

# INDEPENDENT SETS IN HYPERGRAPHS AND RAMSEY PROPERTIES OF GRAPHS AND THE INTEGERS

ROBERT HANCOCK, KATHERINE STADEN AND ANDREW TREGLOWN

ABSTRACT. Many important problems in combinatorics and other related areas can be phrased in the language of independent sets in hypergraphs. Recently Balogh, Morris and Samotij [3], and independently Saxton and Thomason [54] developed very general container theorems for independent sets in hypergraphs; both of which have seen numerous applications to a wide range of problems. In this paper we use the container method to prove results that correspond to problems concerning tuples of *disjoint* independent sets in hypergraphs. In particular:

- We generalise the *random Ramsey theorem* of Rödl and Ruciński [46, 47, 48] by providing a resilience analogue. This result also implies the *random version of Turán’s theorem* due to Conlon and Gowers [13] and Schacht [55].
- We prove a general subcase of the *asymmetric random Ramsey conjecture* of Kohayakawa and Kreuter [35].
- Both of the above results in fact hold for uniform hypergraphs.
- For a (hyper)graph  $H$ , we determine, up to an error term in the exponent, the number of  $n$ -vertex (hyper)graphs  $G$  that have the Ramsey property with respect to  $H$  (that is, whenever  $G$  is  $r$ -coloured, there is a monochromatic copy of  $H$  in  $G$ ).
- We strengthen the *random Rado theorem* of Friedgut, Rödl and Schacht [21] by proving a resilience version of the result.
- For partition regular matrices  $A$  we determine, up to an error term in the exponent, the number of subsets of  $\{1, \dots, n\}$  for which there exists an  $r$ -colouring which contains no monochromatic solutions to  $Ax = 0$ .

Along the way a number of open problems are posed.

## 1. INTRODUCTION

Recently, the *container method* has developed as a powerful tool for attacking problems which reduce to counting independent sets in (hyper)graphs. Loosely speaking, container results typically state that the independent sets of a given (hyper)graph  $H$  lie only in a ‘small’ number of subsets of the vertex set of  $H$  (referred to as *containers*), where each of these containers is an ‘almost independent set’. The method has been of particular importance because a diverse range of problems in combinatorics and other areas can be rephrased into this setting. For example, container results have been used to tackle problems arising in Ramsey theory, combinatorial number theory, positional games, list colourings of graphs and  $H$ -free graphs.

Although the container method has seen an explosion in applications over the last few years, the technique actually dates back to work of Kleitman and Winston [33, 34] from more than 30 years ago; they constructed a relatively simple algorithm that can be used to produce *graph* container results. The catalysts for recent advances in the area are the *hypergraph* container theorems of Balogh, Morris and Samotij [3] and Saxton and Thomason [54]. Both works yield very general container theorems for hypergraphs whose edge distribution satisfies certain boundedness conditions. These results are also related to general transference theorems of Conlon and Gowers [13]

---

*Date:* January 17, 2017.

RH: University of Birmingham, United Kingdom, [rah410@bham.ac.uk](mailto:rah410@bham.ac.uk). KS: University of Warwick, United Kingdom, [k.l.staden@warwick.ac.uk](mailto:k.l.staden@warwick.ac.uk). Research supported by ERC Grant 306493. AT: University of Birmingham, United Kingdom, [a.c.treglown@bham.ac.uk](mailto:a.c.treglown@bham.ac.uk). Research supported by EPSRC grant EP/M016641/1.

and Schacht [55]. In particular, these container and transference theorems can be used to prove a range of combinatorial results in a random setting.

In this paper we use the container method to prove results that correspond to problems concerning tuples of *disjoint* independent sets in hypergraphs. In particular, from the container theorem of Balogh, Morris and Samotij one can obtain an analogous result for tuples of disjoint independent sets in hypergraphs (see Theorem 3.2). It turns out that many Ramsey-type questions (and other problems) can be naturally phrased in this setting. For example, by Schur’s theorem we know that, if  $n$  is large, then whenever one  $r$ -colours the elements of  $[n] := \{1, \dots, n\}$  there is a monochromatic solution to  $x + y = z$ . This raises the question of how large can a subset  $S \subseteq [n]$  be whilst failing to have this property? (This problem was first posed back in 1977 by Abbott and Wang [1].) Let  $H$  be the hypergraph with vertex set  $[n]$  in which edges precisely correspond to solutions to  $x + y = z$ . (Note  $H$  will have edges of size 2 and 3.) Then sets  $S \subseteq [n]$  without this property are precisely the union of  $r$  disjoint independent sets in  $H$ .

In Section 3 we state the container theorem for tuples of independent sets in hypergraphs. In Sections 4 and 5 we give some applications of this container result to enumeration and resilience questions arising in Ramsey theory for graphs and the integers.

### 1.1. Resilience in hypergraphs and the integers.

1.1.1. *Resilience in graphs.* The notion of *graph resilience* has received significant attention in recent years. Roughly speaking, resilience concerns the question of how ‘strongly’ a graph  $G$  satisfies a certain monotone graph property  $\mathcal{P}$ . *Global resilience* concerns how many edges one can delete and still ensure the resulting graph has property  $\mathcal{P}$  whilst *local resilience* considers how many edges one can delete at each vertex whilst ensuring the resulting graph has property  $\mathcal{P}$ . More precisely, we define the *global resilience of  $G$  with respect to  $\mathcal{P}$* ,  $\text{res}(G, \mathcal{P})$ , to be the minimum number  $t$  such that by deleting  $t$  edges from  $G$ , one can obtain a set not having  $\mathcal{P}$ . Many classical results in extremal combinatorics can be rephrased in terms of resilience. For example, Turán’s theorem determines the global resilience of  $K_n$  with respect to the property of containing  $K_r$  (where  $r < n$ ) as a subgraph.

The systematic study of graph resilience was initiated in a paper of Sudakov and Vu [60], though such questions had been studied before this. In particular, a key question in the area is to establish the resilience of various properties of the Erdős–Rényi random graph  $G_{n,p}$ . (Recall that  $G_{n,p}$  has vertex set  $[n]$  in which each possible edge is present with probability  $p$ , independent of all other choices.) The local resilience of  $G_{n,p}$  has been investigated, for example, with respect to Hamiltonicity e.g. [60, 39], almost spanning trees [2] and embedding subgraphs of small bandwidth [7]. See [60] and the surveys [12, 59] for further background on the subject. In this paper we study the global resilience of  $G_{n,p}$  with respect to Ramsey properties (in fact, as we explain later, we will consider its hypergraph analogue  $G_{n,p}^{(k)}$  for  $k \geq 2$ ). First we will focus on the graph case.

1.1.2. *Ramsey properties of random graphs.* An event occurs in  $G_{n,p}$  with *high probability* (w.h.p.) if its probability tends to 1 as  $n \rightarrow \infty$ . For many properties  $\mathcal{P}$  of  $G_{n,p}$ , the probability that  $G_{n,p}$  has the property exhibits a phase transition, changing from 0 to 1 over a small interval. That is, there is a *threshold* for  $\mathcal{P}$ : a function  $p_0 = p_0(n)$  such that  $G_{n,p}$  has  $\mathcal{P}$  w.h.p. when  $p \gg p_0$  (the *1-statement*), while  $G_{n,p}$  does not have  $\mathcal{P}$  w.h.p. when  $p \ll p_0$  (the *0-statement*). Indeed, Bollobás and Thomason [6] proved that every *monotone* property  $\mathcal{P}$  has a threshold.

Given a graph  $H$ , set  $d_2(H) := 0$  if  $e(H) = 0$ ;  $d_2(H) := 1/2$  when  $H$  is precisely an edge and define  $d_2(H) := (e(H) - 1)/(v(H) - 2)$  otherwise. Then define  $m_2(H) := \max_{H' \subseteq H} d_2(H')$  to be the *2-density* of  $H$ . This graph parameter turns out to be very important when determining the threshold for certain properties in  $G_{n,p}$  concerning the containment of a small subgraph  $H$ , which we explain further below.

Given  $\varepsilon > 0$  and a graph  $H$ , we say that a graph  $G$  is  $(H, \varepsilon)$ -Turán if every subgraph of  $G$  with at least  $(1 - \frac{1}{\chi(H)-1} + \varepsilon)e(G)$  edges contains a copy of  $H$ . Note that the Erdős–Stone theorem implies that  $K_n$  is  $(H, \varepsilon)$ -Turán for any fixed  $H$  provided  $n$  is sufficiently large. To motivate the definition, consider any graph  $G$ . Then by considering a random partition of  $V(G)$  into  $\chi(H) - 1$  parts (and then removing any edge contained within a part) we see that there is a subgraph  $G'$  of  $G$  that is  $(\chi(H) - 1)$ -partite where  $e(G') \geq (1 - \frac{1}{\chi(H)-1})e(G)$ . In particular,  $H \not\subseteq G'$ . Intuitively speaking, this implies that (up to the  $\varepsilon$  term),  $(H, \varepsilon)$ -Turán graphs are those graphs that most *strongly* contain  $H$ .

Rephrasing to the language of resilience, we see that if, for any  $\varepsilon > 0$ ,  $G$  is  $(H, \varepsilon)$ -Turán, then  $\text{res}(G, G \supseteq H) = (\frac{1}{\chi(H)-1} \pm \varepsilon)e(G)$ , and vice versa. The global resilience of  $G_{n,p}$  with respect to the Turán problem has been extensively studied. Indeed, a recent trend in combinatorics and probability concerns so-called *sparse random* analogues of extremal theorems (see [12]), and determining when  $G_{n,p}$  is  $(H, \varepsilon)$ -Turán is an example of such a result.

If  $p \leq cn^{-1/m_2(H)}$  for some small constant  $c$ , then it is not hard to show that w.h.p.  $G_{n,p}$  is not  $(H, \varepsilon)$ -Turán. In [27, 28, 36] it was conjectured that w.h.p.  $G_{n,p}$  is  $(H, \varepsilon)$ -Turán provided that  $p \geq Cn^{-1/m_2(H)}$ , where  $C$  is a (large) constant. After a number of partial results, this conjecture was confirmed by Schacht [55] and Conlon and Gowers [13] (in the case when  $H$  is strictly 2-balanced, i.e.  $m_2(H') < m_2(H)$  for all  $H' \subset H$ ).

**Theorem 1.1** ([55, 13]). *For any graph  $H$  with  $\Delta(H) \geq 2$  and any  $\varepsilon > 0$ , there are positive constants  $c, C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p}, (H, \varepsilon)\text{-Turán}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_2(H)}; \\ 1 & \text{if } p > Cn^{-1/m_2(H)}. \end{cases}$$

Given an integer  $r$ , an  $r$ -colouring of a graph  $G$  is a function  $\sigma : E(G) \rightarrow [r]$ . (So this is not necessarily a proper colouring.) We say that  $G$  is  $(H, r)$ -Ramsey if every  $r$ -colouring of  $G$  yields a monochromatic copy of  $H$  in  $G$ . Observe that being  $(H, 1)$ -Ramsey is the same as containing  $H$  as a subgraph. So the 1-statement of Theorem 1.1 says that, given  $\varepsilon > 0$ , there exists a positive constant  $C$  such that, if  $p > Cn^{-1/m_2(H)}$ , then

$$(1.1) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res}(G_{n,p}, (H, 1)\text{-Ramsey})}{e(G_{n,p})} = \frac{1}{\chi(H) - 1} \pm \varepsilon \right] = 1.$$

The following result of Rödl and Ruciński [46, 47, 48] yields a random version of Ramsey's theorem.

**Theorem 1.2** ([46, 47, 48]). *Let  $r \geq 2$  be a positive integer and let  $H$  be a graph that is not a forest consisting of stars and paths of length 3. There are positive constants  $c, C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H, r)\text{-Ramsey}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_2(H)}; \\ 1 & \text{if } p > Cn^{-1/m_2(H)}. \end{cases}$$

Thus  $n^{-1/m_2(H)}$  is again the threshold for the  $(H, r)$ -Ramsey property. Let us provide some intuition as to why. The expected number of copies of  $H$  in  $G_{n,p}$  is  $\Theta(n^{v(H)}p^{e(H)})$ , while the expected number of edges in  $G_{n,p}$  is  $\Theta(pn^2)$ . When  $p = \Theta(n^{-1/d_2(H)})$ , these quantities agree up to a constant. Suppose that  $H$  is 2-balanced, i.e.  $d_2(H) = m_2(H)$ . For small  $c > 0$ , when  $p < cn^{-1/m_2(H)}$ , most copies of  $H$  in  $G_{n,p}$  contain an edge which appears in no other copy. Thus we can hope to colour these special edges blue and colour the remaining edges red to eliminate all monochromatic copies of  $H$ . For large  $C > 0$ , most edges lie in many copies of  $H$ , so the copies of  $H$  are highly overlapping and we cannot avoid monochromatic copies. In general, when  $H$  is not

necessarily 2-balanced, the threshold is  $n^{-1/d_2(H')}$  for the ‘densest’ subgraph  $H'$  of  $H$  since, roughly speaking, the appearance of  $H$  is governed by the appearance of its densest part.

We remark that Nenadov and Steger [43] recently gave a short proof of Theorem 1.2 using the container method.

1.1.3. *Asymmetric Ramsey properties in random graphs.* It is natural to ask for an *asymmetric* analogue of Theorem 1.2. Now, for graphs  $H_1, \dots, H_r$ , a graph  $G$  is  $(H_1, \dots, H_r)$ -Ramsey if for any  $r$ -colouring of  $G$  there is a copy of  $H_i$  in colour  $i$  for some  $i \in [r]$ . (This definition coincides with that of  $(H, r)$ -Ramsey when  $H_1 = \dots = H_r = H$ .) Kohayakawa and Kreuter [35] conjectured an analogue of Theorem 1.2 in the asymmetric case. To state it, we need to introduce the *asymmetric density* of  $H_1, H_2$  where  $m_2(H_1) \geq m_2(H_2)$  via

$$(1.2) \quad m_2(H_1, H_2) := \max \left\{ \frac{e(H'_1)}{v(H'_1) - 2 + 1/m_2(H_2)} : H'_1 \subseteq H_1 \text{ and } e(H'_1) \geq 1 \right\}.$$

**Conjecture 1.3** ([35]). *For any graphs  $H_1, \dots, H_r$  with  $m_2(H_1) \geq \dots \geq m_2(H_r) > 1$ , there are positive constants  $c, C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H_1, \dots, H_r)\text{-Ramsey}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_2(H_1, H_2)}; \\ 1 & \text{if } p > Cn^{-1/m_2(H_1, H_2)}. \end{cases}$$

So the conjectured threshold only depends on the ‘joint density’ of the densest two graphs  $H_1, H_2$ . The intuition for this threshold is discussed in detail e.g. in Section 1.1 in [26]. One can show that  $m_2(H_1) \geq m_2(H_1, H_2) \geq m_2(H_2)$  with equality if and only if  $m_2(H_1) = m_2(H_2)$ . Thus Conjecture 1.3 would generalise Theorem 1.2. Kohayakawa and Kreuter [35] have confirmed Conjecture 1.3 when the  $H_i$  are cycles. In [41] it was observed that the approach used by Kohayakawa and Kreuter [35] implies the 1-statement of Conjecture 1.3 holds when  $H_1$  is strictly 2-balanced *provided* the so-called KLR conjecture holds. This latter conjecture was proven by Balogh, Morris and Samotij [3] thereby proving the 1-statement of Conjecture 1.3 holds in this case.

1.1.4. *Ramsey properties of random hypergraphs.* Consider now the  $k$ -uniform analogue  $G_{n,p}^{(k)}$  of  $G_{n,p}$  which has vertex set  $[n]$  and in which every  $k$ -element subset of  $[n]$  appears as an edge with probability  $p$ , independent of all other choices. Here, we wish to obtain analogues of Theorems 1.1, 1.2 and Conjecture 1.3 by determining the threshold for being  $(H, \varepsilon)$ -Turán,  $(H, r)$ -Ramsey, and more generally being  $(H_1, \dots, H_r)$ -Ramsey. The definitions of  $(H, r)$ -Ramsey and  $(H_1, \dots, H_r)$ -Ramsey extend from graphs in the obvious way. Given a  $k$ -uniform hypergraph  $H$ , let  $\text{ex}(n; H)$  be the maximum size of an  $n$ -vertex  $H$ -free hypergraph. A simple averaging argument shows that the limit

$$\pi(H) := \frac{\text{ex}(n; H)}{\binom{n}{k}}$$

exists. Now we say that a  $k$ -uniform hypergraph  $G$  is  $(H, \varepsilon)$ -Turán if every subhypergraph of  $G$  with at least  $(\pi(H) + \varepsilon)e(G)$  edges contains a copy of  $H$ . (Since  $\pi(H) = 1 - \frac{1}{\chi(H)-1}$  when  $k = 2$ , this generalises the definition we gave earlier.) We also need to generalise the notion of 2-density to  $k$ -density: Given a  $k$ -graph  $H$ , define

$$d_k(H) := \begin{cases} 0 & \text{if } e(H) = 0; \\ 1/k & \text{if } v(H) = k \text{ and } e(H) = 1; \\ \frac{e(H)-1}{v(H)-k} & \text{otherwise,} \end{cases}$$

and let

$$m_k(H) := \max_{H' \subseteq H} d_k(H').$$

The techniques of Conlon–Gowers [13] and of Schacht [55] actually extended to a proof of a version of Theorem 1.1 for hypergraphs:

**Theorem 1.4** ([13, 55]). *For any  $k$ -uniform hypergraph  $H$  with maximum vertex degree at least two and any  $\varepsilon > 0$ , there are positive constants  $c, C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p}^{(k)} \text{ is } (H, \varepsilon)\text{-Turán}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_k(H)}; \\ 1 & \text{if } p > Cn^{-1/m_k(H)}. \end{cases}$$

The 1-statement of Theorem 1.2 was generalised to hypergraphs by Friedgut, Rödl and Schacht [21] and by Conlon and Gowers [13], proving a conjecture of Rödl and Ruciński [50]. (The special cases of the complete 3-uniform hypergraph  $K_4^{(3)}$  on four vertices and of  $k$ -partite  $k$ -uniform hypergraphs were already proved in [50], [51] respectively.)

**Theorem 1.5** ([13, 21]). *Let  $r, k \geq 2$  be integers and let  $H$  be a  $k$ -uniform hypergraph with maximum vertex degree at least two. There is a positive constant  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p}^{(k)} \text{ is } (H, r)\text{-Ramsey}] = 1 \quad \text{if } p > Cn^{-1/m_k(H)}.$$

In [26], sufficient conditions are given for a corresponding 0-statement. However, the authors further show that, for  $k \geq 4$ , there is a  $k$ -uniform hypergraph  $H$  such that the threshold for  $G_{n,p}^{(k)}$  to be  $(H, r)$ -Ramsey is *not*  $n^{-1/m_k(H)}$ , and *nor* does it correspond to the exceptional case in the graph setting of certain forests, where there is a coarse threshold due to the appearance of small subgraphs.

For the asymmetric Ramsey problem, we need to suitably generalise (1.2), in the obvious way: for any  $k$ -uniform hypergraphs  $H_1, H_2$  with non-empty edge sets and  $m_k(H_1) \geq m_k(H_2)$ , let

$$(1.3) \quad m_k(H_1, H_2) := \max \left\{ \frac{e(H'_1)}{v(H'_1) - k + 1/m_k(H_2)} : H'_1 \subseteq H_1 \text{ and } e(H'_1) \geq 1 \right\}$$

be the *asymmetric  $k$ -density* of  $(H_1, H_2)$ . Again,

$$m_k(H_1) \geq m_k(H_1, H_2) \geq m_k(H_2),$$

so, in particular,  $m_k(H_1, H_2) = m_k(H_1)$  if and only if  $H_1$  and  $H_2$  have the same  $k$ -density.

Recently, Gugelmann, Nenadov, Person, Steger, Škorić and Thomas [26] generalised the 1-statement of Conjecture 1.3 to  $k$ -uniform hypergraphs, in the case when  $H'_1 = H_1$  is the unique maximiser in (1.3), i.e.  $H_1$  is *strictly  $k$ -balanced with respect to  $m_k(\cdot, H_2)$* .

**Theorem 1.6** ([26]). *For all positive integers  $r, k$  with  $k \geq 2$  and  $k$ -uniform hypergraphs  $H_1, \dots, H_r$  with  $m_k(H_1) \geq \dots \geq m_k(H_r)$  where  $H_1$  is strictly  $k$ -balanced with respect to  $m_k(\cdot, H_2)$ , there exists  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ G_{n,p}^{(k)} \text{ is } (H_1, \dots, H_r)\text{-Ramsey} \right] = 1 \quad \text{if } p > Cn^{-1/m_k(H_1, H_2)}.$$

They further prove a version of Theorem 1.6 with the weaker bound  $p > Cn^{-1/m_k(H_1, H_2)} \log n$  when  $H_1$  is not required to be strictly  $k$ -balanced with respect to  $m_k(\cdot, H_2)$ .

1.1.5. *New resilience result.* Our main result here is Theorem 1.7, which generalises, fully and partially, all of the 1-statements of the results discussed in this section, giving a unified setting for both the random Ramsey theorem and the random Turán theorem. Once we have obtained a container theorem for Ramsey graphs (Theorem 5.11), the proof is short (see Section 5.6).

For  $k$ -uniform hypergraphs  $H_1, \dots, H_r$  and a positive integer  $n$ , let  $\text{ex}^r(n; H_1, \dots, H_r)$  be the maximum size of an  $n$ -vertex  $k$ -uniform hypergraph  $G$  which is not  $(H_1, \dots, H_r)$ -Ramsey. Define the  $r$ -coloured Turán density

$$(1.4) \quad \pi(H_1, \dots, H_r) := \lim_{n \rightarrow \infty} \frac{\text{ex}^r(n; H_1, \dots, H_r)}{\binom{n}{k}}.$$

Observe that  $\text{ex}^1(n; H) = \text{ex}(n; H)$  since a hypergraph is  $H$ -free if and only if it is not  $(H, 1)$ -Ramsey. Note further that  $\pi(\cdot, \dots, \cdot)$  generalises  $\pi(\cdot)$ . So when  $k = 2$ , we have  $\pi(H) = 1 - \frac{1}{\chi(H)-1}$ . We will show in Section 5.2 that the limit in (1.4) does indeed exist, so  $\pi(\cdot, \dots, \cdot)$  is well-defined. Further, crucially for  $k$ -uniform hypergraphs  $H_1, \dots, H_r$ , there exists an  $\varepsilon = \varepsilon(H_1, \dots, H_r) > 0$  so that  $\pi(H_1, \dots, H_r) < 1 - \varepsilon$  (see (5.3) in Section 5.2).

**Theorem 1.7** (Resilience for random Ramsey). *Let  $\delta > 0$ , let  $r, k$  be positive integers with  $k \geq 2$  and let  $H_1, \dots, H_r$  be  $k$ -uniform hypergraphs each with maximum vertex degree at least two, and such that  $m_k(H_1) \geq \dots \geq m_k(H_r)$ . There exists  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res} \left( G_{n,p}^{(k)}, (H_1, \dots, H_r)\text{-Ramsey} \right)}{e \left( G_{n,p}^{(k)} \right)} = 1 - \pi(H_1, \dots, H_r) \pm \delta \right] = 1 \quad \text{if } p > Cn^{-1/m_k(H_1)}.$$

Thus, when  $p > Cn^{-1/m_k(H_1)}$ , the random hypergraph  $G_{n,p}^{(k)}$  is w.h.p such that every subhypergraph  $G'$  with at least a  $\pi(H_1, \dots, H_r) + \Omega(1)$  fraction of the edges is  $(H_1, \dots, H_r)$ -Ramsey. Conversely, there is a subgraph of  $G_{n,p}^{(k)}$  whose edge density is slightly smaller than this which does not have the Ramsey property.

Let us describe the importance of Theorem 1.7 (in the case  $k = 2$  and  $H_1 = \dots = H_r = H$ ) in conjunction with Theorem 1.2. The 0-statement of Theorem 1.2 says that a typical sparse graph, i.e. one with density at most  $cn^{2-1/m_2(H)}$ , is not  $(H, r)$ -Ramsey. On the other hand, by Theorem 1.7, a typical dense graph, i.e. one with density at least  $Cn^{2-1/m_2(H)}$ , has the Ramsey property in a sense which is *as strong as possible* with respect to subgraphs: every sufficiently dense subgraph is  $(H, r)$ -Ramsey, and this minimum density is the largest we could hope to require.

The relationship between Theorem 1.7 and the previous results stated in this section can be summarised as follows:

- The 1-statement of Theorem 1.1 is recovered when  $k = 2$  and  $r = 1$ . This follows from (1.1) and the relation between  $\pi(H)$  and  $\chi(H)$ .
- In the case  $k = 2$  and  $H_1 = \dots = H_r = H$ , we obtain a stronger statement in place of the 1-statement of Theorem 1.2 as described above.
- Theorem 1.7 proves Conjecture 1.3 in the case when  $m_2(H_1) = m_2(H_2)$  in the same stronger sense as above.
- The 1-statement of Theorem 1.4 is recovered when  $r = 1$ .
- Theorem 1.7 implies Theorem 1.5, yielding a resilience version of this result.
- Theorem 1.7 implies a version of Theorem 1.6 when  $m_k(H_1) = m_k(H_2)$  but now  $H_1$  is not required to be strictly  $k$ -balanced with respect to  $m_k(\cdot, H_2)$ .

1.1.6. *Resilience in the integers.* An important branch of Ramsey theory concerns partition properties of sets of integers. Schur's classical theorem [56] states that if  $\mathbb{N}$  is  $r$ -coloured there exists a monochromatic solution to  $x + y = z$ ; later van der Waerden [61] showed that the same hypothesis ensures a monochromatic arithmetic progression of arbitrary length. More generally, Rado's theorem [45] characterises all those systems of homogeneous linear equations  $\mathcal{L}$  for which every finite colouring of  $\mathbb{N}$  yields a monochromatic solution to  $\mathcal{L}$ .

As in the graph case, there has been interest in proving random analogues of such results from arithmetic Ramsey theory. Before we describe the background of this area we will introduce some notation and definitions. Throughout we will assume that  $A$  is an  $\ell \times k$  integer matrix where  $k \geq \ell$  of full rank  $\ell$ . We will let  $\mathcal{L}(A)$  denote the associated system of linear equations  $Ax = 0$ , noting that for brevity we will simply write  $\mathcal{L}$  if it is clear from the context which matrix  $A$  it refers to. Let  $S$  be a set of integers. If a vector  $x = (x_1, \dots, x_k) \in S^k$  satisfies  $Ax = 0$  (i.e. it is a solution to  $\mathcal{L}$ ) and the  $x_i$  are distinct we call  $x$  a  $k$ -distinct solution to  $\mathcal{L}$  in  $S$ .

We call a set  $S$  of integers  $(\mathcal{L}, r)$ -free if there exists an  $r$ -colouring of  $S$  such that it contains no monochromatic  $k$ -distinct solution to  $\mathcal{L}$ . Otherwise we call  $S$   $(\mathcal{L}, r)$ -Rado. In the case when  $r = 1$ , we write  $\mathcal{L}$ -free instead of  $(\mathcal{L}, 1)$ -free. Define  $\mu(n, \mathcal{L}, r)$  to be the size of the largest  $(\mathcal{L}, r)$ -free subset of  $[n]$ .

A matrix  $A$  is *partition regular* if for any finite colouring of  $\mathbb{N}$ , there is always a monochromatic solution to  $\mathcal{L}$ . As mentioned above, Rado's theorem characterises all those integer matrices  $A$  that are partition regular. A matrix  $A$  is *irredundant* if there exists a  $k$ -distinct solution to  $\mathcal{L}$  in  $\mathbb{N}$ . Otherwise  $A$  is *redundant*. The study of random versions of Rado's theorem has focused on irredundant partition regular matrices. This is natural since for every redundant  $\ell \times k$  matrix  $A$  there exists an irredundant  $\ell' \times k'$  matrix  $A'$  for some  $\ell' < \ell$  and  $k' < k$  with the same family of solutions (viewed as sets). See [49, Section 1] for a full explanation.

Another class of matrices that have received attention in relation to this problem are so-called density regular matrices: An irredundant, partition regular matrix  $A$  is *density regular* if any subset  $F \subseteq \mathbb{N}$  with positive upper density, i.e.,

$$\limsup_{n \rightarrow \infty} \frac{|F \cap [n]|}{n} > 0,$$

contains a  $k$ -distinct solution to  $\mathcal{L}$ .

Index the columns of  $A$  by  $[k]$ . For a partition  $W \dot{\cup} \overline{W} = [k]$  of the columns of  $A$ , we denote by  $A_{\overline{W}}$  the matrix obtained from  $A$  by restricting to the columns indexed by  $\overline{W}$ . Let  $\text{rank}(A_{\overline{W}})$  be the rank of  $A_{\overline{W}}$ , where  $\text{rank}(A_{\overline{W}}) = 0$  for  $\overline{W} = \emptyset$ . We set

$$(1.5) \quad m(A) := \max_{\substack{W \dot{\cup} \overline{W} = [k] \\ |\overline{W}| \geq 2}} \frac{|W| - 1}{|W| - 1 + \text{rank}(A_{\overline{W}}) - \text{rank}(A)}.$$

We now describe some random analogues of results from arithmetic Ramsey theory. Recall that  $[n]_p$  denotes a set where each element  $a \in [n]$  is included with probability  $p$  independently of all other elements. Rödl and Ruciński [49] showed that for irredundant partition regular matrices  $A$ ,  $m(A)$  is an important parameter for determining whether  $[n]_p$  is  $(\mathcal{L}, r)$ -Rado or  $(\mathcal{L}, r)$ -free.

**Theorem 1.8** ([49]). *For all irredundant partition regular full rank matrices  $A$  and all positive integers  $r \geq 2$ , there exists a constant  $c > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[[n]_p \text{ is } (\mathcal{L}, r)\text{-Rado}] = 0 \quad \text{if } p < cn^{-1/m(A)}.$$

We remark it is important that  $r \geq 2$  in Theorem 1.8. That is, the corresponding statement for  $r = 1$  is not true in general. Roughly speaking, Theorem 1.8 implies that almost all subsets of  $[n]$  with significantly fewer than  $n^{1-1/m(A)}$  elements are  $(\mathcal{L}, r)$ -free for any irredundant partition regular matrix  $A$ . The following theorem of Friedgut, Rödl and Schacht [21] complements this result, implying that almost all subsets of  $[n]$  with significantly more than  $n^{1-1/m(A)}$  elements are  $(\mathcal{L}, r)$ -Rado for any irredundant partition regular matrix  $A$ .

**Theorem 1.9** ([21]). *For all irredundant partition regular full rank matrices  $A$  and all positive integers  $r$ , there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[[n]_p \text{ is } (\mathcal{L}, r)\text{-Rado}] = 1 \quad \text{if } p > Cn^{-1/m(A)}.$$

Earlier, Theorem 1.9 was confirmed by Graham, Rödl and Ruciński [22] in the case where  $\mathcal{L}$  is  $x + y = z$  and  $r = 2$ , and then by Rödl and Ruciński [49] in the case when  $A$  is density regular.

Together Theorems 1.8 and 1.9 show that the threshold for the property of being  $(\mathcal{L}, r)$ -Rado is  $p = n^{-1/m(A)}$ . In light of this, it is interesting to ask if above this threshold the property of being  $(\mathcal{L}, r)$ -Rado is resilient to the deletion of a significant number of elements. To be precise, given a set  $S$ , we define the *resilience of  $S$  with respect to  $\mathcal{P}$* ,  $\text{res}(S, \mathcal{P})$ , to be the minimum number  $t$  such that by deleting  $t$  elements from  $S$ , one can obtain a set not having  $\mathcal{P}$ . For example, when  $\mathcal{P}$  is the property of containing an arithmetic progression of length  $k$ , then Szemerédi's theorem can be phrased in terms of resilience; it states that for all  $k \geq 3$  and  $\varepsilon > 0$ , there exists  $n_0 > 0$  such that for all integers  $n \geq n_0$ , we have  $\text{res}([n], \mathcal{P}) \geq (1 - \varepsilon)n$ .

The following result of Schacht [55] provides a resilience strengthening of Theorem 1.9 in the case of density regular matrices.

**Theorem 1.10** ([55]). *For all irredundant density regular full rank matrices  $A$ , all positive integers  $r$  and all  $\varepsilon > 0$ , there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res}([n]_p, (\mathcal{L}, r)\text{-Rado})}{|[n]_p|} \geq 1 - \varepsilon \right] = 1 \quad \text{if } p > Cn^{-1/m(A)}.$$

Note that in [55] the result is stated in the  $r = 1$  case only, but the general result follows immediately from this special case.

Our next result gives a resilience strengthening of Theorem 1.9 for all irredundant partition regular matrices.

**Theorem 1.11.** *For all irredundant partition regular full rank matrices  $A$ , all positive integers  $r$  and all  $\delta > 0$ , there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res}([n]_p, (\mathcal{L}, r)\text{-Rado})}{|[n]_p|} = 1 - \frac{\mu(n, \mathcal{L}, r)}{n} \pm \delta \right] = 1 \quad \text{if } p > Cn^{-1/m(A)}.$$

It is well known that for all irredundant partition regular full rank matrices  $A$  and all positive integers  $r$ , there exist  $n_0 = n_0(A, r), \eta = \eta(A, r) > 0$ , such that for all integers  $n \geq n_0$ , we have  $\mu(n, \mathcal{L}, r) \leq (1 - \eta)n$ . (This follows from a supersaturation lemma of Frankl, Graham and Rödl [20, Theorem 1].) Thus, Theorem 1.11 does imply Theorem 1.9. Further, in the case when  $A$  is density regular, [20, Theorem 2] immediately implies that  $\mu(n, \mathcal{L}, r) = o(n)$  for any fixed  $r \in \mathbb{N}$ . Thus Theorem 1.11 implies Theorem 1.10. Theorem 1.11 in the case when  $r = 1$  and  $\mathcal{L}$  is  $x + y = z$  was proved by Schacht [55]. In fact, the method of Schacht can be used to prove the theorem for  $r = 1$  and every irredundant partition regular matrix  $A$ .

Intuitively, the reader can interpret Theorem 1.11 as stating that almost all subsets of  $[n]$  with significantly more than  $n^{1-1/m(A)}$  elements strongly possess the property of being  $(\mathcal{L}, r)$ -Rado for any irredundant partition regular matrix  $A$ . The 'strength' here depends on the parameter  $\mu(n, \mathcal{L}, r)$ . In light of this it is natural to seek good bounds on  $\mu(n, \mathcal{L}, r)$  (particularly in the cases when  $\mu(n, \mathcal{L}, r) = \Omega(n)$ ). In general, not too much is known about this parameter. However, as mentioned earlier, in the case when  $A = (1, 1, -1)$  (i.e.  $\mathcal{L}$  is  $x + y = z$ ), this is (essentially) a 40-year-old problem of Abbott and Wang [1]. In Section 4.6 we give an upper bound on  $\mu(n, \mathcal{L}, r)$  in this case for all  $r \in \mathbb{N}$ .

Instead of proving Theorem 1.11 directly, in Section 4 we will prove a version of the result that holds for a more general class of matrices  $A$ , namely Theorem 4.1.



**Additional note.** Just before submitting the paper we were made aware of simultaneous and independent work of Spiegel [58]. In [58] the case  $r = 1$  of Theorem 4.1 is proven. Spiegel also used the container method to give an alternative proof of Theorem 1.9.

**1.2. Enumeration questions for Ramsey problems.** A fundamental question in combinatorics is to determine the number of structures with a given property. For example, Erdős, Frankl and Rödl [16] proved that the number of  $n$ -vertex  $H$ -free graphs is  $2^{\binom{n}{2}(1-\frac{1}{r-1}+o(1))}$  for any graph  $H$  of chromatic number  $r$ . Here the lower bound follows by considering all the subgraphs of the  $(r-1)$ -partite Turán graph. Given any  $k, r, n \in \mathbb{N}$  with  $k \geq 2$  and  $k$ -uniform hypergraphs  $H_1, \dots, H_r$ , define  $\text{Ram}(n; H_1, \dots, H_r)$  to be the collection of all  $k$ -uniform hypergraphs on vertex set  $[n]$  that are  $(H_1, \dots, H_r)$ -Ramsey and  $\overline{\text{Ram}}(n; H_1, \dots, H_r)$  to be all those  $k$ -uniform hypergraphs on  $[n]$  that are not  $(H_1, \dots, H_r)$ -Ramsey. A natural question is to determine the size of  $\text{Ram}(n; H_1, \dots, H_r)$ . Surprisingly, we are unaware of any *explicit* results in this direction for  $r \geq 2$ . The next application of the container method fully answers this question up to an error term in the exponent.

**Theorem 1.12.** *Let  $k, r, n \in \mathbb{N}$  with  $k \geq 2$  and  $H_1, \dots, H_r$  be  $k$ -uniform hypergraphs. Then*

$$|\overline{\text{Ram}}(n; H_1, \dots, H_r)| = 2^{\text{ex}^r(n, H_1, \dots, H_r) + o(n^k)} = 2^{\pi(H_1, \dots, H_r) \binom{n}{k} + o(n^k)}.$$

Note that in the case when  $k = 2$  and  $r = 1$ , Theorem 1.12 is precisely the above mentioned result of Erdős, Frankl and Rödl [16]. In fact, one can also obtain Theorem 1.12 by using the work from [42], a hypergraph analogue of the result in [16]; see Section 5.5 for a proof of this.

Our final application of the container method determines, up to an error term in the exponent, the number of  $(\mathcal{L}, r)$ -free subsets of  $[n]$ .

**Theorem 1.13.** *Let  $A$  be an irredundant partition regular matrix of full rank and let  $r \in \mathbb{N}$  be fixed. There are  $2^{\mu(n, \mathcal{L}, r) + o(n)}$   $(\mathcal{L}, r)$ -free subsets of  $[n]$ .*

As an illustration, a result of Hu [29] implies that  $\mu(n, \mathcal{L}, 2) = 4n/5 + o(n)$  in the case when  $\mathcal{L}$  is  $x + y = z$ . Thus, Theorem 1.13 tells us all but  $2^{(4/5+o(1))n}$  subsets of  $[n]$  are  $(\mathcal{L}, 2)$ -Rado in this case.

## 2. NOTATION

For a (hyper)graph  $H$ , we define  $V(H)$  and  $E(H)$  to be the vertex and edge sets of  $H$  respectively, and set  $v(H) := |V(H)|$  and  $e(H) := |E(H)|$ . For a set  $A \subseteq V(H)$ , we define  $H[A]$  to be the induced subgraph of  $H$  on the vertex set  $A$ . For an edge set  $X \subseteq E(H)$ , we define  $H - X$  to be hypergraph with vertex set  $V(H)$  and edge set  $E(H) \setminus X$ .

For a set  $A$  and a positive integer  $x$ , we define  $\binom{A}{x}$  to be the set of all subsets of  $A$  of size  $x$ , and we define  $\binom{A}{\leq x}$  to be the set of all subsets of  $A$  of size at most  $x$ . We use  $\mathcal{P}(X)$  to denote the powerset of  $X$ , that is, the set of all subsets of  $X$ . If  $B$  is a family of subsets of  $A$ , then we define  $\overline{B}$  to be the complement family, that is, precisely the subsets of  $A$  which are not in  $B$ .

Given a hypergraph  $\mathcal{H}$ , for each  $T \subseteq V(\mathcal{H})$ , we define  $\text{deg}_{\mathcal{H}}(T) := |\{e \in E(\mathcal{H}) : T \subseteq e\}|$ , and let  $\Delta_{\ell}(\mathcal{H}) := \max\{\text{deg}_{\mathcal{H}}(T) : T \subseteq V(\mathcal{H}) \text{ and } |T| = \ell\}$ .

We write  $x = a \pm b$  to say that the value of  $x$  is some real number in the interval  $[a - b, a + b]$ .

We will make use of the following Chernoff inequality (see e.g. [30, Theorem 2.1, Corollary 2.3]).

**Proposition 2.1.** *Suppose  $X$  has binomial distribution and  $\lambda \geq 0$ . Then*

$$\mathbb{P}[X > \mathbb{E}[X] + \lambda] \leq \exp\left(-\frac{\lambda^2}{2(\mathbb{E}[X] + \lambda/3)}\right).$$

Further, if  $0 < \varepsilon \leq 3/2$  then

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq \varepsilon \mathbb{E}[X]] \leq 2 \exp\left(-\frac{\varepsilon^2}{3} \mathbb{E}[X]\right).$$

### 3. CONTAINER RESULTS FOR DISJOINT INDEPENDENT SETS

Let  $\mathcal{H}$  be a  $k$ -uniform hypergraph with vertex set  $V$ . A family of sets  $\mathcal{F} \subseteq \mathcal{P}(V)$  is called *increasing* if it is closed under taking supersets; in other words for every  $A, B \subseteq V$ , if  $A \in \mathcal{F}$  and  $A \subseteq B$ , then  $B \in \mathcal{F}$ . Suppose  $\mathcal{F}$  is 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  $A \in \mathcal{F}$ . We define  $\mathcal{I}(\mathcal{H})$  to be the set of all independent sets in  $\mathcal{H}$ .

The next result is the general hypergraph container theorem of Balogh, Morris and Samotij [3].

**Theorem 3.1** ([3], Theorem 2.2). *For every  $k \in \mathbb{N}$  and all positive  $c$  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, for every  $\ell \in [k]$ ,*

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

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  $I \in \mathcal{I}(\mathcal{H})$ , we have that  $g(I) \subseteq I$  and  $I \setminus g(I) \subseteq f(g(I))$ .

Using the above notation, we refer to the set  $\mathcal{C} := \{f(g(I)) \cup g(I) : I \in \mathcal{I}(\mathcal{H})\}$  as a set of *containers* and the  $g(I) \in \mathcal{S}$  as *fingerprints*.

Throughout the paper, when we consider  $r$ -tuples of sets, the  $r$ -tuples are *always* ordered. For two  $r$ -tuples of sets  $(I_1, \dots, I_r)$  and  $(J_1, \dots, J_r)$  we write  $(I_1, \dots, I_r) \subseteq (J_1, \dots, J_r)$  if  $I_x \subseteq J_x$  for each  $x \in [r]$ . We write  $(I_1, \dots, I_r) \cup (J_1, \dots, J_r) := (I_1 \cup J_1, \dots, I_r \cup J_r)$ . Define

$$\mathcal{P}((J_1, \dots, J_r)) := \{(I_1, \dots, I_r) : (I_1, \dots, I_r) \subseteq (J_1, \dots, J_r)\}.$$

Given a set  $X$  and  $r \in \mathbb{N}$  we write  $X^r := (X, \dots, X)$  for the  $r$ -tuple with  $X$  in every coordinate. If  $\mathcal{X}$  is a collection of sets then we write  $\mathcal{X}^r$  for the collection of  $r$ -tuples  $(X_1, \dots, X_r)$  so that  $X_i \in \mathcal{X}$  for all  $1 \leq i \leq r$ . We write  $ij$  to denote the pair  $\{i, j\}$ . For a hypergraph  $\mathcal{H}$  define

$$\mathcal{I}_r(\mathcal{H}) := \left\{ (I_1, \dots, I_r) \in \mathcal{P}(V(\mathcal{H})^r) : I_x \in \mathcal{I}(\mathcal{H}) \text{ and } I_i \cap I_j = \emptyset \text{ for all } x \in [r], ij \in \binom{[r]}{2} \right\}.$$

Whereas Theorem 3.1 provides a set of containers for the independent sets of a hypergraph, the following theorem is an analogous result for the  $r$ -tuples of disjoint independent sets of a hypergraph. It is a straightforward consequence of Theorem 3.1.

**Theorem 3.2.** *For every  $k, r \in \mathbb{N}$  and all positive  $c$  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, for every  $\ell \in [k]$ ,*

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

Then there exists a family  $\mathcal{S}_r \subseteq \mathcal{P}(V(\mathcal{H})^r)$  and functions  $f : \mathcal{S}_r \rightarrow (\overline{\mathcal{F}})^r$  and  $g : \mathcal{I}_r(\mathcal{H}) \rightarrow \mathcal{S}_r$  such that the following conditions hold:

- (i) If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum |S_i| \leq Cp \cdot v(\mathcal{H})$ ;
- (ii) every  $S \in \mathcal{S}_r$  satisfies  $S \in \mathcal{I}_r(\mathcal{H})$ ;

(iii) for every  $(I_1, \dots, I_r) \in \mathcal{I}_r(\mathcal{H})$ , we have that  $S \subseteq (I_1, \dots, I_r) \subseteq S \cup f(S)$  where  $S := g(I_1, \dots, I_r)$ .

*Proof.* Apply Theorem 3.1 with  $k, c, \varepsilon$  to obtain a positive constant  $C_1$ . Let  $C := rC_1$ . We will show that  $C$  has the required properties. Let  $\mathcal{H}$  be a  $k$ -uniform hypergraph which together with a set  $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$  satisfies the hypotheses of Theorem 3.2. Since  $\mathcal{H}, \mathcal{F}$  also satisfy the hypotheses of Theorem 3.1, there exists a family  $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leq C_1 p \cdot v(\mathcal{H})}$  and functions  $f' : \mathcal{S} \rightarrow \overline{\mathcal{F}}$  and  $g' : \mathcal{I}(\mathcal{H}) \rightarrow \mathcal{S}$  such that for every  $I \in \mathcal{I}(\mathcal{H})$  we have  $g'(I) \subseteq I$  and  $I \setminus g'(I) \subseteq f'(g'(I))$ . Define

$$\mathcal{S}' := \{S \in \mathcal{S} : \text{there exists } I \in \mathcal{I}(\mathcal{H}) \text{ such that } g'(I) = S\},$$

and

$$\mathcal{S}_r := \left\{ (S_1, \dots, S_r) \in \mathcal{P}(V(\mathcal{H})^r) : S_x \in \mathcal{S}' \text{ and } S_i \cap S_j = \emptyset \text{ for all } x \in [r], ij \in \binom{[r]}{2} \right\}.$$

Let  $(S_1, \dots, S_r) \in \mathcal{S}_r$ . First note that

$$\sum_{x \in [r]} |S_x| \leq C_1 r \cdot pv(\mathcal{H}) = Cp \cdot v(\mathcal{H}),$$

so (i) holds. Also since  $S_x \in \mathcal{S}'$  for all  $x \in [r]$ , we have  $S_x \in \mathcal{I}(\mathcal{H})$  and so by definition of  $\mathcal{S}_r$  we have  $\mathcal{S}_r \subseteq \mathcal{I}_r(\mathcal{H})$  proving (ii).

Consider any  $(S_1, \dots, S_r) \in \mathcal{S}_r$  and any  $(I_1, \dots, I_r) \in \mathcal{I}_r(\mathcal{H})$ . Define  $f : \mathcal{S}_r \rightarrow (\overline{\mathcal{F}})^r$  by setting  $f(S_1, \dots, S_r) := (f'(S_1), \dots, f'(S_r))$  and define  $g : \mathcal{I}_r(\mathcal{H}) \rightarrow \mathcal{S}_r$  by setting  $g(I_1, \dots, I_r) := (g'(I_1), \dots, g'(I_r))$ .

Note that since  $f'(S_x) \in \overline{\mathcal{F}}$ ,  $g'(I_x) \in \mathcal{S}'$  and  $g'(I_i) \cap g'(I_j) = \emptyset$  for all  $x \in [r]$  and  $ij \in \binom{[r]}{2}$ , we do indeed have  $(f'(S_1), \dots, f'(S_r)) \in (\overline{\mathcal{F}})^r$  and  $(g'(I_1), \dots, g'(I_r)) \in \mathcal{S}_r$ .

Now for (iii), since  $g'(I_x) \subseteq I_x$  and  $I_x \setminus g'(I_x) \subseteq f'(g'(I_x))$  for all  $x \in [r]$ , we have  $g(I_1, \dots, I_r) = (g'(I_1), \dots, g'(I_r)) \subseteq (I_1, \dots, I_r)$ . Since  $f(g(I_1, \dots, I_r)) = (f'(g'(I_1)), \dots, f'(g'(I_r)))$  we also have  $(I_1, \dots, I_r) \subseteq f(g(I_1, \dots, I_r)) \cup g(I_1, \dots, I_r)$  as required.  $\square$

In the proof of e.g. Theorem 1.7, instead of considering tuples of disjoint independent sets from the *same* hypergraph  $\mathcal{H}$ , we are actually concerned with disjoint independent sets from *different* hypergraphs but which have the *same vertex set*: For all  $i \in [r]$ , let  $\mathcal{H}_i$  be a  $k_i$ -uniform hypergraph, each on the same vertex set  $V$ , and define  $\mathcal{I}(\mathcal{H}_1, \dots, \mathcal{H}_r)$  to be the set of all  $r$ -tuples  $(I_1, \dots, I_r) \in \prod_{i \in [r]} \mathcal{I}(\mathcal{H}_i)$  such that  $I_i \cap I_j = \emptyset$  for all  $1 \leq i < j \leq r$ .

For such applications we will instead use the following asymmetric version of Theorem 3.2. We omit its proof since it follows from Theorem 3.1 as in the proof of Theorem 3.2.

**Theorem 3.3.** *For every  $r, k_1, \dots, k_r \in \mathbb{N}$  with  $k_i \geq 2$  for all  $i \in [r]$ , and all  $c, \varepsilon > 0$ , there exists a positive constant  $C$  such that the following holds. For all  $i \in [r]$ , let  $\mathcal{H}_i$  be a  $k_i$ -uniform hypergraph, each on the same vertex set  $V$ . For all  $i \in [r]$ , let  $\mathcal{F}_i \subseteq \mathcal{P}(V)$  be an increasing family of sets such that  $|A| \geq \varepsilon|V|$  for all  $A \in \mathcal{F}_i$ . Suppose that each  $\mathcal{H}_i$  is  $(\mathcal{F}_i, \varepsilon)$ -dense. Further suppose  $p \in (0, 1)$  is such that, for every  $i \in [r]$  and  $\ell \in [k_i]$ ,*

$$\Delta_\ell(\mathcal{H}_i) \leq c \cdot p^{\ell-1} \frac{e(\mathcal{H}_i)}{|V|}.$$

*Then there exists a family  $\mathcal{S}_r \subseteq \mathcal{I}(\mathcal{H}_1, \dots, \mathcal{H}_r)$  and functions  $f : \mathcal{S}_r \rightarrow \prod_{i \in [r]} \overline{\mathcal{F}_i}$  and  $g : \mathcal{I}(\mathcal{H}_1, \dots, \mathcal{H}_r) \rightarrow \mathcal{S}_r$  such that the following conditions hold:*

- (i) *If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum |S_i| \leq Cp|V|$ ;*
- (ii) *for every  $(I_1, \dots, I_r) \in \mathcal{I}(\mathcal{H}_1, \dots, \mathcal{H}_r)$ , we have that  $S \subseteq (I_1, \dots, I_r) \subseteq S \cup f(S)$  where  $S := g(I_1, \dots, I_r)$ .*

#### 4. APPLICATIONS OF THE CONTAINER METHOD TO $(\mathcal{L}, r)$ -FREE SETS

In this section we will prove Theorems 1.11 and 1.13 by using the container theorem for  $r$ -tuples of disjoint independent sets, applied with irredundant partition regular matrices  $A$ . (Recall that we assume  $A$  is an  $\ell \times k$  integer matrix where  $k \geq \ell$  of full rank  $\ell$  throughout.) Suppose that we have a  $k$ -uniform hypergraph  $\mathcal{H}$  whose vertex set is a subset of  $\mathbb{N}$  and where the edges correspond to the  $k$ -distinct solutions of  $\mathcal{L}$ . Then in this setting, an  $(\mathcal{L}, r)$ -free set is precisely an  $r$ -tuple of disjoint independent sets in  $\mathcal{H}$ .

Theorems 1.11 and 1.13 will be deduced from a container theorem, Theorem 4.7, which in turn follows from Theorem 3.2. Theorem 4.7 actually holds for a class of irredundant matrices of which partition regular matrices are a subclass. Let  $(*)$  be the following matrix property:

$(*)$  Under Gaussian elimination  $A$  does not have any row which consists of precisely two non-zero rational entries.

All irredundant partition regular matrices  $A$  satisfy  $(*)$ . So instead, we will prove the following strengthenings of Theorems 1.11 and 1.13.

**Theorem 4.1.** *For all irredundant full rank matrices  $A$  which satisfy  $(*)$ , all positive integers  $r$  and all  $\delta > 0$ , there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res}([n]_p, (\mathcal{L}, r)\text{-Rado})}{|[n]_p|} = 1 - \frac{\mu(n, \mathcal{L}, r)}{n} \pm \delta \right] = 1 \quad \text{if } p > Cn^{-1/m(A)}.$$

**Theorem 4.2.** *Let  $A$  be an irredundant matrix of full rank which satisfies  $(*)$  and let  $r \in \mathbb{N}$  be fixed. There are  $2^{\mu(n, \mathcal{L}, r) + o(n)}$   $(\mathcal{L}, r)$ -free subsets of  $[n]$ .*

Given a system of linear equations  $\mathcal{L}$ , a *strongly  $\mathcal{L}$ -free* subset of  $[n]$  is a subset that contains no solution to  $\mathcal{L}$ . Although this is not quite the same definition as  $\mathcal{L}$ -free, we remark that Theorem 4.2 implies a result of Green [24, Theorem 9.3] in the case where  $k \geq 3$ , on the number of strongly  $\mathcal{L}$ -free subsets of  $[n]$  for homogeneous linear equations  $\mathcal{L}$ .

**Additional note.** As mentioned in the introduction, Spiegel [58] independently proved the case  $r = 1$  of Theorem 4.1. (Note in [58] this result is mentioned in terms of *abundant* matrices  $A$ . That is every  $\ell \times (k - 2)$  submatrix of  $A$  has rank  $\ell$ . But this is clearly equivalent to  $(*)$  in the case of irredundant full rank matrices.)

**4.1. Matrices which satisfy  $(*)$ .** First we prove that irredundant partition regular matrices are a strict subclass of irredundant matrices which satisfy  $(*)$ .

Suppose that an irredundant matrix  $A$  does not satisfy  $(*)$ . Then for all solutions  $(x_1, \dots, x_k)$  to  $\mathcal{L}$ , there exists a pair  $ij \in \binom{[k]}{2}$  and non-zero rationals  $\alpha, \beta$  such that  $\alpha x_i = \beta x_j$ . If  $\alpha = \beta$  then no solution to  $\mathcal{L}$  is  $k$ -distinct and so  $A$  is redundant, a contradiction. Otherwise, without loss of generality, assume that  $\alpha > \beta$ , and devise the following 2-colouring of  $\mathbb{N}$ : greedily colour the numbers  $\{1, 2, 3, \dots\}$  so that when colouring  $x$ , we always give it a different colour to  $\beta x / \alpha$  (if  $\beta x / \alpha \in \mathbb{N}$ ). Such a colouring ensures that no solution to  $\mathcal{L}$  is monochromatic, and so  $A$  is not partition regular.

Note that the converse is not true. An  $\ell \times k$  matrix with columns  $a^{(1)}, \dots, a^{(k)}$  satisfies the *columns property* if there is a partition of  $[k]$ , say  $[k] = D_1 \cup \dots \cup D_t$  such that

$$\sum_{i \in D_1} a^{(i)} = 0$$

and for every  $r \in [t]$  we have

$$\sum_{i \in D_r} a^{(i)} \in \langle a^{(j)} : j \in D_1 \cup \dots \cup D_{r-1} \rangle.$$

Rado's theorem [45] states that a matrix is partition regular if and only if it satisfies the columns property. Now, for example  $A := \begin{pmatrix} 2 & 2 & -1 \end{pmatrix}$  clearly satisfies (\*), and additionally does not have the columns property, so is not partition regular.

Call a matrix  $A$   $r$ -regular if all  $r$ -colourings of  $\mathbb{N}$  yield a monochromatic solution to  $\mathcal{L}$ . Observe that a matrix is  $r$ -regular for all  $r \in \mathbb{N}$  if and only if it is partition regular. The argument above actually implies that if an irredundant matrix  $A$  is 2-regular, then it satisfies (\*). So Theorems 4.1 and 4.2 consider *all* pairs  $(A, r)$  such that  $A$  is an irredundant  $r$ -regular matrix and  $r \geq 2$ .

**4.2. Useful matrix lemmas.** Before we can prove our container result (Theorem 4.7), we require some matrix lemmas. Note that all of these lemmas hold for irredundant matrices which satisfy (\*). As a consequence, Theorem 1.8 was actually implicitly proven for irredundant matrices which satisfy (\*), since in [49] the only necessity of the matrix being partition regular was so that the results stated below could be applied.

Recall the definition of  $m(A)$  given by (1.5). Parts (i) and (ii) of the following proposition were verified for irredundant partition regular matrices by Rödl and Ruciński (see Proposition 2.2 in [49]). In fact their result easily extends to matrices which satisfy (\*). We give the full proof for completeness, and add further facts ((iii)–(v)) which will be useful in the proof of Theorem 4.7.

**Proposition 4.3.** *Let  $A$  be an  $\ell \times k$  irredundant matrix of full rank  $\ell$  which satisfies (\*). Then for every  $W \subseteq [k]$ , the following hold.*

- (i) *If  $|W| = 1$ , then  $\text{rank}(A_{\overline{W}}) = \ell$ .*
- (ii) *If  $|W| \geq 2$ , then  $\ell - \text{rank}(A_{\overline{W}}) + 2 \leq |W|$ .*
- (iii) *If  $|W| \geq 2$ , then*

$$-|W| - \text{rank}(A_{\overline{W}}) \leq -\ell - 1 - \frac{|W| - 1}{m(A)}.$$

*Furthermore,*

- (iv)  $k \geq \ell + 2$ ;
- (v)  $m(A) > 1$ .

*Proof.* For (i), suppose that  $\text{rank}(A_{\overline{W}}) = \ell - 1$  for some  $W \subseteq [k]$  with  $|W| = 1$ . Since  $A_{\overline{W}}$  is an  $\ell \times (k - 1)$  matrix of rank  $\ell - 1$ , under Gaussian elimination it must contain a row of zeroes. Hence  $A$  under Gaussian elimination contains a row with at most one non-zero entry. If there is one, then there are no positive solutions to  $\mathcal{L}$ , which contradicts  $A$  being irredundant. If there are none, then  $A$  does not have rank  $\ell$ , also a contradiction.

For (ii) proceed by induction on  $|W|$ . Assume first that there is a  $W \subseteq [k]$  with  $|W| = 2$ , such that  $\text{rank}(A_{\overline{W}}) < \ell$ . Using a similar argument to (i), under Gaussian elimination  $A$  contains a row with at most two non-zero entries. If there are two non-zero entries this contradicts  $A$  satisfying (\*). Otherwise we again get a contradiction to either  $A$  being irredundant or of rank  $\ell$ . Assume now that  $|W| \geq 3$  and that the statement holds for  $|W| - 1$ . The rank of a matrix drops by at most one when a column is deleted, hence the required inequality follows by induction.

For (iii), note that for  $|W| \geq 2$ , by definition we have  $m(A) \geq (|W| - 1) / (|W| - 1 + \text{rank}(A_{\overline{W}}) - \ell)$ . This can be rearranged to give the required inequality. For (iv), suppose that  $k \leq \ell + 1$ . Then under Gaussian elimination  $A$  must have a row with at most two non-zero entries. This leads to the same contradiction as the base case of the induction for (ii). For (v), take  $W = [k]$ . Then by definition  $m(A) \geq (k - 1) / (k - \ell - 1) > 1$ , where the second inequality follows since the denominator is positive by (iv).  $\square$

The following supersaturation lemma follows easily from the (1-colour) removal lemma proved for integer matrices by Král', Serra and Vena (Theorem 2 in [38]).

**Lemma 4.4.** *Let  $A$  be an  $\ell \times k$  integer matrix of rank  $\ell$ . For every  $\delta > 0$  and  $r \in \mathbb{N}$ , there exist  $n_0, \varepsilon > 0$  with the following property. Suppose  $n \geq n_0$  is an integer and  $X \subseteq [n]$  is  $r$ -coloured, and  $|X| \geq \mu(n, \mathcal{L}, r) + \delta n$ . Then there are more than  $\varepsilon n^{k-\ell}$  monochromatic  $k$ -distinct solutions to  $\mathcal{L}$  in  $X$ .*

Finally we need the following well known result (and a simple corollary of it), which gives a useful upper bound on the number of solutions to a system of linear equations. Note that in this lemma only, we *do not* assume  $A$  to be necessarily of full rank (as we will apply the result directly to matrices formed by deleting columns from our original matrix of full rank).

**Lemma 4.5.** *For an  $\ell \times k$  matrix  $A$  not necessarily of full rank, an  $\ell$ -dimensional integer vector  $b$  and a set  $X \subseteq [n]$ , the system  $Ax = b$  has at most  $|X|^{k-\text{rank}(A)}$  solutions in  $X$ .*

*Proof.* Let  $m := \text{rank}(A)$ . Proceed by induction on  $k$ . If  $k = 1$ , then  $m = 1$  or  $m = 0$ . If  $m = 1$ , then there is a unique solution to  $Ax = b$ , so  $Ax = b$  has at most  $1 = |X|^{k-m}$  solution in  $X$ . If  $m = 0$ , then  $A$  is the zero matrix, and so each element in  $X$  could be a solution (if  $b = 0$ ), but then  $Ax = b$  has at most  $|X| = |X|^{k-m}$  solutions in  $X$ .

For the inductive step, pick any  $c \in [k]$  and fix some  $x_c \in X$ . Form a new system of linear equations  $A'x' = b'$ , where  $A'$  is formed from  $A := (a_{ij})$  by removing the  $c^{\text{th}}$  column, and  $b' := (b'_1, \dots, b'_\ell)$  is formed from  $b := (b_1, \dots, b_\ell) \in \mathbb{Z}^\ell$  by setting  $b'_r := b_r - a_{rc}x_c$ . (So  $x := (x_1, \dots, x_k)$  is a solution to  $Ax = b$  if and only if  $x' := (x_1, \dots, x_{c-1}, x_{c+1}, \dots, x_k)$  is a solution to  $A'x' = b'$ .)

Note that  $A'$  must have rank  $m$  or  $m - 1$ . If  $A'$  has rank  $m$  then by the induction hypothesis  $A'x' = b'$  has at most  $|X|^{(k-1)-\text{rank}(A')} = |X|^{k-m-1}$  solutions in  $X$ . Since there are  $|X|$  valid choices for  $x_c$ , there are at most  $|X|^{k-m}$  solutions to  $Ax = b$  in  $X$ , as required.

Now suppose  $A'$  has rank  $m - 1$ . Then under Gaussian elimination,  $A'$  and  $A$  have  $\ell - m + 1$ , and respectively  $\ell - m$  rows consisting entirely of zeroes, and in particular,  $A$  has a row with precisely one non-zero entry which is in the  $c^{\text{th}}$  column. Hence there is at most one value  $x_c$  can take in any solution  $x = (x_1, \dots, x_k)$  to  $Ax = b$ . So for this choice of  $x_c$ ,  $Ax = b$  and  $A'x' = b'$  have precisely the same number of solutions in  $X$ . Since  $A'$  is an  $\ell \times (k - 1)$  matrix of rank  $m - 1$ , the induction hypothesis implies that there are  $|X|^{(k-1)-(m-1)} = |X|^{k-\text{rank}(A)}$  solutions to  $A'x' = b'$  and thus  $Ax = b$ , as desired.  $\square$

**Corollary 4.6.** *Consider an  $\ell \times k$  matrix  $A$  of rank  $\ell$ , a set  $X \subseteq [n]$  and an integer  $1 \leq t \leq k$ . Fix distinct  $y_1, \dots, y_t \in X$  and consider any  $W = \{s_1, \dots, s_t\} \subseteq [k]$ . The system  $Ax = 0$  has at most  $|X|^{k-t-\text{rank}(A_{\overline{W}})}$  solutions  $(x_1, \dots, x_k)$  in  $X$  for which  $x_{s_j} = y_j$  for each  $j \in [t]$ . Moreover, if  $A$  is irredundant and satisfies  $(*)$  and  $t = 1$ , then the system  $Ax = 0$  has at most  $|X|^{k-\ell-1}$  solutions  $(x_1, \dots, x_k)$  in  $X$  for which  $x_{s_1} = y_1$ .*

*Proof.* Write  $A = (a_{ij})$ . Consider the system of linear equations  $A_{\overline{W}}x' = b$  where, for each  $r \in [\ell]$ , the  $r^{\text{th}}$  term in  $b$  is

$$b_r := - \sum_{s_j \in W} a_{rs_j} y_j.$$

Now by Lemma 4.5 the system of linear equations  $A_{\overline{W}}x' = b$  has at most  $|X|^{k-t-\text{rank}(A_{\overline{W}})}$  solutions in  $X$ . The first part of the corollary then follows since all solutions  $(x_1, \dots, x_k)$  to  $Ax = b$  with  $x_{s_j} = y_j$  for each  $j \in [t]$ , rise from a solution  $x'$  to  $A_{\overline{W}}x' = b$ . For the second part, if  $A$  is irredundant and satisfies  $(*)$  and  $t = 1$ , then by Proposition 4.3(i), we have  $\text{rank}(A_{\overline{W}}) = \ell$  and so the result follows.  $\square$

**4.3. A container theorem for tuples of  $\mathcal{L}$ -free sets.** Recall that an  $\mathcal{L}$ -free set is simply an  $(\mathcal{L}, 1)$ -free set. Let  $\mathcal{I}_r(n, \mathcal{L})$  denote the set of all ordered  $r$ -tuples  $(X_1, \dots, X_r) \in \mathcal{P}([n]^r)$  so that each  $X_i$  is  $\mathcal{L}$ -free and  $X_i \cap X_j = \emptyset$  for all distinct  $i, j \in [r]$ . Note that any  $(\mathcal{L}, r)$ -free subset  $X$  of

$[n]$  has a partition  $X_1, \dots, X_r$  so that  $(X_1, \dots, X_r) \in \mathcal{I}_r(n, \mathcal{L})$ . We now prove a container theorem for the elements of  $\mathcal{I}_r(n, \mathcal{L})$ .

**Theorem 4.7.** *Let  $r \in \mathbb{N}$  and  $0 < \delta < 1$ . Let  $A$  be an  $\ell \times k$  irredundant matrix of full rank  $\ell$  which satisfies  $(*)$ . Then there exists  $D > 0$  such that the following holds. For all  $n \in \mathbb{N}$ , there is a collection  $\mathcal{S}_r \subseteq \mathcal{P}([n]^r)$  and a function  $f : \mathcal{S}_r \rightarrow \mathcal{P}([n]^r)$  such that:*

- (i) *For all  $(I_1, \dots, I_r) \in \mathcal{I}_r(n, \mathcal{L})$ , there exists  $S \in \mathcal{S}_r$  such that  $S \subseteq (I_1, \dots, I_r) \subseteq f(S)$ .*
- (ii) *If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum_{i \in [r]} |S_i| \leq Dn^{\frac{m(A)-1}{m(A)}}$ .*
- (iii) *Every  $S \in \mathcal{S}_r$  satisfies  $S \in \mathcal{I}_r(n, \mathcal{L})$ .*
- (iv) *Given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$ , write  $f(S) =: (f(S_1), \dots, f(S_r))$ . Then*
  - (a) *for each  $1 \leq i \leq r$ ,  $f(S_i)$  contains at most  $\delta n^{k-\ell}/r$   $k$ -distinct solutions to  $\mathcal{L}$ ; and*
  - (b)  *$|\cup_{i \in [r]} f(S_i)| \leq \mu(n, \mathcal{L}, r) + \delta n$ .*

We emphasise that (iv)(b) does not necessarily guarantee  $\sum_{i \in [r]} |f(S_i)| \leq \mu(n, \mathcal{L}, r) + \delta n$ . Rather it ensures at most  $\mu(n, \mathcal{L}, r) + \delta n$  elements of  $[n]$  appear in at least one of the co-ordinates of  $f(S)$ . This property is crucial for our applications.

*Proof.* It suffices to prove the theorem when  $n$  is sufficiently large; otherwise we can set  $\mathcal{S}_r$  to be  $\mathcal{I}_r(n, \mathcal{L})$ ; set  $f$  to be the identity function and choose  $D$  to be large.

Let  $0 < \delta < 1$  and  $r \in \mathbb{N}$  be given and apply Lemma 4.4 to obtain  $n_0, \varepsilon > 0$ . Without loss of generality we may assume  $\varepsilon \leq \delta$ . Let

$$\varepsilon' := \frac{\varepsilon}{2r} \quad \text{and} \quad c := \frac{k!}{\varepsilon'}.$$

Apply Theorem 3.2 with  $k, r, c, \varepsilon'$  playing the roles of  $k, r, c, \varepsilon$  respectively to obtain  $D > 0$ . Increase  $n_0$  if necessary so that  $0 < 1/n_0 \ll 1/D, 1/k, 1/r, \varepsilon, \delta$  and let  $n \geq n_0$  be an integer.

If  $n \in \mathbb{N}$  is such that  $[n]$  contains at most  $\varepsilon' n^{k-\ell}$   $k$ -distinct solutions to  $\mathcal{L}$ , then we can set  $\mathcal{S}_r$  to consist only of  $\emptyset^r = (\emptyset, \dots, \emptyset)$  and  $f(\emptyset^r) := [n]^r$  (and  $D := 1$ , say). Then (i)–(iii) and (iv)(a) hold, and by Lemma 4.4,  $|[n]| \leq \mu(n, \mathcal{L}, r) + \delta n$  so we are done. Thus we may assume that there are at least  $\varepsilon' n^{k-\ell}$   $k$ -distinct solutions to  $\mathcal{L}$  in  $[n]$ .

Let  $\mathcal{H}_n$  be the hypergraph with  $V(\mathcal{H}_n) := [n]$  and an edge set which consists of all  $k$ -distinct solutions to  $\mathcal{L}$  in  $[n]$ . Observe that  $\mathcal{H}_n$  is  $k$ -uniform and an independent set in  $\mathcal{H}_n$  is an  $\mathcal{L}$ -free set. Note by our assumption on the number of  $k$ -distinct solutions to  $\mathcal{L}$  and by Lemma 4.5, we have

$$(4.1) \quad \varepsilon' n^{k-\ell} \leq e(\mathcal{H}_n) \leq n^{k-\ell}.$$

Let  $\mathcal{F}_n := \{F \subseteq V(\mathcal{H}_n) : e(\mathcal{H}_n[F]) \geq \varepsilon' e(\mathcal{H}_n)\}$ . We claim that  $\mathcal{H}_n$  and  $\mathcal{F}_n$  satisfy the hypotheses of Theorem 3.2 with parameters chosen as above with

$$p = p(n) := n^{-1/m(A)}.$$

Clearly  $\mathcal{F}_n$  is increasing, and  $\mathcal{H}_n$  is  $(\mathcal{F}_n, \varepsilon')$ -dense. By Lemma 4.5, a set  $F \subseteq V(\mathcal{H}_n)$  contains at most  $|F|^{k-\ell}$  solutions to  $\mathcal{L}$  (so  $e(\mathcal{H}_n[F]) \leq |F|^{k-\ell}$ ). Hence for all  $F \in \mathcal{F}_n$ , we have

$$|F| \geq e(\mathcal{H}_n[F])^{\frac{1}{k-\ell}} \geq (\varepsilon' e(\mathcal{H}_n))^{\frac{1}{k-\ell}} \stackrel{(4.1)}{\geq} ((\varepsilon')^2 n^{k-\ell})^{\frac{1}{k-\ell}} \geq \varepsilon' n$$

where the last inequality follows by Proposition 4.3(iv).

For each  $j \in [k]$ , we wish to bound the number of hyperedges containing some  $\{y_1, \dots, y_j\} \subseteq V(\mathcal{H}_n)$ . Suppose  $(x_1, \dots, x_k)$  is a  $k$ -distinct solution to  $\mathcal{L}$  so that  $\{y_1, \dots, y_j\} \subseteq \{x_1, \dots, x_k\}$ . There are  $k!/(k-j)!$  choices for picking the  $j$  roles the  $y_i$  play in  $(x_1, \dots, x_k)$ . Let  $W$  be one such choice for the set of indices of the  $x_i$  used by  $\{y_1, \dots, y_j\}$ . In this case, Corollary 4.6 implies there are at

most  $n^{k-j-\text{rank}(A_{\overline{W}})}$  such solutions to  $\mathcal{L}$ , and if  $j = 1$ , there are at most  $n^{k-\ell-1}$  such solutions. So for  $j = 1$  this yields

$$\deg_{\mathcal{H}_n}(y_1) \leq kn^{k-\ell-1} \stackrel{(4.1)}{\leq} \frac{k}{\varepsilon'} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \leq c \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

For  $j \geq 2$ , by Proposition 4.3(iii) we have  $k - j - \text{rank}(A_{\overline{W}}) \leq k - \ell - 1 - (j - 1)/m(A)$ , and hence we have

$$\deg_{\mathcal{H}_n}(\{y_1, \dots, y_j\}) \leq k!n^{k-\ell-1-(j-1)/m(A)} \leq \frac{k!}{\varepsilon'^j} p^{j-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} = cp^{j-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

Since  $\{y_1, \dots, y_j\}$  was arbitrary, we therefore have  $\Delta_j(\mathcal{H}_n) \leq cp^{j-1}e(\mathcal{H}_n)/v(\mathcal{H}_n)$ , as required. We have therefore shown that  $\mathcal{H}_n$  and  $\mathcal{F}_n$  satisfy the hypotheses of Theorem 3.2.

Then Theorem 3.2 implies that there exists a family  $\mathcal{S}_r \subseteq \mathcal{P}(V(\mathcal{H}_n)^r) = \mathcal{P}([n]^r)$  and functions  $f' : \mathcal{S}_r \rightarrow (\overline{\mathcal{F}_n})^r$  and  $g : \mathcal{I}_r(\mathcal{H}_n) \rightarrow \mathcal{S}_r$  such that the following conditions hold:

- (a) If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum_{i \in [r]} |S_i| \leq Dpn$ ;
- (b) every  $S \in \mathcal{S}_r$  satisfies  $S \in \mathcal{I}_r(\mathcal{H}_n)$ ;
- (c) for every  $(I_1, \dots, I_r) \in \mathcal{I}_r(\mathcal{H}_n)$ , we have that  $S \subseteq (I_1, \dots, I_r) \subseteq S \cup f'(S)$  where  $S := g(I_1, \dots, I_r)$ .

Note that  $\mathcal{I}_r(\mathcal{H}_n) = \mathcal{I}_r(n, \mathcal{L})$ . For each  $S \in \mathcal{S}_r$ , define

$$f(S) := S \cup f'(S).$$

So  $f : \mathcal{S}_r \rightarrow \mathcal{P}([n]^r)$ . Thus, (a)–(c) immediately imply that (i)–(iii) hold.

Given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  write  $f(S) =: (f(S_1), \dots, f(S_r))$  and  $f'(S) =: (f'(S_1), \dots, f'(S_r))$ . (Note the slight abuse of the use of the  $f$  and  $f'$  notation here.) For each  $i \in [r]$ , we have that  $f'(S_i) \in \overline{\mathcal{F}_n}$ , so it contains at most  $\varepsilon' n^{k-\ell}$   $k$ -distinct solutions to  $\mathcal{L}$ . By Corollary 4.6, the number of  $k$ -distinct solutions to  $\mathcal{L}$  in  $[n]$  that use at least one element from  $S_i$  is at most  $kn^{k-\ell-1}|S_i|$ . Further,

$$kn^{k-\ell-1}|S_i| \leq kDpn^{k-\ell} \leq \varepsilon' n^{k-\ell}.$$

Here, the first inequality holds by (a), and the second since  $p = n^{-1/m(A)}$  and  $m(A) > 0$  by Proposition 4.3(v). Thus, in total  $f(S_i) = S_i \cup f'(S_i)$  contains at most  $2\varepsilon' n^{k-\ell} \leq \delta n^{k-\ell}/r$   $k$ -distinct solutions to  $\mathcal{L}$ , so (iv)(a) holds.

In fact, the argument above implies that there is an  $r$ -colouring of the set  $\cup_{i \in [r]} f(S_i)$  so that there are at most  $2\varepsilon' r n^{k-\ell} = \varepsilon n^{k-\ell}$  monochromatic  $k$ -distinct solutions to  $\mathcal{L}$  in  $\cup_{i \in [r]} f(S_i)$ . Hence, Lemma 4.4 ensures (iv)(b), as desired.  $\square$

**4.4. The number of  $(\mathcal{L}, r)$ -free subsets of  $[n]$ .** Our first application of Theorem 4.7 yields an enumeration result (Theorem 4.2) for the number of  $(\mathcal{L}, r)$ -free subsets of  $[n]$ .

*Proof of Theorem 4.2.* By definition of  $\mu(n, \mathcal{L}, r)$  there are at least  $2^{\mu(n, \mathcal{L}, r)}$   $(\mathcal{L}, r)$ -free subsets of  $[n]$ . So it suffices to prove the upper bound.

For this, note that we may assume  $n$  is sufficiently large. Let  $0 < \delta < 1$  be arbitrary and let  $D > 0$  be obtained from Theorem 4.7 applied to  $A$  with parameter  $\delta$ . We obtain a collection  $\mathcal{S}_r$  and function  $f$  as in Theorem 4.7. Consider any  $(\mathcal{L}, r)$ -free subset  $X$  of  $[n]$ . Note that  $X$  has a partition  $X_1, \dots, X_r$  so that  $(X_1, \dots, X_r) \in \mathcal{I}_r(n, \mathcal{L})$ . So by Theorem 4.7(i) this means there is some  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  so that  $X \subseteq \cup_{i \in [r]} f(S_i)$ .



Further, given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$ , we have that  $|\cup_{i \in [r]} f(S_i)| \leq \mu(n, \mathcal{L}, r) + \delta n$ . Thus, each such  $\cup_{i \in [r]} f(S_i)$  contains at most  $2^{\mu(n, \mathcal{L}, r) + \delta n}$   $(\mathcal{L}, r)$ -free subsets of  $[n]$ . Note that, by Theorem 4.7(ii),

$$|\mathcal{S}_r| \leq \left( Dn \sum_{s=0}^{\frac{m(A)-1}{m(A)}} \binom{n}{s} \right)^r < 2^{\delta n},$$

where the last inequality holds since  $n$  is sufficiently large.

Altogether, this implies that the number of  $(\mathcal{L}, r)$ -free subsets of  $[n]$  is at most

$$2^{\delta n} \times 2^{\mu(n, \mathcal{L}, r) + \delta n} = 2^{\mu(n, \mathcal{L}, r) + 2\delta n}.$$

Since the choice of  $0 < \delta < 1$  was arbitrary this proves the theorem.  $\square$

**4.5. The resilience of being  $(\mathcal{L}, r)$ -Rado.** Recall that the resilience of  $S$  with respect to  $\mathcal{P}$ ,  $\text{res}(S, \mathcal{P})$ , is the minimum number  $t$  such that by deleting  $t$  elements from  $S$ , one can obtain a set not having  $\mathcal{P}$ . In this section we will determine  $\text{res}([n]_p, (\mathcal{L}, r)\text{-Rado})$  for irredundant matrices  $A$  which satisfy  $(*)$ . We now use Theorem 4.7 to deduce Theorem 4.1.

*Proof of Theorem 4.1.* Let  $0 < \delta < 1$ ,  $r \in \mathbb{N}$  and  $A$  be a matrix as in the statement of the theorem. Given  $n$ , if  $p > n^{-1/m(A)}$  then since  $m(A) > 1$  by Proposition 4.3(v), Proposition 2.1 implies that, w.h.p.,

$$(4.2) \quad |[n]_p| = \left( 1 \pm \frac{\delta}{4} \right) pn.$$

We first show that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\text{res}([n]_p, (\mathcal{L}, r)\text{-Rado})}{|[n]_p|} \leq 1 - \frac{\mu(n, \mathcal{L}, r)}{n} + \delta \right] = 1 \quad \text{if } p > n^{-1/m(A)}.$$

For this, we must show that the probability of the event that there exists a set  $S \subseteq [n]_p$  such that  $|S| \geq (\mu(n, \mathcal{L}, r)/n - \delta)|[n]_p|$  and  $S$  is  $(\mathcal{L}, r)$ -free, tends to one as  $n$  tends to infinity. This indeed follows: Let  $T$  be an  $(\mathcal{L}, r)$ -free subset of  $[n]$  of maximum size  $\mu(n, \mathcal{L}, r)$ . Then, by Proposition 2.1, w.h.p. we have  $|T \cap [n]_p| = (\mu(n, \mathcal{L}, r)/n \pm \delta)|[n]_p|$ , and  $T \cap [n]_p$  is  $(\mathcal{L}, r)$ -free, as required.

For the remainder of the proof, we will focus on the lower bound, namely that there exists  $C > 0$  such that whenever  $p > Cn^{-1/m(A)}$ ,

$$(4.3) \quad \mathbb{P} \left[ \text{res}([n]_p, (\mathcal{L}, r)\text{-Rado}) \geq \left( 1 - \frac{\mu(n, \mathcal{L}, r)}{n} - \delta \right) |[n]_p| \right] \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Suppose  $n$  is sufficiently large. Apply Theorem 4.7 with parameters  $r, \delta/8, A$  to obtain  $D > 0$ , a collection  $\mathcal{S}_r \subseteq \mathcal{P}([n]^r)$  and a function  $f$  satisfying (i)–(iv). Now choose  $C$  such that  $0 < 1/C \ll 1/D, \delta, 1/r$ . Let  $p \geq Cn^{-1/m(A)}$ .

Since (4.2) holds with high probability, to prove (4.3) holds it suffices to show that the probability  $[n]_p$  contains an  $(\mathcal{L}, r)$ -free subset of size at least  $(\frac{\mu(n, \mathcal{L}, r)}{n} + \delta/2)np$  tends to one as  $n$  tends to infinity.

Suppose that  $[n]_p$  does contain an  $(\mathcal{L}, r)$ -free subset  $I$  of size at least  $(\frac{\mu(n, \mathcal{L}, r)}{n} + \delta/2)np$ . Note that  $I$  has a partition  $I_1, \dots, I_r$  so that  $(I_1, \dots, I_r) \in \mathcal{I}_r(n, \mathcal{L})$ . Further, there is some  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  such that  $S \subseteq (I_1, \dots, I_r) \subseteq f(S)$ . Thus,  $[n]_p$  must contain  $\cup_{i \in [r]} S_i$  as well as at least  $(\frac{\mu(n, \mathcal{L}, r)}{n} + \delta/4)np$  elements from  $(\cup_{i \in [r]} f(S_i)) \setminus (\cup_{i \in [r]} S_i)$ . (Note here we are using that  $|\cup_{i \in [r]} S_i| \leq \delta np/4$ , which holds by Theorem 4.7(ii) and since  $0 < 1/C \ll 1/D, \delta$ .) Writing  $s := |\cup_{i \in [r]} S_i|$ , the probability  $[n]_p$  contains  $\cup_{i \in [r]} S_i$  is  $p^s$ . Note that  $|\cup_{i \in [r]} f(S_i) \setminus (\cup_{i \in [r]} S_i)| \leq \mu(n, \mathcal{L}, r) + \delta n/8$  by Theorem 4.7(iv)(b). So by the first part of Proposition 2.1, the probability  $[n]_p$  contains at least  $(\frac{\mu(n, \mathcal{L}, r)}{n} + \delta/4)np$  elements from  $(\cup_{i \in [r]} f(S_i)) \setminus (\cup_{i \in [r]} S_i)$ , is at most  $\exp(-\delta^2 np/256)$ .

Write  $N := n^{(m(A)-1)/m(A)}$  and  $\gamma := \delta^2/256$ . Given some  $0 \leq s \leq DN$ , there are at most  $r^s \binom{n}{s}$  elements  $(S_1, \dots, S_r) \in \mathcal{S}_r$  such that  $|\cup_{i \in [r]} S_i| = s$ . Indeed, this follows by Theorem 4.7(iii) and since there are  $r^s$  ways to partition a set of size  $s$  into  $r$  classes. Thus, the probability  $[n]_p$  does contain an  $(\mathcal{L}, r)$ -free subset  $I$  of size at least  $(\frac{\mu(n, \mathcal{L}, r)}{n} + \delta/2)np$  is at most

$$\begin{aligned} \sum_{s=0}^{DN} r^s \binom{n}{s} \cdot p^s \cdot e^{-\gamma np} &\leq \sum_{s=0}^{DN} \left(\frac{repn}{s}\right)^s e^{-\gamma np} \leq (DN+1) \left(\frac{repn}{DN}\right)^{DN} e^{-\gamma np} \\ &\leq (DN+1) \left(\frac{reC}{D}\right)^{DN} e^{-\gamma CN} \leq e^{\gamma CN/2} e^{-\gamma CN} = e^{-\gamma CN/2} \end{aligned}$$

which tends to zero as  $n$  tends to infinity. This completes the proof.  $\square$

**4.6. The size of the largest  $(\mathcal{L}, r)$ -free set.** Both as a natural question in itself, and in light of Theorems 4.1 and 4.2, it is of interest to obtain good bounds on  $\mu(n, \mathcal{L}, r)$ . For the rest of this section we assume that  $A$  is a  $1 \times k$  matrix, i.e. we are interested in solutions to a linear equation  $a_1x_1 + \dots + a_kx_k = 0$ . Such  $\mathcal{L}$  are called *translation-invariant* if the coefficients  $a_i$  sum to zero. It is known that  $\mu(n, \mathcal{L}, 1) = o(n)$  if  $\mathcal{L}$  is translation-invariant and  $\mu(n, \mathcal{L}, 1) = \Omega(n)$  otherwise (see [52]). Determining exact bounds remains open in many cases, famously including *progression-free sets* (where  $\mathcal{L}$  is  $x + y = 2z$ ). See [5, 14, 25] for the state-of-the-art lower and upper bounds for this case.

Call  $S \subseteq [n]$  *strongly  $(\mathcal{L}, r)$ -free* if there exists an  $r$ -colouring of  $S$  which contains no monochromatic solutions to  $\mathcal{L}$  of any type (that is, solutions are not required to be  $k$ -distinct). Define  $\mu^*(n, \mathcal{L}, r)$  to be the size of the largest strongly  $(\mathcal{L}, r)$ -free subset  $S \subseteq [n]$ . Note that for any density regular matrix  $A$ ,  $(x, \dots, x)$  is a solution to  $\mathcal{L}$  for all  $x \in [n]$  (as observed by Frankl, Graham and Rödl [20, Fact 4]) and so we have  $\mu^*(n, \mathcal{L}, r) = 0$ . (Note that this result implies that all density regular  $1 \times k$  matrices gives rise to an equation  $\mathcal{L}$  which is translation-invariant.) In fact, if  $A$  is any  $1 \times k$  irredundant integer matrix, then for all  $\varepsilon > 0$  there exists an  $n_0 > 0$  such that for all integers  $n \geq n_0$  we have

$$\mu^*(n, \mathcal{L}, r) \leq \mu(n, \mathcal{L}, r) \leq \mu^*(n, \mathcal{L}, r) + \varepsilon n.$$

This follows from e.g. [38, Theorem 2], since such  $\mathcal{L}$  have  $o(n^{k-\ell})$  non- $k$ -distinct solutions in  $[n]$  (i.e. a solution  $(x_1, \dots, x_k)$  where there is an  $i \neq j$  such that  $x_i = x_j$ ).

Consequently it is equally interesting to study  $\mu^*(n, \mathcal{L}, r)$  in the case when  $\mu(n, \mathcal{L}, r) = \Omega(n)$ . In the case of sum-free sets (where  $\mathcal{L}$  is  $x + y = z$ ), the study of  $\mu^*(n, \mathcal{L}, r)$  is a classical problem of Abbott and Wang [1]. (Note that the only difference between  $\mu(n, \mathcal{L}, r)$  and  $\mu^*(n, \mathcal{L}, r)$  in this case is that  $\mu(n, \mathcal{L}, r)$  allows non-distinct sums  $x + x = z$  whereas  $\mu^*(n, \mathcal{L}, r)$  does not.) Let  $\mu(n, r) := \mu^*(n, \mathcal{L}, r)$  where  $\mathcal{L}$  is  $x + y = z$ . An easy proof shows that  $\mu(n, 1) = \lceil n/2 \rceil$ .

The following definitions help motivate the study of  $\mu(n, r)$  for  $r \geq 2$ . Let  $f(r)$  denote the largest positive integer  $m$  for which there exists a partition of  $[m]$  into  $r$  sum-free sets, and let  $h(r)$  denote the largest positive integer  $m$  for which there exists a partition of  $[m]$  into  $r$  sets which are sum-free modulo  $h(r) + 1$ .

Abbott and Wang [1] conjectured that  $h(r) = k(r)$ , and showed that  $\mu(n, r) \geq n - \lfloor n/(h(r)+1) \rfloor$ . They also proved the following upper bound.

**Theorem 4.8** ([1]). *We have  $\mu(n, r) \leq n - \lfloor cn/((f(r)+1) \log(f(r)+1)) \rfloor$  where  $c := e^{-\gamma} \approx 0.56$  ( $\gamma$  denotes the Euler-Mascheroni constant).*

We provide an alternate upper bound, which is a modification of Hu's [29] proof that  $\mu(n, 2) = n - \lfloor \frac{n}{5} \rfloor$ . (To see why this is a lower bound, consider the set  $\{x \in [n] : x \equiv 1 \text{ or } 4 \pmod{5}\} \cup \{y \in [n] : y \equiv 2 \text{ or } 3 \pmod{5}\}$ .) First we need the following fact. Given  $x \in [n]$  and  $T \subseteq [n]$ , write

$x + T := \{x + y : y \in T\}$ . Given  $S, T \subseteq [n]$ , say that  $T$  is a *difference set* of  $S$  if there exists  $x \in S$  such that  $x + T \subseteq S$ .

**Fact 4.9.** *Let  $n \in \mathbb{N}$  and  $S, T, T' \subseteq [n]$ .*

- (i) *If  $T$  is a difference set of a sum-free set  $S$ , then  $S \cap T = \emptyset$ .*
- (ii) *If  $T'$  is a difference set of  $T$ , and  $T$  is a difference set of  $S$ , then  $T'$  is a difference set of  $S$ .*

*Proof.* If there exists  $x \in S$  such that  $x + T \subseteq S$  and moreover there exists  $y \in S \cap T$ , then  $x + y \in S$ , proving (i). For (ii), suppose that there is  $x' \in T$  and  $x \in S$  such that  $x' + T' \subseteq T$  and  $x + T \subseteq S$ . Then  $x + x' + T' \subseteq S$  and  $x + x' \in x + T \subseteq S$ , proving (ii).  $\square$

**Theorem 4.10.** *We have  $\mu(n, r) \leq n - \lfloor \frac{n}{r!e} \rfloor$ .*

Note that Theorem 4.10 does indeed recover Hu's bound [29] for the case  $r = 2$ .

*Proof.* Fix  $n, r \in \mathbb{N}$ . Let  $\ell(0) := 1$ . For all integers  $i \geq 1$ , define

$$\ell(i) := i! \left( 1 + \sum_{t \in [i]} \frac{1}{t!} \right) = \lfloor i!e \rfloor.$$

Note that  $\ell(i) = i\ell(i-1) + 1$  for all  $i \geq 1$ . Choose the unique  $q \in \mathbb{N} \cup \{0\}$  and  $0 \leq k \leq \ell(r) - 1$  such that  $n = \ell(r)q + k$ . Consider any partition  $S_1 \dot{\cup} \dots \dot{\cup} S_r \dot{\cup} R = [n]$ , where each  $S_i$  is sum-free. We wish to show that  $|R| \geq q$ , since then  $\mu(\ell(r)q + k, r) \leq (\ell(r) - 1)q + k$  and so  $\mu(n, r) \leq n - \lfloor n/\ell(r) \rfloor$ .

Suppose not. We will obtain integers  $\{j_1, \dots, j_r\} = [r]$  and subsets  $D_0, D_1, \dots, D_r$  of  $[n]$  such that the following properties hold for all  $0 \leq i \leq r$ .

- $P_1(i)$   $|D_i| \geq \ell(r - i)q$ ;
- $P_2(i)$   $D_i$  is a difference set of  $S_{j_t}$  for all  $t \in [i]$ ;
- $P_3(i)$   $D_i \cap S_{j_t} = \emptyset$  for all  $t \in [i]$ .

Let  $D_0 := [n]$ . Then  $P_1(0)$  holds by definition, and  $P_2(0)$  and  $P_3(0)$  are vacuous. Suppose, for some  $0 \leq i < r$ , we have obtained distinct  $\{j_1, \dots, j_i\} \subseteq [r]$  and  $D_0, D_1, \dots, D_i$  such that  $P_1(t)$ – $P_3(t)$  hold for all  $t \in [i]$ .

Suppose that  $|D_i \cap \bigcup_{i+1 \leq t \leq r} S_{j_t}| \leq (\ell(r - i) - 1)q$ . Then we have that

$$|D_i \cap R| \stackrel{P_3(i)}{\geq} |D_i| - (\ell(r - i) - 1)q \stackrel{P_1(i)}{\geq} q,$$

a contradiction. So by averaging, there exists  $j_{i+1} \in [r] \setminus \{j_1, \dots, j_i\}$  such that

$$|D_i \cap S_{j_{i+1}}| \geq \left\lceil \frac{(\ell(r - i) - 1)q + 1}{r - i} \right\rceil = \ell(r - i - 1)q + 1.$$

Thus we can write  $D_i \cap S_{j_{i+1}} \supseteq \{s_{i,0} < \dots < s_{i,\ell(r-i-1)q}\}$ . Let  $D_{i+1} := \{s_{i,x} - s_{i,0} : x \in [\ell(r-i-1)q]\}$ . We claim that  $P_1(i+1)$ – $P_3(i+1)$  hold. Property  $P_1(i+1)$  is clear by definition. For  $P_2(i+1)$ , note that  $D_{i+1}$  is a difference set of both  $D_i$  and  $S_{j_{i+1}}$ . Then Fact 4.9(ii) and  $P_2(i)$  imply that additionally  $D_{i+1}$  is a difference set of  $S_{j_t}$  for all  $t \in [i]$ . Fact 4.9(i) implies that  $D_{i+1} \cap S_{j_t} = \emptyset$  for all  $t \in [i+1]$ , proving  $P_3(i+1)$ .

Thus we obtain  $D_r$  satisfying  $P_1(r)$ – $P_3(r)$ . By  $P_1(r)$  and  $P_3(r)$  we have that  $|D_r| \geq \ell(0)q = q$  and  $D_r \subseteq R$ , a contradiction.  $\square$

**4.7. Open Problem.** We conclude the section with an open problem. Recall Hu [29] showed that  $\mu(n, 2) = n - \lfloor \frac{n}{5} \rfloor$ . So in the case when  $\mathcal{L}$  is  $x + y = z$ , Theorem 4.2 implies that there are  $2^{4n/5 + o(n)}$   $(\mathcal{L}, 2)$ -free subsets of  $[n]$ . We believe the error term in the exponent here can be replaced by a constant.

**Conjecture 4.11.** *Let  $\mathcal{L}$  denote  $x + y = z$ . There are  $\Theta(2^{4n/5})$   $(\mathcal{L}, 2)$ -free subsets of  $[n]$ .*

Note that Conjecture 4.11 can be viewed as a 2-coloured analogue of the Cameron–Erdős conjecture [10] which was famously resolved by Green [23] and independently Sapozhenko [53].

## 5. APPLICATIONS OF THE CONTAINER METHOD TO GRAPH RAMSEY THEORY

In this section we answer some questions in hypergraph Ramsey theory, introduced in Sections 1.1 and 1.2. How many  $n$ -vertex hypergraphs are not Ramsey, and what does a typical such hypergraph look like? How dense must the Erdős–Rényi random hypergraph be to have the Ramsey property with high probability, and above this threshold, how strongly does it possess the Ramsey property?

Our main results here are applications of the asymmetric container theorem (Theorem 3.3). For arbitrary  $k$ -uniform hypergraphs  $H_1, \dots, H_r$ , we first prove Theorem 5.11, a container theorem for non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -uniform hypergraphs. To see how one might prove such a theorem, observe that, if  $\mathcal{H}_i$  is the hypergraph of copies of  $H_i$  on  $n$  vertices (i.e. vertices correspond to  $k$ -subsets of  $[n]$ , and edges correspond to copies of  $E(H_i)$ ; see Definition 5.9), then every non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -uniform hypergraph  $G$  corresponds to a set in  $\mathcal{I}(\mathcal{H}_1, \dots, \mathcal{H}_r)$ . We then use Theorem 5.11 to:

- (1) count the number of  $k$ -uniform hypergraphs on  $n$  vertices which are not  $(H_1, \dots, H_r)$ -Ramsey (Theorem 1.12);
- (2) determine the global resilience of  $G_{n,p}^{(k)}$  with respect to the property of being  $(H_1, \dots, H_r)$ -Ramsey (Theorem 1.7). That is, we show that there is a constant  $C$  such that whenever  $p \geq Cn^{-1/m_k(H_1)}$ , we obtain a function  $t$  of  $n$  and  $p$  such that, with high probability, any subhypergraph  $G \subseteq G_{n,p}^{(k)}$  with  $e(G) > t + \Omega(n^k)$  is  $(H_1, \dots, H_r)$ -Ramsey, while there is some  $G' \subseteq G_{n,p}^{(k)}$  with  $e(G') < t - o(n^k)$  which is *not*  $(H_1, \dots, H_r)$ -Ramsey.
- (3) As a corollary of (3), we see that, whenever  $p \geq Cn^{-1/m_k(H_1)}$ , the random hypergraph  $G_{n,p}^{(k)}$  is  $(H_1, \dots, H_r)$ -Ramsey with high probability.

The statements of (1)–(3) all involve a common parameter: the maximum size  $\text{ex}^r(n; H_1, \dots, H_r)$  of an  $n$ -vertex  $k$ -uniform hypergraph which is not  $(H_1, \dots, H_r)$ -Ramsey. For this reason, we generalise the classical supersaturation result of Erdős and Simonovits [17] to show that any  $n$ -vertex  $k$ -uniform hypergraph  $G$  with at least  $\text{ex}^r(n; H_1, \dots, H_r) + \Omega(n^k)$  edges is somehow ‘strongly’  $(H_1, \dots, H_r)$ -Ramsey. In the graph case, an old result of Burr, Erdős and Lovász [8] allows us to quite accurately determine  $\text{ex}^r(n; H_1, \dots, H_r)$ .

**5.1. Definitions and notation.** In this section,  $k \geq 2$  is an integer and we use  $k$ -*graph* as shorthand for  $k$ -uniform hypergraph. Recall from Section 1.1 that, given  $r \in \mathbb{N}$  and a  $k$ -graph  $G$ , an  $r$ -*colouring* is a function  $\sigma : E(G) \rightarrow [r]$ . Given  $k$ -graphs  $H_1, \dots, H_r$ , we say that  $\sigma$  is  $(H_1, \dots, H_r)$ -*free* if  $\sigma^{-1}(i)$  is  $H_i$ -free for all  $i \in [r]$ . Then  $G$  is  $(H_1, \dots, H_r)$ -*Ramsey* if it has no  $(H_1, \dots, H_r)$ -free  $r$ -colouring.

Given an integer  $\ell \geq k$ , denote by  $K_\ell^{(k)}$  the complete  $k$ -graph on  $\ell$  vertices. A  $k$ -graph  $H$  is  $k$ -*partite* if the vertices of  $H$  can be  $k$ -coloured so that each edge contains one vertex of each colour. Given a  $k$ -graph  $S$ , recall the definitions

$$d_k(S) := \begin{cases} 0 & \text{if } e(S) = 0; \\ 1/k & \text{if } v(S) = k \text{ and } e(S) = 1; \\ \frac{e(S)-1}{v(S)-k} & \text{otherwise} \end{cases}$$

and

$$m_k(S) := \max_{S' \subseteq S} d_k(S').$$

**5.2. The maximum density of a hypergraph which is not Ramsey.** Given integers  $n \geq k$  and a  $k$ -graph  $H$ , we denote by  $\text{ex}(n; H)$  the maximum size of an  $n$ -vertex  $H$ -free  $k$ -graph. Define the *Turán density*  $\pi(H)$  of  $H$  by

$$(5.1) \quad \pi(H) := \lim_{n \rightarrow \infty} \frac{\text{ex}(n; H)}{\binom{n}{k}}$$

(which exists by a simple averaging argument, see [31]). The so-called supersaturation phenomenon discovered by Erdős and Simonovits [17] asserts that any sufficiently large hypergraph with density greater than  $\pi(H)$  contains not just one copy of  $H$ , but in fact a positive fraction of  $v(H)$ -sized sets span a copy of  $H$ .

**Theorem 5.1** ([17]). *For all  $k \in \mathbb{N}$ ;  $\delta > 0$  and all  $k$ -graphs  $H$ , there exist  $n_0, \varepsilon > 0$  such that for all integers  $n \geq n_0$ , every  $n$ -vertex  $k$ -graph  $G$  with  $e(G) \geq (\pi(H) + \delta) \binom{n}{k}$  contains at least  $\varepsilon \binom{n}{v(H)}$  copies of  $H$ .*

When  $k = 2$ , the Erdős–Stone–Simonovits theorem [18] says that for all graphs  $H$ , the value of  $\pi(H)$  is determined by the chromatic number  $\chi(H)$  of  $H$ , via

$$(5.2) \quad \pi(H) = 1 - \frac{1}{\chi(H) - 1}.$$

For  $k \geq 3$ , the value of  $\pi(H)$  is only known for a small family of  $k$ -graphs  $H$ . It remains an open problem to even determine the Turán density of  $K_4^{(3)}$ , the smallest non-trivial complete 3-graph (the widely-believed conjectured value is  $\frac{5}{9}$ ). For more background on this, the so-called *hypergraph Turán problem*, the interested reader should consult the excellent survey of Keevash [32].

In this section, we generalise Theorem 5.1 from  $H$ -free hypergraphs to non- $(H_1, \dots, H_r)$ -Ramsey hypergraphs (note that a hypergraph is  $H$ -free if and only if it is *not*  $(H)$ -Ramsey). Given  $\varepsilon > 0$ , we say that an  $n$ -vertex  $k$ -graph  $G$  is  $\varepsilon$ -*weakly*  $(H_1, \dots, H_r)$ -Ramsey if there exists an  $r$ -colouring  $\sigma$  of  $G$  such that, for all  $i \in [r]$ , the number of copies of  $H_i$  in  $\sigma^{-1}(i)$  is less than  $\varepsilon \binom{n}{v(H_i)}$ . Otherwise,  $G$  is  $\varepsilon$ -*strongly*  $(H_1, \dots, H_r)$ -Ramsey.

Using a well-known averaging argument of Katona, Nemetz and Simonovits [31], we can show that  $\binom{n}{k}^{-1} \text{ex}^r(n; H_1, \dots, H_r)$  converges as  $n$  tends to infinity. Indeed, let  $G$  be an  $n$ -vertex non- $(H_1, \dots, H_r)$ -Ramsey graph with  $e(G) = \text{ex}^r(n; H_1, \dots, H_r)$ . The average density of an  $(n-1)$ -vertex induced subgraph of  $G$  is precisely

$$\binom{n}{n-1}^{-1} \sum_{U \subseteq V(G); |U|=n-1} \frac{e(G[U])}{\binom{n-1}{k}} = (n-k)^{-1} \cdot \binom{n}{k}^{-1} \sum_{U \subseteq V(G); |U|=n-1} e(G[U]) = \binom{n}{k}^{-1} e(G).$$

But the left-hand side is at most  $\binom{n-1}{k}^{-1} \cdot \text{ex}^r(n-1; H_1, \dots, H_r)$ , otherwise  $G$  would contain an  $(n-1)$ -vertex subgraph which is  $(H_1, \dots, H_r)$ -Ramsey, violating the choice of  $G$ . We have shown that

$$\frac{\text{ex}^r(n; H_1, \dots, H_r)}{\binom{n}{k}}$$

is a non-increasing function of  $n$  (which is bounded below, by 0), and so this function has a limit. Therefore we may define the  $r$ -coloured *Turán density*  $\pi(H_1, \dots, H_r)$  of  $(H_1, \dots, H_r)$  by

$$\pi(H_1, \dots, H_r) := \lim_{n \rightarrow \infty} \frac{\text{ex}^r(n; H_1, \dots, H_r)}{\binom{n}{k}}.$$

As for  $k \geq 3$ , the problem of determining  $\pi(H)$  is still out of reach, we certainly cannot evaluate  $\pi(H_1, \dots, H_r)$  in general. However, any non- $(H_1, \dots, H_r)$ -Ramsey graph is  $K_s^{(k)}$ -free, where  $s :=$

$R(H_1, \dots, H_r)$  is the smallest integer  $m$  such that  $K_m^{(k)}$  is  $(H_1, \dots, H_r)$ -Ramsey. Thus

$$(5.3) \quad \pi(H_1, \dots, H_r) \leq \pi(K_s^{(k)}),$$

which is at most  $1 - \binom{s-1}{k-1}^{-1}$  (de Caen [9]). An interesting question is for which  $H_1, \dots, H_r$  the inequality in (5.3) is tight. We discuss the case  $k = 2$  in detail in Section 5.3.

We now state the main result of this subsection, which generalises Theorem 5.1 to  $r \geq 1$ .

**Theorem 5.2.** *For all  $\delta > 0$ , integers  $r \geq 1$  and  $k \geq 2$ , and  $k$ -graphs  $H_1, \dots, H_r$ , there exist  $n_0, \varepsilon > 0$  such that for all integers  $n \geq n_0$ , every  $n$ -vertex  $k$ -graph  $G$  with  $e(G) \geq (\pi(H_1, \dots, H_r) + \delta) \binom{n}{k}$  is  $\varepsilon$ -strongly  $(H_1, \dots, H_r)$ -Ramsey.*

*Proof.* Let  $\delta > 0$  and let  $r, k$  be positive integers with  $k \geq 2$ . By the definition of  $\pi(\cdot)$ , there exists  $m_0 > 0$  such that for all integers  $m \geq m_0$ ,

$$\text{ex}^r(m; H_1, \dots, H_r) < \left( \pi(H_1, \dots, H_r) + \frac{\delta}{2} \right) \binom{m}{k}.$$

Fix an integer  $m \geq m_0$ . Without loss of generality, we may assume that  $m \geq v(H_i)$  for all  $i \in [r]$ . Choose  $\varepsilon > 0$  to be such that

$$\varepsilon \leq \frac{\delta}{2r} \binom{m}{v(H_i)}^{-1}$$

for all  $i \in [r]$ . Let  $n$  be an integer which is sufficiently large compared to  $m$ , and let  $G$  be a  $k$ -graph on  $n$  vertices with  $e(G) = (\pi(H_1, \dots, H_r) + \delta) \binom{n}{k}$ . We need to show that, for every  $r$ -colouring  $\sigma$  of  $G$ , there is  $i \in [r]$  such that  $\sigma^{-1}(i)$  contains at least  $\varepsilon \binom{n}{v(H_i)}$  copies of  $H_i$ ; so fix an arbitrary  $\sigma$ .

Define  $\mathcal{M}$  to be the set of  $M \in \binom{V(G)}{m}$  such that  $e(G[M]) \geq (\pi(H_1, \dots, H_r) + \frac{\delta}{2}) \binom{m}{k}$ . Then

$$\sum_{U \subseteq V(G): |U|=m} e(G[U]) \leq |\mathcal{M}| \binom{m}{k} + \left( \binom{n}{m} - |\mathcal{M}| \right) \left( \pi(H_1, \dots, H_r) + \frac{\delta}{2} \right) \binom{m}{k}.$$

But for every  $e \in E(G)$ , there are exactly  $\binom{n-k}{m-k}$  sets  $U \subseteq V(G)$  with  $|U| = m$  such that  $e \in E(G[U])$ . Thus also

$$\sum_{U \subseteq V(G): |U|=m} e(G[U]) \geq \binom{n-k}{m-k} (\pi(H_1, \dots, H_r) + \delta) \binom{n}{k} = (\pi(H_1, \dots, H_r) + \delta) \binom{n}{m} \binom{m}{k},$$

and so, rearranging, we have  $|\mathcal{M}| \geq \delta \binom{n}{m} / 2$ . By the choice of  $m$ , for every  $M \in \mathcal{M}$ , there exists  $i = i(M) \in [r]$  such that  $\sigma^{-1}(i)$  contains a copy of  $H_i$  with vertices in  $M$ . Choose  $\mathcal{M}' \subseteq \mathcal{M}$  such that the  $i(M')$  are equal for all  $M' \in \mathcal{M}'$  and  $|\mathcal{M}'| \geq |\mathcal{M}|/r$ . Without loss of generality let us assume that  $i(M') = 1$  for all  $M' \in \mathcal{M}'$ . So for each  $M' \in \mathcal{M}'$ , there is a copy of  $H_1 \subseteq G[M']$  which is monochromatic with colour 1 under  $\sigma$ . Each such copy has vertex set contained in at most  $\binom{n-v(H_1)}{m-v(H_1)}$  sets  $M' \in \mathcal{M}'$ . Thus the number of such monochromatic copies of  $H_1$  in  $G$  is at least

$$\frac{\frac{\delta}{2} \cdot \binom{n}{m}}{r \binom{n-v(H_1)}{m-v(H_1)}} = \frac{\delta}{2r} \cdot \binom{m}{v(H_1)}^{-1} \cdot \binom{n}{v(H_1)} \geq \varepsilon \binom{n}{v(H_1)}.$$

So  $G$  is  $\varepsilon$ -strongly  $(H_1, \dots, H_r)$ -Ramsey, as required.  $\square$

**5.3. The special case of graphs: maximum size and typical structure.** The intimate connection between forbidden subgraphs and chromatic number when  $k = 2$  allows us to make some further remarks here. (This section is separate from the remainder of the paper and the results stated here will not be required later on.)

5.3.1. *The maximum number of edges in a graph which is not Ramsey.* Given  $s, n \in \mathbb{N}$ , let  $T_s(n)$  denote the  $s$ -partite Turán (2-)graph on  $n$  vertices; that is, the vertex set of  $T_s(n)$  has a partition into  $s$  parts  $V_1, \dots, V_s$  such that  $\|V_i\| - \|V_j\| \leq 1$  for all  $i, j \in [s]$ ; and  $xy$  is an edge of  $T_s(n)$  if and only if there are  $ij \in \binom{[s]}{2}$  such that  $x \in V_i$  and  $y \in V_j$ . Write  $t_s(n) := e(T_s(n))$ .

We need to define two notions of Ramsey number.

**Definition 5.3** (Ramsey number, chromatic Ramsey number and chromatic Ramsey equivalence). *Given an integer  $r \geq 1$  and families  $\mathcal{H}_1, \dots, \mathcal{H}_r$  of graphs, the Ramsey number  $R(\mathcal{H}_1, \dots, \mathcal{H}_r)$  is the least  $m$  such that  $K_m$  is  $(H_1, \dots, H_r)$ -Ramsey for some  $(H_1, \dots, H_r) \in \mathcal{H}_1 \times \dots \times \mathcal{H}_r$ . If  $\mathcal{H}_i = \{K_{\ell_i}\}$  for all  $i \in [r]$  then we instead write  $R(\ell_1, \dots, \ell_r)$ , and simply  $R^r(\ell)$  in the case when  $\ell_1 = \dots = \ell_r =: \ell$ .*

*Given graphs  $H_1, \dots, H_r$ , the chromatic Ramsey number  $R_\chi(H_1, \dots, H_r)$  is the least  $m$  for which there exists an  $(H_1, \dots, H_r)$ -Ramsey graph with chromatic number  $m$ .*

Trivially, for any  $k$ -graph  $H$ , we have that  $R_\chi(H) = \chi(H)$ . If  $H_1, \dots, H_r$  are graphs, then

$$(5.4) \quad t_{R_\chi(H_1, \dots, H_r)-1}(n) \leq \text{ex}^r(n; H_1, \dots, H_r) \leq t_{R_\chi(H_1, \dots, H_r)-1}(n) + o(n^2).$$

Thus

$$(5.5) \quad \pi(H_1, \dots, H_r) = 1 - \frac{1}{R_\chi(H_1, \dots, H_r) - 1} = \pi(K_{R_\chi(H_1, \dots, H_r)}).$$

The first inequality in (5.4) follows by definition of  $\text{ex}^r(n; H_1, \dots, H_r)$ ; the second from (5.2) applied with a graph  $H$  which is  $(H_1, \dots, H_r)$ -Ramsey and has  $\chi(H) = R_\chi(H_1, \dots, H_r)$ . Clearly, then,  $\pi(H_1, \dots, H_r) = \pi(J_1, \dots, J_r)$  if and only if  $R_\chi(J_1, \dots, J_r) = R_\chi(H_1, \dots, H_r)$ . So, in the graph case, the inequality (5.3) is tight when the Ramsey number and chromatic Ramsey number coincide.

As noted by Bialostocki, Caro and Roditty [4], one can determine  $\text{ex}^r(n; H_1, \dots, H_r)$  exactly in the case when  $H_1, \dots, H_r$  are cliques of equal size.

**Theorem 5.4** ([4]). *For all positive integers  $\ell, n \geq 3$  and  $r \geq 1$ , we have  $\text{ex}^r(n; K_\ell, \dots, K_\ell) = t_{R^r(\ell)-1}(n)$ .*

Thus in this case (5.3) is tight. The chromatic Ramsey number was introduced by Burr, Erdős and Lovász [8] who showed that, in principle, one can determine  $R_\chi$  given the usual Ramsey number  $R$ . A *graph homomorphism* from a graph  $H$  to a graph  $K$  is a function  $\phi: V(H) \rightarrow V(K)$  such that  $\phi(x)\phi(y) \in E(K)$  whenever  $xy \in E(H)$ . Let  $\text{Hom}(H)$  denote the set of all graphs  $K$  such that there exists a graph homomorphism  $\phi$  for which  $K = \phi(H)$ . Since there exists a homomorphism from  $H$  into  $K_\ell$  if and only if  $\chi(H) \leq \ell$ , we also have that  $R(\text{Hom}(H)) = \chi(H)$ . Thus  $R(\text{Hom}(H)) = R_\chi(H)$ . In fact this relationship extends to all  $r \geq 1$ .

**Lemma 5.5** ([8, 11, 40]). *For all integers  $r \in \mathbb{N}$  and graphs  $H_1, \dots, H_r$ ,*

$$R_\chi(H_1, \dots, H_r) = R(\text{Hom}(H_1), \dots, \text{Hom}(H_r)).$$

*Moreover, for all integers  $\ell_1, \dots, \ell_r \geq 3$ , we have that*

$$R_\chi(K_{\ell_1}, \dots, K_{\ell_r}) = R(\ell_1, \dots, \ell_r).$$

The second statement is a corollary of the first since  $\text{Hom}(K_\ell) = \{K_\ell\}$ . Another observation (see [8]) is that for all  $\ell \in \mathbb{N}$ , the chromatic Ramsey number  $R_\chi(C_{2\ell+1}, C_{2\ell+1})$  is equal to 5 if  $\ell = 2$ , and equal to 6 otherwise.

The first inequality in (5.4) is not always tight, for example when  $H$  is the disjoint union of two copies of some graph  $G$ . Indeed,  $\text{Hom}(H) \supseteq \text{Hom}(G)$  and so  $R_\chi(H, \dots, H) = R_\chi(G, \dots, G)$ . Let  $F$  be an  $n$ -vertex graph with  $e(F) = \text{ex}^r(n; G, \dots, G)$  which is not  $(G, r)$ -Ramsey. Obtain a graph  $T$  by adding an edge  $e$  to  $F$ . Then there exists an  $r$ -colouring of  $T$  in which every monochromatic

copy of  $G$  contains  $e$  (the monochromatic- $G$ -free colouring of  $F$ , with  $e$  arbitrarily coloured). Hence  $T$  is not  $(H, r)$ -Ramsey and so

$$\text{ex}^r(n; H, \dots, H) > \text{ex}^r(n; G, \dots, G) \geq t_{R_\chi(G, \dots, G)}(n) = t_{R_\chi(H, \dots, H)}(n).$$

We say that a graph  $H$  is (weakly) colour-critical if there exists  $e \in E(H)$  for which  $\chi(H - e) < \chi(H)$ . Complete graphs and odd cycles are examples of colour-critical graphs. The following conjecture would generalise Theorem 5.4 to provide a large class of graphs where the first inequality in (5.4) is tight.

**Conjecture 5.6.** *Let  $r$  be a positive integer and  $H$  a colour-critical graph. Then, whenever  $n$  is sufficiently large,*

$$\text{ex}^r(n; H, \dots, H) = t_{R_\chi(H, \dots, H)-1}(n).$$

If true, this conjecture would also generalise a well-known result of Simonovits [57] which extends Turán's theorem to colour-critical graphs. It would also determine  $\text{ex}^r(n; H, \dots, H)$  explicitly whenever  $H$  is an odd cycle.

**5.3.2. The typical structure of non-Ramsey graphs.** There has been much interest in determining the typical structure of an  $H$ -free graph. For example, Kolaitis, Prömel and Rothschild [37] proved that almost all  $K_r$ -free graphs are  $(r - 1)$ -partite. It turns out that one can easily obtain a result on the typical structure of non-Ramsey graphs from a result of Prömel and Steger [44].

Given two families  $\mathcal{A}(n), \mathcal{B}(n)$  of  $n$ -vertex graphs such that  $\mathcal{B}(n) \subseteq \mathcal{A}(n)$ , we say that almost all  $n$ -vertex graphs  $G \in \mathcal{A}(n)$  are in  $\mathcal{B}(n)$  if

$$\lim_{n \rightarrow \infty} \frac{|\mathcal{A}(n)|}{|\mathcal{B}(n)|} = 1.$$

The next result of Prömel and Steger [44] immediately tells us the typical structure of non-Ramsey graphs in certain cases.

**Theorem 5.7** ([44]). *For every graph  $H$ , the following holds. Almost all  $H$ -free graphs are  $(\chi(H) - 1)$ -partite if and only if  $H$  is colour-critical.*

**Corollary 5.8.** *For all integers  $r$  and graphs  $H_1, \dots, H_r$ , if there exists an  $(H_1, \dots, H_r)$ -Ramsey graph  $H$  such that  $\chi(H) = R_\chi(H_1, \dots, H_r)$  and  $H$  is colour-critical, then almost every non- $(H_1, \dots, H_r)$ -Ramsey graph is  $(R_\chi(H_1, \dots, H_r) - 1)$ -partite.*

*Proof.* The result follows since every non- $(H_1, \dots, H_r)$ -Ramsey graph  $G$  is  $H$ -free, and every  $(R_\chi(H_1, \dots, H_r) - 1)$ -partite graph is non- $(H_1, \dots, H_r)$ -Ramsey.  $\square$

In particular, if in Corollary 5.8, each  $H_i$  is a clique, say  $H_i = K_{\ell_i}$ , then by Lemma 5.5 we can take  $H := K_{R(\ell_1, \dots, \ell_r)}$ . So, for example, almost every non- $(K_3, 2)$ -Ramsey graph is 5-partite.

**5.4. A container theorem for Ramsey hypergraphs.** Recall that  $\overline{\text{Ram}}(n; H_1, \dots, H_r)$  is the set of  $n$ -vertex  $k$ -graphs which are not  $(H_1, \dots, H_r)$ -Ramsey and  $\text{Ram}(H_1, \dots, H_r)$  is the set of  $(H_1, \dots, H_r)$ -Ramsey  $k$ -graphs (on any number of vertices). Recall further that an  $H$ -free  $k$ -graph is precisely an  $(H, 1)$ -Ramsey graph. Write  $\mathcal{G}^k(n)$  for the set of all  $k$ -graphs on vertex set  $[n]$ . Let  $\mathcal{I}_r(n; H_1, \dots, H_r)$  denote the set of all ordered  $r$ -tuples  $(G_1, \dots, G_r) \in (\mathcal{G}^k(n))^r$  of  $k$ -graphs such that each  $G_i$  is  $H_i$ -free and  $E(G_i) \cap E(G_j) = \emptyset$  for all distinct  $i, j \in [r]$ . Note that for any  $G \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$ , there exist pairwise edge-disjoint  $k$ -graphs  $G_1, \dots, G_r$  such that  $\bigcup_{i \in [r]} G_i = G$  and  $(G_1, \dots, G_r) \in \mathcal{I}_r(n; H_1, \dots, H_r)$ . In this subsection, we prove a container theorem for elements in  $\mathcal{I}_r(n; H_1, \dots, H_r)$ . To do so, we will apply Theorem 3.3 to hypergraphs  $\mathcal{H}_1, \dots, \mathcal{H}_r$ , where  $\mathcal{H}_i$  is the hypergraph of copies of  $H_i$  (see Definition 5.9). In  $\mathcal{H}_i$ , an independent set corresponds to an  $H_i$ -free  $k$ -graph.



**Definition 5.9.** Given an integer  $k \geq 2$ , a  $k$ -graph  $H$  and positive integer  $n$ , the hypergraph  $\mathcal{H}$  of copies of  $H$  in  $K_n^{(k)}$  has vertex set  $V(\mathcal{H}) := \binom{[n]}{k}$ , and  $E \subseteq \binom{V(\mathcal{H})}{e(H)}$  is an edge of  $\mathcal{H}$  if and only if  $E$  is isomorphic to  $E(H)$ .

We will need the following simple proposition from [3].

**Proposition 5.10** ([3], Proposition 7.3). *Let  $H$  be a  $k$ -graph. Then there exists  $c > 0$  such that, for all positive integers  $n$ , the following holds. Let  $\mathcal{H}$  be the  $e(H)$ -uniform hypergraph of copies of  $H$  in  $K_n^{(k)}$ . Then, letting  $p = n^{-1/m_k(H)}$ ,*

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

for every  $\ell \in [e(H)]$ .

We can now prove our container theorem for elements in  $\mathcal{I}_r(n; H_1, \dots, H_r)$ .

**Theorem 5.11.** *Let  $r, k \in \mathbb{N}$  with  $k \geq 2$  and  $\delta > 0$ . Let  $H_1, \dots, H_r$  be  $k$ -graphs such that  $m_k(H_1) \geq \dots \geq m_k(H_r)$  and  $\Delta_1(H_i) \geq 2$  for all  $i \in [r]$ . Then there exists  $D > 0$  such that the following holds. For all  $n \in \mathbb{N}$ , there is a collection  $\mathcal{S}_r \subseteq (\mathcal{G}_k(n))^r$  and a function  $f : \mathcal{S}_r \rightarrow (\mathcal{G}_k(n))^r$  such that:*

- (i) For all  $(I_1, \dots, I_r) \in \mathcal{I}_r(n; H_1, \dots, H_r)$ , there exists  $S \in \mathcal{S}_r$  such that  $S \subseteq (I_1, \dots, I_r) \subseteq f(S)$ .
- (ii) If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum_{i \in [r]} e(S_i) \leq Dn^{k-1/m_k(H_1)}$ .
- (iii) Every  $S \in \mathcal{S}_r$  satisfies  $S \in \mathcal{I}_r(n; H_1, \dots, H_r)$ .
- (iv) Given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$ , write  $f(S) =: (f(S_1), \dots, f(S_r))$ . Then
  - (a)  $\bigcup_{i \in [r]} f(S_i)$  is  $\delta$ -weakly  $(H_1, \dots, H_r)$ -Ramsey; and
  - (b)  $e\left(\bigcup_{i \in [r]} f(S_i)\right) \leq \text{ex}^r(n; H_1, \dots, H_r) + \delta \binom{n}{k}$ .

Note that if  $H$  is a  $k$ -graph with  $\Delta_1(H) = 1$ , then  $H$  is a matching, i.e. a set of vertex-disjoint edges.

*Proof.* We will identify any hypergraph which has vertex set  $[n]$  with its edge set. It suffices to prove the theorem when  $n$  is sufficiently large; otherwise we can set  $\mathcal{S}_r$  to be  $\mathcal{I}_r(n; H_1, \dots, H_r)$ ; set  $f$  to be the identity function and choose  $D$  to be large. We may further assume that there are no isolated vertices in  $H_i$  for any  $i \in [r]$ .

Apply Proposition 5.10 with input hypergraphs  $H_1, \dots, H_r$  to obtain  $c > 0$  such that its conclusion holds with  $H_i$  playing the role of  $H$ , for all  $i \in [r]$ . Let  $\delta > 0$ ,  $r \in \mathbb{N}$  and  $k \geq 2$  be given and apply Theorem 5.2 (with  $\delta/2$  playing the role of  $\delta$ ) to obtain  $n_0, \varepsilon > 0$ . Without loss of generality we may assume  $\varepsilon \leq \delta < 1$ . For each  $i \in [r]$ , let  $v_i := v(H_i)$  and  $m_i := e(H_i)$  for all  $i \in [r]$ . Set  $v := \max_{i \in [r]} v_i$ ;  $m := \max_{i \in [r]} m_i$ ;

$$\varepsilon' := \frac{\varepsilon}{2 \cdot v!}; \quad \text{and} \quad \varepsilon'' := \frac{\varepsilon'}{\binom{v}{k} \cdot v!}.$$

Apply Theorem 3.3 with parameters  $r, m_1, \dots, m_r, c, \varepsilon''$  playing the roles of  $r, k_1, \dots, k_r, c, \varepsilon$  respectively to obtain  $D > 0$ . Increase  $n_0$  if necessary so that  $0 < 1/n_0 \ll 1/D, 1/k, 1/r, \varepsilon, \delta$  and let  $n \geq n_0$  be an integer.

Let  $\mathcal{H}_{n,i}$  be the hypergraph of copies of  $H_i$  in  $K_n^{(k)}$ . That is,  $V(\mathcal{H}_{n,i}) := \binom{[n]}{k}$  and for each  $m_i$ -subset  $E$  of  $\binom{[n]}{k}$ , put  $E \in E(\mathcal{H}_{n,i})$  if and only if  $E$  is isomorphic to a copy of  $H_i$ . By definition,  $\mathcal{H}_{n,i}$  is an  $m_i$ -uniform hypergraph and an independent set in  $\mathcal{H}_{n,i}$  corresponds to an  $H_i$ -free  $k$ -graph

with vertex set  $[n]$ . Since  $H_i$  is a  $k$ -graph with no isolated vertices,

$$(5.6) \quad e(\mathcal{H}_{n,i}) = \frac{v_i!}{|\text{Aut}(H_i)|} \binom{n}{v_i}$$

where  $\text{Aut}(H_i)$  is the automorphism group of  $H_i$ . For all  $i \in [r]$ , let

$$\mathcal{F}_{n,i} := \left\{ A \subseteq \binom{[n]}{k} : e(\mathcal{H}_{n,i}[A]) \geq \varepsilon' e(\mathcal{H}_{n,i}) \right\}.$$

We claim that  $\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r}$  and  $\mathcal{F}_{n,1}, \dots, \mathcal{F}_{n,r}$  satisfy the hypotheses of Theorem 3.3 with the parameters chosen as above and with

$$p = p(n) := n^{-1/m_k(H_1)}.$$

Clearly each family  $\mathcal{F}_{n,i}$  is increasing, and  $\mathcal{H}_{n,i}$  is  $(\mathcal{F}_{n,i}, \varepsilon')$ -dense. Next, we show that  $|A| \geq \varepsilon'' \binom{n}{k}$  for all  $A \in \mathcal{F}_{n,i}$ . In any  $k$ -graph on  $n$  vertices, there are at most  $v_i! \binom{n-k}{v_i-k}$  copies of  $H_i$  that contain some fixed set  $\{x_1, \dots, x_k\}$  of vertices. Therefore, for every  $e \in \binom{[n]}{k}$ , the number of  $E \in E(\mathcal{H}_{n,i})$  containing  $e$  is at most

$$(5.7) \quad v_i! \binom{n-k}{v_i-k}.$$

Thus every  $A \in \mathcal{F}_{n,i}$  satisfies

$$|A| \geq \frac{e(\mathcal{H}_{n,i}[A])}{v_i! \binom{n-k}{v_i-k}} \stackrel{(5.6)}{\geq} \frac{\varepsilon' v_i! \binom{n}{v_i}}{v_i! \binom{n-k}{v_i-k} |\text{Aut}(H_i)|} = \frac{\varepsilon'}{\binom{v}{k} |\text{Aut}(H_i)|} \binom{n}{k} \geq \varepsilon'' \binom{n}{k},$$

where, in the final inequality, we used the fact that  $|\text{Aut}(H_i)| \leq v_i!$ . Note that  $\varepsilon'' < \varepsilon'$ . So  $\mathcal{H}_{n,i}$  is  $(\mathcal{F}_{n,i}, \varepsilon'')$ -dense and  $|A| \geq \varepsilon'' \binom{n}{k}$  for all  $A \in \mathcal{F}_{n,i}$ .

Certainly  $p \geq n^{-1/m_k(H_j)}$  for all  $j \in [r]$ . By the choice of  $c$ , we then have

$$\Delta_\ell(\mathcal{H}_{n,i}) \leq c \cdot p^{\ell-1} \frac{e(\mathcal{H}_{n,i})}{\binom{n}{k}}$$

for all  $i \in [r]$  and  $\ell \in [m_i]$ . We have shown that  $\mathcal{H}_{n,i}$  and  $\mathcal{F}_{n,i}$  satisfy the hypotheses of Theorem 3.3 for all  $i \in [r]$ .

Then Theorem 3.3 implies that there exists a family  $\mathcal{S}_r \subseteq \prod_{i \in [r]} \mathcal{P}(V(\mathcal{H}_{n,i})) = \mathcal{P}(\binom{[n]}{k})^r$  and functions  $f' : \mathcal{S}_r \rightarrow \prod_{i \in [r]} \overline{\mathcal{F}_{n,i}}$  and  $g : \mathcal{I}(\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r}) \rightarrow \mathcal{S}_r$  such that the following conditions hold:

- (a) If  $(S_1, \dots, S_r) \in \mathcal{S}_r$  then  $\sum |S_i| \leq Dp \binom{n}{k}$ ;
- (b) every  $S \in \mathcal{S}_r$  satisfies  $S \in \mathcal{I}(\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r})$ ;
- (c) for every  $(I_1, \dots, I_r) \in \mathcal{I}(\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r})$ , we have that  $S \subseteq (I_1, \dots, I_r) \subseteq S \cup f'(S)$ , where  $S := g(I_1, \dots, I_r)$ .

Note that  $(G_1, \dots, G_r) \in \mathcal{I}(\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r})$  if and only if  $(G_1, \dots, G_r) \in \mathcal{I}_r(n; H_1, \dots, H_r)$  (where we recall the identification of graphs and edge sets). For each  $S \in \mathcal{S}_r$ , define

$$f(S) := S \cup f'(S).$$

So  $f : \mathcal{S}_r \rightarrow \mathcal{P}(\binom{[n]}{k})^r$ . (Note that under the correspondence of graphs and edge sets we can view  $\mathcal{P}(\binom{[n]}{k})^r = (\mathcal{G}_k(n))^r$ .) Thus (a)–(c) immediately imply that (i) and (iii) hold, and additionally for any  $(S_1, \dots, S_r) \in \mathcal{S}_r$  we have

$$\sum_{i \in [r]} e(S_i) \leq Dp \binom{n}{k} \leq Dn^{-1/m_k(H_1)} \cdot \frac{n^k}{k!} < Dn^{k-1/m_k(H_1)},$$

yielding (ii).

Given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  write  $f(S) =: (f(S_1), \dots, f(S_r))$  and  $f'(S) =: (f'(S_1), \dots, f'(S_r))$ . Let  $G := \bigcup_{i \in [r]} f(S_i)$ ; so  $G$  is a  $k$ -graph with vertex set  $[n]$ . To prove (iv)(a), we need to exhibit an  $r$ -colouring  $\sigma$  of  $G$  with the property that  $\sigma^{-1}(i)$  contains less than  $\varepsilon \binom{n}{v_i}$  copies of  $H_i$  for all  $i \in [r]$ . Indeed, consider the  $r$ -colouring  $\sigma$  of  $G$  defined by setting  $\sigma(e) = i$  when  $i$  is the least integer such that  $e \in f(S_i)$ . Then the subgraph of  $G$  coloured  $i$  is  $\sigma^{-1}(i) \subseteq f(S_i) = S_i \cup f'(S_i)$ . Since  $S_i$  is an independent set in  $\mathcal{H}_{n,i}$ , we have that  $S_i$  is  $H_i$ -free. Every copy of  $H_i$  in  $\sigma^{-1}(i)$  either contains at least one edge in  $S_i$ , or has every edge contained in  $f'(S_i)$ . Note that  $m_k(H_1) \leq m$ . By (5.7), the number of copies of  $H_i$  in  $G$  containing at least one edge in  $S_i$  is at most

$$e(S_i) \cdot v_i! \binom{n-k}{v_i-k} \leq Dn^{k-1/m_k(H_1)} \cdot v_i!(n-k)^{v_i-k} \leq Dv_i! \cdot n^{v_i-\frac{1}{m}} < \frac{\varepsilon}{2} \binom{n}{v_i}.$$

For each  $i \in [r]$  we have that  $f'(S_i) \in \overline{\mathcal{F}_{n,i}}$ , and so  $e(\mathcal{H}_{n,i}[f'(S_i)]) < \varepsilon' e(\mathcal{H}_{n,i})$ . That is, the number of copies of  $H_i$  in  $f'(S_i)$  is less than

$$\varepsilon' \cdot \frac{v_i!}{|\text{Aut}(H_i)|} \binom{n}{v_i} \leq \frac{\varepsilon}{2} \binom{n}{v_i}.$$

Thus, in total  $f(S_i) = S_i \cup f'(S_i)$  contains at most  $\varepsilon \binom{n}{v_i}$  copies of  $H_i$ , so  $G$  is  $\varepsilon$ -weakly  $(H_1, \dots, H_r)$ -Ramsey. Since  $\varepsilon \leq \delta$ , this immediately implies (iv)(a), and (iv)(b) follows from Theorem 5.2, our choice of parameters, and since  $n$  is sufficiently large.  $\square$

As in Theorem 3.1, we will call the elements  $S \in \mathcal{S}_r$  *fingerprints*, and each  $\bigcup_{i \in [r]} f(S_i)$  with  $(S_1, \dots, S_r) \in \mathcal{S}_r$  is a *container*.

**5.5. The number of hypergraphs which are not Ramsey.** Our first application of Theorem 5.11 is an enumeration result for non- $(H_1, \dots, H_r)$ -Ramsey hypergraphs (Theorem 1.12), which asymptotically determines the logarithm of  $|\overline{\text{Ram}}(n; H_1, \dots, H_r)|$ .

*Proof of Theorem 1.12.* Let  $0 < \delta < 1$  be arbitrary, and let  $n \in \mathbb{N}$  be sufficiently large. Clearly,  $|\overline{\text{Ram}}(n; H_1, \dots, H_r)| \geq 2^{\text{ex}^r(n; H_1, \dots, H_r)}$  since no subhypergraph of an  $n$ -vertex non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -graph with  $\text{ex}^r(n; H_1, \dots, H_r)$  edges is  $(H_1, \dots, H_r)$ -Ramsey.

For the upper bound, suppose first that  $\Delta_1(H_i) \geq 2$  for all  $i \in [r]$ . Let  $D > 0$  be obtained from Theorem 5.11 applied to  $H_1, \dots, H_r$  with parameter  $\delta$ . We obtain a collection  $\mathcal{S}_r$  and a function  $f$  as in Theorem 5.11. Consider any  $G \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$ . Note that there are pairwise edge-disjoint  $k$ -graphs  $G_1, \dots, G_r$  such that  $\bigcup_{i \in [r]} G_i = G$  and  $(G_1, \dots, G_r) \in \mathcal{I}_r(n; H_1, \dots, H_r)$ . So by Theorem 5.11(i) this means there is some  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  so that  $G \subseteq \bigcup_{i \in [r]} f(S_i)$ . Further, given any  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$ , we have

$$e\left(\bigcup_{i \in [r]} f(S_i)\right) \leq \text{ex}^r(n; H_1, \dots, H_r) + \delta \binom{n}{k}.$$

Thus, each such  $\bigcup_{i \in [r]} f(S_i)$  contains at most  $2^{\text{ex}^r(n; H_1, \dots, H_r) + \delta \binom{n}{k}}$   $k$ -graphs in  $\overline{\text{Ram}}(n; H_1, \dots, H_r)$ . Note that, by Theorem 5.11(ii),

$$|\mathcal{S}_r| \leq \left( \sum_{s=0}^{Dn^{k-1/m_k(H_1)}} \binom{\binom{n}{k}}{s} \right)^r < 2^{\delta \binom{n}{k}},$$

where the last inequality holds since  $n$  is sufficiently large. Altogether, this implies

$$(5.8) \quad |\overline{\text{Ram}}(n; H_1, \dots, H_r)| \leq 2^{\delta n} \times 2^{\text{ex}^r(n; H_1, \dots, H_r) + \delta \binom{n}{k}} = 2^{\text{ex}^r(n; H_1, \dots, H_r) + 2\delta \binom{n}{k}}.$$

Since the choice of  $0 < \delta < 1$  was arbitrary, this proves the theorem in the case when  $\Delta_1(H_i) \geq 2$  for all  $i \in [r]$ .

Suppose now that, say,  $\Delta_1(H_1) = 1$ . Then  $H_1$  is a matching. Certainly every non- $(H_2, \dots, H_r)$ -Ramsey  $k$ -graph is non- $(H_1, \dots, H_r)$ -Ramsey. Let  $H \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$ . Then there exists an  $r$ -colouring  $\sigma$  of  $H$  such that  $\sigma^{-1}(i)$  is  $H_i$ -free for all  $i \in [r]$ . Thus  $H$  is the union of pairwise edge-disjoint  $k$ -graphs  $J \in \overline{\text{Ram}}(n; H_2, \dots, H_r)$  and  $J' := \sigma^{-1}(1)$ . But  $J'$  is  $H_1$ -free and hence does not contain a matching of size  $\lfloor v(H_1)/2 \rfloor =: h$ . A result of Erdős [15] (used here in a weaker form) implies that, for sufficiently large  $n$ ,

$$e(J') \leq (h-1) \binom{n-1}{k-1}.$$

Thus, for large  $n$ ,

$$\begin{aligned} |\overline{\text{Ram}}(n; H_1, \dots, H_r)| &\leq \sum_{J \in \overline{\text{Ram}}(n; H_2, \dots, H_r)} \binom{\binom{n}{k}}{(h-1) \binom{n-1}{k-1}} = |\overline{\text{Ram}}(n; H_2, \dots, H_r)| \cdot \binom{\binom{n}{k}}{\frac{k(h-1)}{n} \binom{n}{k}} \\ &\leq |\overline{\text{Ram}}(n; H_2, \dots, H_r)| \cdot 2^{\delta \binom{n}{k}}. \end{aligned}$$

Iterating this argument, using (5.8) and the fact that  $0 < \delta < 1$  was arbitrary, we obtain the required upper bound in the general case.  $\square$

In fact Theorem 1.12 can be recovered in a different way, which, to the best of our knowledge, has not been explicitly stated elsewhere. Let  $\mathcal{F}$  be a (possibly infinite) family of  $k$ -graphs, and let  $\text{Forb}(n; \mathcal{F})$  be the set of  $n$ -vertex  $k$ -graphs which contain no copy of any  $F \in \mathcal{F}$  as a subhypergraph. The following result of Nagle, Rödl and Schacht [42] asymptotically determines the logarithm of  $|\text{Forb}(n; \mathcal{F})|$ . (This generalises the corresponding result of Erdős, Frankl and Rödl [16] for graphs.) Let

$$\text{ex}(n; \mathcal{F}) := \max\{e(H) : H \in \text{Forb}(n; \mathcal{F})\}.$$

(So when  $\mathcal{F} = \{F\}$  contains a single  $k$ -graph, we have  $\text{ex}(n; \{F\}) = \text{ex}(n; F)$ .)

**Theorem 5.12** (Theorem 2.3, [42]). *Let  $k \geq 2$  be a positive integer and  $\mathcal{F}$  be a (possibly infinite) family of  $k$ -graphs. Then, for all  $n \in \mathbb{N}$ ,*

$$|\text{Forb}(n; \mathcal{F})| = 2^{\text{ex}(n; \mathcal{F}) + o(n^k)}.$$

Since  $G \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$  if and only if  $G$  is an  $n$ -vertex  $k$ -graph without a copy of any  $F \in \text{Ram}(H_1, \dots, H_r)$  as a subhypergraph, Theorem 5.12 immediately implies Theorem 1.12. We remark also that (subcases of) Theorem 1.12 can be obtained from other container results (e.g. [19]).

**5.6. The resilience of being  $(H_1, \dots, H_r)$ -Ramsey.** Recall that  $G_{n,p}^{(k)}$  has vertex set  $[n]$ , where each edge lies in  $\binom{[n]}{k}$  and appears with probability  $p$ , independently of all other edges. In this section we apply Theorem 5.11 to prove Theorem 1.7, which determines  $\text{res}(G_{n,p}^{(k)}, (H_1, \dots, H_r)$ -Ramsey) for given fixed  $k$ -graphs  $H_1, \dots, H_r$ . Explicitly,  $\text{res}(G_{n,p}^{(k)}, (H_1, \dots, H_r)$ -Ramsey) is the minimum integer  $t$  such that one can remove  $t$  edges from  $G_{n,p}^{(k)}$  to obtain a  $k$ -graph  $H$  which has an  $(H_1, \dots, H_r)$ -free  $r$ -colouring.

Observe that Theorem 1.7 together with (5.3) immediately implies the following corollary.

**Corollary 5.13** (Random Ramsey for hypergraphs). *For all positive integers  $r, k$  with  $k \geq 2$  and  $k$ -graphs  $H_1, \dots, H_r$  with  $m_k(H_1) \geq \dots \geq m_k(H_r)$  and  $\Delta_1(H_i) \geq 2$  for all  $i \in [r]$ , there exists  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ G_{n,p}^{(k)} \text{ is } (H_1, \dots, H_r)\text{-Ramsey} \right] = 1 \quad \text{if } p > C n^{-1/m_k(H_1)}.$$

In the case when  $m_k(H_1) = m_k(H_2)$ , Corollary 5.13 generalises Theorem 1.6 since we do not require  $H_1$  to be strictly  $k$ -balanced. Further, Corollary 5.13 resolves (the 1-statement part) of Conjecture 1.3 in the case when  $m_2(H_1) = m_2(H_2)$ .

*Proof of Theorem 1.7.* Let  $0 < \delta < 1$  be arbitrary,  $r, k \in \mathbb{N}$  with  $k \geq 2$ , and let  $H_1, \dots, H_r$  be  $k$ -graphs as in the statement of the theorem. Given  $n \in \mathbb{N}$ , if  $p > n^{-1/m_k(H_1)}$ , then  $p > n^{-(k-1)}$  since  $\Delta_1(H_1) \geq 2$ . Proposition 2.1 implies that, w.h.p.,

$$(5.9) \quad e(G_{n,p}^{(k)}) = \left(1 \pm \frac{\delta}{4}\right) p \binom{n}{k}.$$

For brevity, write  $\pi := \pi(H_1, \dots, H_r)$ . We will first prove the upper bound

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \text{res}(G_{n,p}^{(k)}, (H_1, \dots, H_r)\text{-Ramsey}) \leq (1 - \pi + \delta) e(G_{n,p}^{(k)}) \right] = 1 \quad \text{if } p > n^{-1/m_k(H_1)}.$$

For this, we must show that the probability of the event that there exists an  $n$ -vertex  $k$ -graph  $G \subseteq G_{n,p}^{(k)}$  such that  $e(G) \geq (\pi - \delta) e(G_{n,p}^{(k)})$  and  $G \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$ , tends to one as  $n$  tends to infinity. This indeed follows: Let  $n$  be sufficiently large so that  $\text{ex}^r(n; H_1, \dots, H_r) \geq (\pi - \delta/2) \binom{n}{k}$ . Let  $G^*$  be an  $n$ -vertex non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -graph with  $e(G^*) = \text{ex}^r(n; H_1, \dots, H_r)$ . Then, by Proposition 2.1, w.h.p. we have  $e(G^* \cap G_{n,p}^{(k)}) = (\pi \pm \delta) e(G_{n,p}^{(k)})$ , and  $G^* \cap G_{n,p}^{(k)} \in \overline{\text{Ram}}(n; H_1, \dots, H_r)$ , as required.

For the remainder of the proof, we will focus on the lower bound, namely that there exists  $C > 0$  such that whenever  $p > Cn^{-1/m_k(H_1)}$ ,

$$(5.10) \quad \mathbb{P} \left[ \text{res}(G_{n,p}^{(k)}, (H_1, \dots, H_r)\text{-Ramsey}) \geq (1 - \pi - \delta) e(G_{n,p}^{(k)}) \right] \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Suppose  $n$  is sufficiently large. Apply Theorem 5.11 with parameters  $r, k, \delta/16, (H_1, \dots, H_r)$  to obtain  $D > 0$  and for each  $n \in \mathbb{N}$ , a collection  $\mathcal{S}_r$  and a function  $f$  satisfying (i)–(iv). Now choose  $C$  such that  $0 < 1/C \ll 1/D, \delta, 1/k, 1/r$ . Let  $p \geq Cn^{-1/m_k(H_1)}$ .

Since (5.9) holds with high probability, to prove (5.10) holds it suffices to show that the probability  $G_{n,p}^{(k)}$  contains a non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -graph with at least  $(\pi + \delta/2)p \binom{n}{k}$  edges tends to zero as  $n$  tends to infinity.

Suppose that  $G_{n,p}^{(k)}$  does contain a non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -graph  $I$  with at least  $(\pi + \delta/2)p \binom{n}{k}$  edges. Then there exist pairwise edge-disjoint  $k$ -graphs  $I_1, \dots, I_r$  such that  $\bigcup_{i \in [r]} I_i = I$  and  $(I_1, \dots, I_r) \in \mathcal{I}_r(n; H_1, \dots, H_r)$ . Further, there is some  $S = (S_1, \dots, S_r) \in \mathcal{S}_r$  such that  $S \subseteq (I_1, \dots, I_r) \subseteq f(S)$ . Thus,  $G_{n,p}^{(k)}$  must contain (the edges of)  $\bigcup_{i \in [r]} S_i$  as well as at least  $(\pi + \delta/4)p \binom{n}{k}$  edges from  $(\bigcup_{i \in [r]} f(S_i)) \setminus (\bigcup_{i \in [r]} S_i)$ . (Note here we are using that  $e(\bigcup_{i \in [r]} S_i) \leq \delta p \binom{n}{k} / 4$ , which holds by Theorem 5.11(ii) and since  $0 < 1/C \ll 1/D, 1/k, \delta$ .) Writing  $s := e(\bigcup_{i \in [r]} S_i)$ , the probability  $G_{n,p}^{(k)}$  contains  $\bigcup_{i \in [r]} S_i$  is  $p^s$ . Note that  $e((\bigcup_{i \in [r]} f(S_i)) \setminus (\bigcup_{i \in [r]} S_i)) \leq (\pi + \delta/8) \binom{n}{k}$  by Theorem 5.11(iv)(b) and since  $n$  is sufficiently large. So by the first part of Proposition 2.1, the probability  $G_{n,p}^{(k)}$  contains at least  $(\pi + \delta/4)p \binom{n}{k}$  edges from  $(\bigcup_{i \in [r]} f(S_i)) \setminus (\bigcup_{i \in [r]} S_i)$  is at most  $\exp(-\delta^2 p \binom{n}{k} / 256) \leq \exp(-\delta^2 p n^k / 256 k^k)$ .

Write  $N := n^{k-1/m_k(H_1)}$  and  $\gamma := \delta^2 / 256 k^k$ . Given some integer  $0 \leq s \leq DN$ , there are at most  $r^s \binom{n}{s}$  elements  $(S_1, \dots, S_r) \in \mathcal{S}_r$  such that  $e(\sum_{i \in [r]} S_i) = s$ . Indeed, this follows by Theorem 5.11(iii) and since there are  $r^s$  ways to partition a set of size  $s$  into  $r$  classes. Thus, the probability that  $G_{n,p}^{(k)}$  does contain a non- $(H_1, \dots, H_r)$ -Ramsey  $k$ -graph  $I$  with at least  $(\pi + \delta/2)p \binom{n}{k}$  edges is at most

$$\sum_{s=0}^{DN} r^s \binom{n}{s} \cdot p^s \cdot e^{-\gamma n^k p} \leq \sum_{s=0}^{DN} \left( \frac{r e^{k+1} p n^k}{s k^k} \right)^s e^{-\gamma n^k p} \leq (DN + 1) \left( \frac{r e^{k+1} p n^k}{DN k^k} \right)^{DN} e^{-\gamma n^k p}$$

$$\leq (DN + 1) \left( \frac{r e^{k+1} C}{D k^k} \right)^{DN} e^{-\gamma CN} \leq e^{\gamma CN/2} e^{-\gamma CN} = e^{-\gamma CN/2},$$

which tends to zero as  $n$  tends to infinity. This completes the proof.  $\square$

**Remark.** In the proof it was vital that there is a one-to-one correspondence between containers and fingerprints. Given a container  $C$ , let  $S(C) =: (S_1(C), \dots, S_r(C))$  be the fingerprint of  $C$ . We know that the number of containers whose fingerprint has  $e(\bigcup_{i \in [r]} S_i(C)) = s$  is at most  $r^s \binom{n}{s}$ . So if there are a lot of containers, many of them have large fingerprints. But if there is  $H \subseteq G_{n,p}^{(k)}$  such that  $C$  is a container for  $H$  and  $e(\bigcup_{i \in [r]} S_i(C)) = s$ , then  $\bigcup_{i \in [r]} S_i(C) \subseteq H$  and so  $\bigcup_{i \in [r]} S_i(C) \subseteq G_{n,p}^{(k)}$ , which occurs with probability  $p^s$ . So this small term  $p^s$  compensates for the large term  $r^s \binom{n}{s}$ . Note also that this correspondence between containers and fingerprints plays the same crucial role in the proof of Theorem 1.11.

#### ACKNOWLEDGEMENTS

The authors are grateful to Victor Falgas-Ravry for a helpful conversation on [19], and to Hong Liu for a helpful conversation on the graph Ramsey problems considered in this paper.

#### REFERENCES

- [1] H.L. Abbott and E.T.H. Wang, Sum-free sets of integers, *Proc. Amer. Math. Soc.* **67** (1977), 11–16.
- [2] J. Balogh, B. Csaba and W. Samotij, Local resilience of almost spanning trees in random graphs, *Random Structures & Algorithms* **38** (2011), 121–139.
- [3] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* **28** (2015), 669–709.
- [4] A. Bialostocki, Y. Caro, and Y. Roditty, On zero sum Turan numbers, *Ars Combin.* **29** (1990), 117–127.
- [5] T.F. Bloom, A quantitative improvement for Roth’s theorem on arithmetic progressions, *J. London Math. Soc.*, **93** (3), (2016), 643–663.
- [6] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* **7** (1) (1987), 35–38.
- [7] J. Böttcher, Y. Kohayakawa and A. Taraz, Almost spanning subgraphs of random graphs after adversarial edge removal, *Combin. Probab. Comput.* **22** (2013), 639–683.
- [8] S. A. Burr, P. Erdős, L. Lovász, On graphs of Ramsey type, *Ars Combinatoria* **1** (1976) no. 1, 167–190.
- [9] D. de Caen, Extension of a theorem of Moon and Moser on complete subgraphs, *Ars Combin.* **16** (1983), 5–10.
- [10] P. Cameron and P. Erdős, On the number of sets of integers with various properties, in *Number Theory* (R.A. Mollin, ed.), 61–79, Walter de Gruyter, Berlin, 1990.
- [11] V. Chvátal, personal communication to the authors of [8].
- [12] D. Conlon, Combinatorial theorems relative to a random set, *Proceedings of the International Congress of Mathematicians 2014*, Vol. 4, 303–328.
- [13] D. Conlon and W.T. Gowers, Combinatorial theorems in sparse random sets, *Ann. Math.* **84** (2016), 367–454.
- [14] M. Elkin, An improved construction of progression-free sets, *Israel J. Math.* **184** (2011), 93–128.
- [15] P. Erdős, A problem on independent  $r$ -tuples, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **8** (1965), 93–95.
- [16] 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) no. 2, 113–121.
- [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, *Bulletin American Math. Soc.* **52** (12) (1946), 1087–1091.
- [19] V. Falgas-Ravry, K. O’Connell and A. Uzzell, Multicolour containers, extremal entropy and counting, submitted.
- [20] P. Frankl, R.L. Graham and V. Rödl, Quantitative theorems for regular systems of equations, *J. Combin. Theory A* **47** (1988), 246–261.
- [21] E. Friedgut, V. Rödl and M. Schacht, Ramsey properties of random discrete structures, *Random Structures & Algorithms* **37** (2010), 407–436.
- [22] R. Graham, V. Rödl, and A. Ruciński, On Schur properties of random subsets of integers, *J. Number Theory* **61** (1996), 388–408.
- [23] B. Green, The Cameron-Erdős conjecture, *Bull. London Math. Soc.* **36** (2004), 769–778.

- [24] B. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Anal.* **15** (2005), 340–376.
- [25] B. Green and J. Wolf, A note on Elkins improvement of Behrend’s construction, *Additive number theory: Festschrift in honor of the sixtieth birthday of Melwyn B. Nathanson*, pages 141–144. Springer-Verlag, 1st edition, 2010.
- [26] L. Gugelmann, R. Nenadov, Y. Person, A. Steger, N. Škorić and H. Thomas, Symmetric and asymmetric Ramsey properties in random hypergraphs, preprint.
- [27] P.E. Haxell, Y. Kohayakawa and T. Łuczak, Turán’s extremal problem in random graphs: forbidding even cycles, *J. Combin. Theory B* **64** (1995), 273–287.
- [28] P.E. Haxell, Y. Kohayakawa and T. Łuczak, Turán’s extremal problem in random graphs: forbidding odd cycles, *Combinatorica* **16** (1996), 107–122.
- [29] M.C. Hu, A note on sum-free sets of integers, *Proc. Amer. Math. Soc.* **80** (1980), 711–712.
- [30] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, 2000.
- [31] G. Katona, T. Nemetz and M. Simonovits, On a problem of Turán in the theory of graphs, *Mat. Lapok* **15** (1964), 228–238.
- [32] P. Keevash, Hypergraph Turán problems, *Surveys in Combinatorics*, Cambridge University Press, 2011, 83–140.
- [33] D. Kleitman and K.J. Winston, The asymptotic number of lattices, *Ann. Discrete Math.* **6** (1980), 243–249, Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978).
- [34] D. Kleitman and K.J. Winston, On the number of graphs without 4-cycles, *Discrete Math.* **6** (1982), 167–172.
- [35] Y. Kohayakawa and B. Kreuter, Threshold functions for asymmetric Ramsey properties involving cycles, *Random Structures & Algorithms* **11** (1997), 245–276.
- [36] Y. Kohayakawa, T. Łuczak and V. Rödl, On  $K_4$ -free subgraphs of random graphs, *Combinatorica* **17** (1997), 173–213.
- [37] P. Kolaitis, H.J. Promel and B. Rothschild,  $K_{l+1}$ -free graphs: asymptotic structure and a 0 – 1 law, *Trans. Amer. Math. Soc.* **303** (1987), 637–671.
- [38] D. Král’, O. Serra and L. Vena, A removal lemma for systems of linear equations over finite fields, *Israel J. Math.* **187** (2012), 193–207.
- [39] C. Lee and B. Sudakov, Dirac’s theorem for random graphs, *Random Structures & Algorithms* **41** (2012), 293–305.
- [40] S. Lin, On Ramsey numbers and  $K_r$ -coloring of graphs, *J. Combin. Theory* **12** (1972), 82–92.
- [41] M. Marcinişzyn, J. Skokan, R. Spöhel and A. Steger, Asymmetric Ramsey properties of random graphs involving cliques, *Random Structures & Algorithms* **34** (2009), 419–453.
- [42] B. Nagle, V. Rödl, M. Schacht, Extremal hypergraph problems and the regularity method, *Topics in Discrete Mathematics*, vol. 26 (2006) *Algorithms Combin.*, 247–278.
- [43] R. Nenadov and A. Steger, A Short Proof of the Random Ramsey Theorem, *Combin. Probab. Comput.* **25** (2016), 130–144.
- [44] H. Prömel and A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, *Combinatorica* **12** (4) (1992), 463–473.
- [45] R. Rado, Studien zur kombinatorik, *Mathematische Zeitschrift* **36** (1933), 424–470.
- [46] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, in *Combinatorics, Paul Erdős is Eighty*, Vol. 1, 317–346, Bolyai Soc. Math. Studies, János Bolyai Math. Soc., Budapest, 1993.
- [47] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures & Algorithms* **5** (1994), 253–270.
- [48] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.
- [49] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers, *Proc. London Math. Soc.* **74** (3) (1997), 481–502.
- [50] V. Rödl and A. Ruciński, Ramsey properties of random hypergraphs, *J. Combin. Theory Ser. A* **81** (1) (1998), 1–33.
- [51] V. Rödl, A. Ruciński, and M. Schacht, Ramsey properties of random  $k$ -partite,  $k$ -uniform hypergraphs, *SIAM J. Discrete Math.* **21** (2), (2007) 442–460.
- [52] I.Z. Ruzsa, Solving a linear equation in a set of integers I, *Acta Arith.* **65** (1993), 259–282.
- [53] A.A. Sapozhenko, The Cameron-Erdős conjecture, (Russian) *Dokl. Akad. Nauk.* **393** (2003), 749–752.
- [54] D. Saxton and A. Thomason, Hypergraph containers, *Invent. Math.* **201** (2015), 925–992.
- [55] M. Schacht, Extremal results for random discrete structures, *Ann. Math.* **184** (2016), 331–363.
- [56] I. Schur, Über die Kongruenz  $x^m + y^m \equiv z^m \pmod{p}$ , *ber. Deutsch. Mat. Verein.* **25** (1916), 114–117.
- [57] M. Simonovits, ‘A method for solving extremal problems in graph theory, stability problems’, in *1968 Theory of Graphs* (Proc. Colloq., Tihany, 1966) 279–319, Academic Press, New York.

- [58] C. Spiegel, A note on sparse supersaturation and extremal results for linear homogeneous systems, preprint.
- [59] B. Sudakov, Robustness of graph properties, submitted.
- [60] B. Sudakov and V.H. Vu, Local resilience of graphs, *Random Structures & Algorithms* **33** (2008), 409–433.
- [61] B.L. van der Warden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wisk.* **15** (1927), 212–216.