G}) - e_{\tilde{G}}(A, B)}{m - \ell} \leq \sum_{\ell \geq \alpha^2 m/2} \binom{\delta(\alpha)\alpha^2 n^2}{\ell} \binom{n^2}{m - \ell}. \quad (42)$$

Note that the right-hand side of (42) is zero if  $m > 2\delta(\alpha)n^2$ , so we may assume that  $m \leq 2\delta(\alpha)n^2 \leq n^2/2$ . Thus, using (14) and (16), (42) implies that

$$|\mathcal{G}^*| \leq \sum_{\ell \geq \alpha^2 m/2} \left( \frac{e\delta(\alpha)\alpha^2 n^2}{\ell} \right)^\ell \left( \frac{m}{n^2 - m} \right)^\ell \binom{n^2}{m} \leq \sum_{\ell \geq \alpha^2 m/2} \left( \frac{2e\delta(\alpha)\alpha^2 m}{\ell} \right)^\ell \binom{n^2}{m}. \quad (43)$$

Since  $\delta(\alpha) < 1/4e$ , the summand in the right-hand side of (43) is decreasing in  $\ell$  on  $(\alpha^2 m/2, \infty)$  and hence

$$|\mathcal{G}^*| \leq m(4e\delta(\alpha))^{\alpha^2 m/2} \binom{n^2}{m} \leq \beta^m \binom{n^2}{m},$$

as required, since  $(4e\delta(\alpha))^{\alpha^2/2} = \beta/2$ .  $\square$

We can now easily deduce Theorem 1.8 from Theorem 2.2.

*Proof of Theorem 1.8.* Let  $H$  be a fixed 2-balanced graph, let  $n \in \mathbb{N}$ , and let  $H(n)$  be the largest graph in the family  $\mathcal{G}(H; n, \dots, n)$ , i.e., the complete blow-up of  $H$ , where each vertex of  $H$  is replaced by an independent set of size  $n$  and each edge of  $H$  is replaced by the complete bipartite graph  $K_{n,n}$ . Let  $\mathcal{H}$  be the  $e(H)$ -uniform hypergraph on the vertex set  $E(H(n))$  whose edges are all  $n^{v(H)}$  canonical copies of  $H$  in  $H(n)$ .

Fix an arbitrary positive constant  $\beta$ , let  $\delta: (0, 1] \rightarrow (0, 1)$  be the function defined in (41) with  $\beta$  replaced by  $\beta^2/4$ , i.e., set

$$\delta(x) = \frac{1}{4e} \left( \frac{\beta^2}{8} \right)^{2/x^2}$$

for each  $x \in (0, 1]$ , and let  $\alpha_0 = (\alpha_0)_{9.1}(H, \delta)$ ,  $\xi = \xi_{9.1}(H, \delta)$ , and  $N = N_{9.1}(H, \delta)$ . Let  $\mathcal{F}$  be the family of all subgraphs of  $H(n)$ , i.e., graphs in  $\mathcal{G}(H; n, \dots, n)$ , for which (b) in Lemma 9.1 is *not* satisfied. Clearly  $\mathcal{F}$  is an upset, and so, by Lemma 9.1,  $\mathcal{H}$  is  $(\mathcal{F}, \xi)$ -dense provided that  $n \geq N$ .

Now, since  $\mathcal{H}$  is contained in the hypergraph of all copies of  $H$  in the complete graph on  $v(H)n$  vertices and contains a positive proportion of those copies, it follows from Proposition 7.3 that  $\mathcal{H}$  satisfies the assumptions of Theorem 2.2 with  $p = n^{-1/m_2(H)}$  and  $\varepsilon = \xi$ , for some constants  $c$  and  $c'$  depending only on  $H$ . Therefore, there is a constant  $C'$ , a family  $\mathcal{S} \subseteq \binom{E(H(n))}{\leq C'n^{2-1/m_2(H)}}$ , and functions  $f: \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g: \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{S}$  such that

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

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

Let  $\varepsilon$  be a sufficiently small positive constant such that, in particular,  $\varepsilon \leq \varepsilon_{9.2}(\alpha_0, \beta^2/4)$ , let  $C = C'/\varepsilon$ , and suppose that  $m \geq Cn^{2-1/m_2(H)}$ . Let  $\mathcal{G}^* = \mathcal{G}^*(H, n, m, m/n^2, \varepsilon)$  and note that  $\mathcal{G}^* \subseteq \mathcal{I}(\mathcal{H})$ . We are required to bound from above the number of graphs in  $\mathcal{G}^*$ .

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

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

and let  $G_S = f(S)$ . For each  $\{u, w\} \in E(H)$ , let  $s(u, w) = e_S(V_u, V_w)$  and note that  $\sum_{\{u, w\} \in E(H)} s(u, w) = |S|$ . Since

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

then  $s(u, v) \leq \varepsilon m$  for every  $uv \in E(H)$ .

Now, since  $G_S \in \overline{\mathcal{F}}$ , it follows that there exist an  $\alpha \in [\alpha_0, 1]$ , an edge  $\{i, j\} \in E(H)$ , and sets  $A_i \subseteq V_i$ ,  $A_j \subseteq V_j$  such that  $|A_i|, |A_j| \geq \alpha n$  and  $d_{G_S}(A_i, A_j) < \delta(\alpha)$ . By Lemma 9.2, it



follows that there are at most

$$\left(\frac{\beta^2}{4}\right)^{m-s(i,j)} \binom{n^2}{m-s(i,j)}$$

choices for the edges between  $V_i$  and  $V_j$  such that  $G[V_i, V_j] \in \mathcal{G}(K_2, n, m, m/n^2, \varepsilon)$  and  $G[V_i, V_j] \subset G_S[V_i, V_j]$ . Since  $m - s(i, j) \geq m/2$ , it follows immediately that

$$|\mathcal{G}_S^*| \leq \left(\frac{\beta}{2}\right)^m \prod_{uv \in E(H)} \binom{n^2}{m-s(u,v)}.$$

Summing over sets  $S \in \mathcal{S}$ , and using (14) and (16), we obtain

$$\begin{aligned} |\mathcal{G}^*| &\leq \sum_{S \in \mathcal{S}} \left(\frac{\beta}{2}\right)^m \prod_{uv \in E(H)} \left(\frac{m}{n^2-m}\right)^{s(u,v)} \binom{n^2}{m} = \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{S \in \mathcal{S}} \left(\frac{m}{n^2-m}\right)^{|S|} \\ &\leq \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{s \leq \varepsilon m} \binom{e(H)n^2}{s} \left(\frac{2m}{n^2}\right)^s \leq \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{s \leq \varepsilon m} \left(\frac{2e \cdot e(H)m}{s}\right)^s. \end{aligned}$$

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

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

as required. □

**Acknowledgement.** The third author would like to thank Noga Alon and David Conlon for stimulating discussions. The authors would also like to thank David Conlon and Yoshiharu Kohayakawa for helpful comments on the manuscript.

REFERENCES

1. N. Alon, J. Balogh, R. Morris, and W. Samotij, *Counting sum-free sets in Abelian groups*, arXiv:1201.6654v1 [math.CO].
2. ———, *A refinement of the Cameron-Erdős Conjecture*, arXiv:1202.5200v2 [math.CO].
3. N. Alon and J.H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience Series in Discrete Math. and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős.
4. L. Babai, M. Simonovits, and J. Spencer, *Extremal subgraphs of random graphs*, J. Graph Theory 14 (1990), 599–622.
5. J. Balogh, B. Bollobás, and M. Simonovits, *The number of graphs without forbidden subgraphs*, J. Combin. Theory Ser. B 91 (2004), 1–24.
6. ———, *The typical structure of graphs without given excluded subgraphs*, Random Structures Algorithms 34 (2009), 305–318.
7. J. Balogh and D. Mubayi, *Almost all triangle-free triple systems are tripartite*, to appear in Combinatorica.
8. ———, *Almost all triple systems with independent neighborhoods are semi-bipartite*, J. Combin. Theory Ser. A 118 (2011), 1494–1518.
9. J. Balogh and W. Samotij, *The number of  $K_{m,m}$ -free graphs*, Combinatorica 31 (2011), 131–150.
10. ———, *The number of  $K_{s,t}$ -free graphs*, J. Lond. Math. Soc. (2) 83 (2011), 368–388.

11. M. Behrisch, *Random graphs without a short cycle*, Master's thesis, Humboldt-Universität zu Berlin, 2002.
12. D. Conlon and W.T. Gowers, *Combinatorial theorems in sparse random sets*, arXiv:1011.4310v1 [math.CO].
13. D. Conlon, W.T. Gowers, W. Samotij, and M. Schacht, *On the KLR conjecture in random graphs*, submitted.
14. P. Erdős, *Some recent results on extremal problems in graph theory. Results*, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York, 1967, pp. 117–123 (English); pp. 124–130 (French).
15. 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 Combin. **2** (1986), 113–121.
16. P. Erdős, D.J. Kleitman, and B.L. Rothschild, *Asymptotic enumeration of  $K_n$ -free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Accad. Naz. Lincei, Rome, 1976, pp. 19–27. Atti dei Convegni Lincei, No. 17.
17. P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192.
18. P. Erdős and A.H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.
19. J. Fox, *A new proof of the graph removal lemma*, Ann. of Math. **174** (2011), 561–579.
20. P. Frankl and V. Rödl, *Large triangle-free subgraphs in graphs without  $K_4$* , Graphs Combin. **2** (1986), 135–144.
21. Z. Füredi, *Random Ramsey graphs for the four-cycle*, Discrete Math. **126** (1994), 407–410.
22. Z. Füredi, O. Pikhurko, and M. Simonovits, *On triple systems with independent neighbourhoods*, Combin. Probab. Comput. **14** (2005), 795–813.
23. Z. Füredi and M. Simonovits, *Triple systems not containing a Fano configuration*, Combin. Probab. Comput. **14** (2005), 467–484.
24. H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.
25. H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291.
26. S. Gerke, *Random graphs with constraints*, 2005, Habilitationsschrift, Institut für Informatik, Technische Universität München.
27. S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse  $\epsilon$ -regularity*, J. Combin. Theory Ser. B **97** (2007), 34–56.
28. S. Gerke, H.J. Prömel, T. Schickinger, A. Steger, and A. Taraz,  *$K_4$ -free subgraphs of random graphs revisited*, Combinatorica **27** (2007), 329–365.
29. S. Gerke, T. Schickinger, and A. Steger,  *$K_5$ -free subgraphs of random graphs*, Random Structures Algorithms **24** (2004), 194–232.
30. S. Gerke and A. Steger, *The sparse regularity lemma and its applications*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 227–258.
31. W.T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. **166** (2007), 897–946.
32. M. Hamel and I. Laba, *Arithmetic structures in random sets*, Integers **8** (2008), A04, 21.
33. P.E. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: forbidding even cycles*, J. Combin. Theory Ser. B **64** (1995), 273–287.
34. ———, *Turán's extremal problem in random graphs: forbidding odd cycles*, Combinatorica **16** (1996), 107–122.
35. P. Keevash and D. Mubayi, *Stability theorems for cancellative hypergraphs*, J. Combin. Theory Ser. B **92** (2004), 163–175.
36. P. Keevash and B. Sudakov, *The Turán number of the Fano plane*, Combinatorica **25** (2005), 561–574.
37. D.J. Kleitman and K.J. Winston, *On the number of graphs without 4-cycles*, Discrete Math. **41** (1982), 167–172.

38. Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.
39. Y. Kohayakawa and B. Kreuter, *Threshold functions for asymmetric Ramsey properties involving cycles*, Random Structures Algorithms **11** (1997), 245–276.
40. Y. Kohayakawa, T. Łuczak, and V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), 133–163.
41. ———, *On  $K^4$ -free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.
42. Y. Kohayakawa and V. Rödl, *Regular pairs in sparse random graphs. I*, Random Structures Algorithms **22** (2003), 359–434.
43. Y. Kohayakawa, V. Rödl, and M. Schacht, *The Turán theorem for random graphs*, Combin. Probab. Comput. **13** (2004), 61–91.
44. Y. Kohayakawa, M. Schacht, and R. Spöhel, *Upper bounds on probability thresholds for asymmetric Ramsey properties*, to appear in Random Structures Algorithms.
45. T. Łuczak, *On triangle-free random graphs*, Random Structures Algorithms **16** (2000), 260–276.
46. M. Marcinişzyn, J. Skokan, R. Spöhel, and A. Steger, *Asymmetric Ramsey properties of random graphs involving cliques*, Random Structures Algorithms **34** (2009), 419–453.
47. B. Nagle, V. Rödl, and M. Schacht, *Extremal hypergraph problems and the regularity method*, Topics in discrete mathematics, Algorithms Combin., vol. 26, Springer, Berlin, 2006, pp. 247–278.
48. D. Osthus, H.J. Prömel, and A. Taraz, *For which densities are random triangle-free graphs almost surely bipartite?*, Combinatorica **23** (2003), 105–150, Paul Erdős and his mathematics (Budapest, 1999).
49. Y. Person and M. Schacht, *Almost all hypergraphs without Fano planes are bipartite*, Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms (Philadelphia, PA), SIAM, 2009, pp. 217–226.
50. F.P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. **30** (1930), 264–286.
51. V. Rödl and A. Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, pp. 317–346.
52. ———, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), 917–942.
53. V. Rödl and J. Skokan, *Applications of the regularity lemma for uniform hypergraphs*, Random Structures Algorithms **28** (2006), 180–194.
54. 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.
55. W. Samotij, *Stability results for random discrete structures*, arXiv:1111.6885v3 [math.CO].
56. A. Sárközy, *On difference sets of sequences of integers. III*, Acta Math. Acad. Sci. Hungar. **31** (1978), 355–386.
57. D. Saxton and A. Thomason, *Hypergraph containers*, submitted.
58. M. Schacht, *Extremal results for random discrete structures*, submitted.
59. A. Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combin. Probab. Comput. **20** (2011), 455–466.
60. M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 279–319.
61. T. Szabó and V.H. Vu, *Turán's theorem in sparse random graphs*, Random Structures Algorithms **23** (2003), 225–234.
62. E. Szemerédi, *On sets of integers containing no  $k$  elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
63. ———, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.
64. P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.
65. P. Varnavides, *On certain sets of positive density*, J. London Math. Soc. **34** (1959), 358–360.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS, 1409 W. GREEN STREET, URBANA, IL 61801

*E-mail address:* jobal@math.uiuc.edu

IMPA, ESTRADA DONA CASTORINA 110, JARDIM BOTÂNICO, RIO DE JANEIRO, RJ, BRASIL

*E-mail address:* rob@impa.br

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL; AND TRINITY COLLEGE, CAMBRIDGE CB2 1TQ, UK

*E-mail address:* ws299@cam.ac.uk