Independent sets in hypergraphs and Ramsey properties of graphs and the integers

Hancock, Robert; Staden, Katherine; Treglown, Andrew

DOI:
10.1137/18M119642X

SIAM J. DISCRETE MATH.
Vol. 33, No. 1, pp. 153–188

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 [*J. Amer. Math. Soc.*, 28 (2015), pp. 669–709], and independently Saxton and Thomason [*Invent. Math.*, 201 (2015), pp. 925–992], 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 give relatively short and elementary proofs of a number of results concerning Ramsey (and Turán) properties of (hyper)graphs and the integers. In particular we do the following: (a) We generalize the *random Ramsey theorem* of Rödl and Ruciński [*Combinatorics, Paul Erdős Is Eighty*, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Mathematical Society, Budapest, 1993, pp. 317–346; *Random Structures Algorithms*, 5 (1994), pp. 253–270; *J. Amer. Math. Soc.*, 8 (1995), pp. 917–942] by providing a resilience analogue. Our result unifies and generalizes several fundamental results in the area including the *random version of Turán's theorem* due to Conlon and Gowers [*Ann. of Math.*, 184 (2016), pp. 367–454] and Schacht [*Ann. of Math.*, 184 (2016), pp. 331–363]. (b) The above result also resolves a general subcase of the *asymmetric random Ramsey conjecture* of Kohayakawa and Kreuter [*Random Structures Algorithms*, 11 (1997), pp. 245–276]. (c) All of the above results in fact hold for uniform hypergraphs. (d) 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-colored, there is a monochromatic copy of H in G). (e) We strengthen the *random Rado theorem* of Friedgut, Rödl, and Schacht [*Random Structures Algorithms*, 37 (2010), pp. 407–436] by proving a resilience version of the result. (f) For partition regular matrices A we determine, up to an error term in the exponent, the number of subsets of $\{1, \ldots, n\}$ for which there exists an r-coloring which contains no monochromatic solutions to $Ax = 0$. Along the way a number of open problems are posed.

Key words. Ramsey theory, container method, random graphs, random sets of integers

AMS subject classifications. 05C30, 05C55, 05D10, 11B75

DOI. 10.1137/18M119642X

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 colorings of graphs, and H-free graphs.

*Received by the editors June 25, 2018; accepted for publication (in revised form) November 27, 2018; published electronically January 15, 2019.
http://www.siam.org/journals/sidma/33-1/M119642.html
Funding: The second author's research supported by ERC grant 306493. The third author's research supported by EPSRC grant EP/M016641/1.
†University of Birmingham, Birmingham, United Kingdom, and Institute of Mathematics, Czech Academy of Sciences, Praha, 110 00, Czechia (hancock@math.cas.cz).
‡Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom (staden@maths.ox.ac.uk).
§University of Birmingham, Birmingham, United Kingdom (a.c.treglown@bham.ac.uk).

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 [38, 39] 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 [62]. 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 [14] and Schacht [64]. In particular, these container and transference theorems can be used to prove a range of combinatorial results in a random setting. See [4] for a survey on the container method.

An overarching aim of the paper is to demonstrate that with the container method at hand, one can give relatively short and elementary proofs of fundamental results concerning Ramsey properties of graphs and the integers. Moreover, our results give us a precise understanding about how *resiliently* typical graphs and sets of integers of a given density possess a given Ramsey property. In particular, one of our main results is a resilience random Ramsey theorem (Theorem 1.7). This result provides a unified framework for studying both the Ramsey and Turán problems in the setting of random (hyper)graphs. In particular, Theorem 1.7 implies the (so-called 1-statements of the) random Ramsey theorem due to Rödl and Ruciński [54, 55, 56] and the random version of Turán's theorem [14, 64]. Moreover, Theorem 1.7 also resolves a general subcase of the asymmetric random Ramsey conjecture of Kohayakawa and Kreuter [40]. Since Theorem 1.7 unifies and generalizes several fundamental results concerning Ramsey and Turán properties of random (hyper)graphs, we survey these topics in sections 1.1.2–1.1.4 before we state this result in section 1.1.5.

We also prove a sister result to Theorem 1.7, a resilience strengthening of the random Rado theorem (Theorem 1.11). Again the container method allows us to give a rather short proof of this result. We further provide results on the enumeration of Ramsey graphs (Theorem 1.12) and sets of integers without a given Ramsey property (Theorem 1.13).

The results we prove all correspond to problems concerning tuples of disjoint independent sets in hypergraphs. In particular, from the container theorem of Balogh, Morris, and Samotij one can easily obtain an analogous result for tuples of independent sets in hypergraphs (see Proposition 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-colors 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 while 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 our 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 $\mathcal{P}$ while *local resilience* considers how many edges one can delete at each vertex while ensuring the resulting graph has property $\mathcal{P}$. More precisely, we define the *global resilience of G with respect to $\mathcal{P}$*, $\mathrm{res}(G,\mathcal{P})$, to be the minimum number t such that by deleting t edges from G, one can obtain a graph 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 [69], 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., [69, 45], almost spanning trees [2], and embedding subgraphs of small bandwidth [8]. See [69] and the surveys [13, 68] 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 $G^{(k)}_{n,p}$ 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 $n \to \infty$. 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 [7] 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,ε)-*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,ε)-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 ε term), (H,ε)-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,ε)-Turán, then $\mathrm{res}(G,\mathcal{P}) = (\frac{1}{\chi(H)-1} \pm \varepsilon)e(G)$, and vice versa, where $\mathcal{P}$ is the property of containing H as a subgraph. (Note that we write $x = a \pm b$ to say that the value of x is some real number in the interval $[a-b, a+b]$.) 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 [13]), and determining when $G_{n,p}$ is (H,ε)-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,ε)-Turán. In [30, 31, 41] it was conjectured that w.h.p. $G_{n,p}$ is (H,ε)-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 [64] and (in the case when H is strictly 2-balanced, i.e., $m_2(H') < m_2(H)$ for all $H' \subset H$) by Conlon and Gowers [14].

THEOREM 1.1 (see [64, 14]). *For any graph H with $\Delta(H) \geq 2$ and any $\varepsilon > 0$, there are positive constants c, C such that*

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

Given an integer r, an *r-coloring* of a graph G is a function $\sigma : E(G) \to [r]$. (So this is not necessarily a proper coloring.) We say that G is *(H,r)-Ramsey* if every r-coloring 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

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

The following result of Rödl and Ruciński [54, 55, 56] yields a random version of Ramsey's theorem.

THEOREM 1.2 (see [54, 55, 56]). *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*

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

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 color these special edges blue and color 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 [51] 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 $H_1, \ldots, H_r$, a graph G is *$(H_1, \ldots, H_r)$-Ramsey* if for any r-coloring of G there is a copy of H_i in color i for some $i \in [r]$. (This definition coincides with that of (H,r)-Ramsey

when $H_1 = \cdots = H_r = H$.) Kohayakawa and Kreuter [40] 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

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

CONJECTURE 1.3 (see [40]). *For any graphs $H_1, \ldots, H_r$ with $m_2(H_1) \geq \cdots \geq m_2(H_r) > 1$, there are positive constants $c, C > 0$ such that*

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

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 [29]. 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 generalize Theorem 1.2. Kohayakawa and Kreuter [40] have confirmed Conjecture 1.3 when the H_i are cycles. In [47] it was observed that the approach used by Kohayakawa and Kreuter [40] implies the 1-statement of Conjecture 1.3 holds when H_1 is strictly 2-balanced *provided* the so-called KŁR conjecture holds. This latter conjecture was proved by Balogh, Morris, and Samotij [3], thereby proving that the 1-statement of Conjecture 1.3 holds in this case.

Additional note. Since the paper was submitted the 1-statement of Conjecture 1.3 has been proved by Mousset, Nenadov, and Samotij [49].

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 and 1.2 and Conjecture 1.3 by determining the threshold for being (H, ε)-Turán, (H, r)-Ramsey, and more generally being $(H_1, \ldots, H_r)$-Ramsey. The definitions of (H, r)-Ramsey and $(H_1, \ldots, H_r)$-Ramsey extend from graphs in the obvious way. Given a k-uniform hypergraph H, let $\mathrm{ex}(n; H)$ be the maximum size of an n-vertex H-free hypergraph. A simple averaging argument shows that the limit

$$\pi(H) := \lim_{n\to\infty} \frac{\mathrm{ex}(n; H)}{\binom{n}{k}}$$

exists. Now we say that a k-uniform hypergraph G is (H, ε)-*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 generalizes the definition we gave earlier.) We also need to generalize 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 and Gowers [14] and of Schacht [64] actually extended to a proof of a version of Theorem 1.1 for hypergraphs:

THEOREM 1.4 (see [14, 64]). *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*

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

The 1-statement of Theorem 1.2 was generalized to hypergraphs by Friedgut, Rödl, and Schacht [24] and by Conlon and Gowers [14], proving a conjecture of Rödl and Ruciński [58]. (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 [58] and [59], respectively. Also in [51] Nenadov and Steger remark that their proof of the 1-statement of Theorem 1.2 extends to Theorem 1.5.)

THEOREM 1.5 (see [14, 24]). *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*

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

In [29], 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. (This H is the disjoint union of a tight cycle and hypergraph triangle.)

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

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

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 et al. [29] generalized the 1-statement of Conjecture 1.3 to k-uniform hypergraphs, in the case when $H_1' = H_1$ is the unique maximizer in (1.3), i.e., H_1 is *strictly k-balanced with respect to* $m_k(\cdot, H_2)$.

THEOREM 1.6 (see [29]). *For all positive integers r, k with $k \geq 2$ and k-uniform hypergraphs $H_1, \ldots, H_r$ with $m_k(H_1) \geq \cdots \geq m_k(H_r)$, where H_1 is strictly k-balanced with respect to $m_k(\cdot, H_2)$, there exists $C > 0$ such that*

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

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 generalizes, 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 $\mathrm{ex}^r(n; H_1, \dots, H_r)$ be the maximum size of an n-vertex k-uniform hypergraph G which is not $(H_1, \dots, H_r)$-Ramsey. Define the *r-colored Turán density*

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

Observe that $\mathrm{ex}^1(n; H) = \mathrm{ex}(n; H)$ since a hypergraph is H-free if and only if it is not $(H, 1)$-Ramsey. Note further that $\pi(\cdot, \dots, \cdot)$ generalizes $\pi(\cdot)$. So when $k = 2$, we have $\pi(H) = 1 - \frac{1}{\chi(H)-1}$. We will observe 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 $m_k(H_1) \geq \cdots \geq m_k(H_r)$. There exists $C > 0$ such that*

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

Thus, when $p > Cn^{-1/m_k(H_1)}$, the random hypergraph $G^{(k)}_{n,p}$ 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 $G^{(k)}_{n,p}$ whose edge density is slightly smaller than this which *does not* have the Ramsey property.

Note that the threshold of $p > Cn^{-1/m_k(H_1)}$ in Theorem 1.7 is tight up to the multiplicative constant C. Indeed, consider the random hypergraph $G^{(k)}_{n,p}$ with $p \ll n^{-1/m_k(H_1)}$. Let $H_1' \subseteq H_1$ be such that $m_k(H_1) = d_k(H_1')$. Then the expected number of copies of H_1' in $G^{(k)}_{n,p}$ is much smaller than the expected number of edges in $G^{(k)}_{n,p}$, so w.h.p. we can delete every copy of H_1' (and therefore H_1) by removing $o(e(G^{(k)}_{n,p}))$ edges. So the hypergraph G that remains has $(1 - o(1))e(G^{(k)}_{n,p})$ edges and is not $(H_1, \dots, H_r)$-Ramsey because we can color every edge of G with color 1. Then, since G is H_1-free, there is no copy of H_i in color i in G.

Let us describe the importance of Theorem 1.7 (in the case $k = 2$ and $H_1 = \cdots = H_r = H$) 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 summarized 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 $H_1 = \cdots = H_r = H$, we obtain a stronger statement in place of the 1-statement of Theorem 1.2 as described above.
- Theorem 1.7 proves the 1-statement of 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 $m_k(\cdot, H_2)$.

Note that even though Theorem 1.7 implies many of the known results concerning Ramsey properties of random (hyper)graphs, often the resilience random Ramsey problem is different from the random Ramsey problem. In particular, we have determined the threshold for the former problem, while we have seen above examples of (hyper)graphs $H_1, \ldots, H_r$ where a lower value of p still ensures that $G^{(k)}_{n,p}$ is w.h.p. $(H_1, \ldots, H_r)$-Ramsey.

1.1.6. Resilience in the integers. An important branch of Ramsey theory concerns partition properties of sets of integers. Schur's classical theorem [65] states that if $\mathbb{N}$ is r-colored there exists a monochromatic solution to $x+y = z$; later van der Waerden [72] showed that the same hypothesis ensures a monochromatic arithmetic progression of arbitrary length. More generally, Rado's theorem [53] characterizes all those systems of homogeneous linear equations $\mathcal{L}$ for which every finite coloring 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 ℓ. 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 $x = (x_1, \ldots, x_k) \in S^k$ 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-coloring 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 coloring of $\mathbb{N}$, there is always a monochromatic solution to $\mathcal{L}$. As mentioned above, Rado's theorem characterizes 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 [57, 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.,

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

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 $\mathrm{rank}(A_{\overline{W}})$ be the rank of $A_{\overline{W}}$, where $\mathrm{rank}(A_{\overline{W}}) = 0$ for $\overline{W} = \emptyset$. We set

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

We remark that the denominator of $m(A)$ is strictly positive provided that A is irredundant and partition regular.

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 [57] 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 (see [57]). *For all irredundant partition regular full rank matrices A and all positive integers $r \geq 2$, there exists a constant $c > 0$ such that*

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

We remark that 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 [24] 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 (see [24]). *For all irredundant partition regular full rank matrices A and all positive integers r, there exists a constant $C > 0$ such that*

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

Earlier, Theorem 1.9 was confirmed by Graham, Rödl, and Ruciński [25] in the case where $\mathcal{L}$ is $x + y = z$ and $r = 2$, and then by Rödl and Ruciński [57] 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 $\mathcal{P}$*, $\mathrm{res}(S, \mathcal{P})$, 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 $\mathrm{res}([n], \mathcal{P}) \geq (1 - \varepsilon)n$.

The following result of Schacht [64] provides a resilience strengthening of Theorem 1.9 in the case of density regular matrices.

THEOREM 1.10 (see [64]). *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*

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

Note that in [64] 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*

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

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 [23, Theorem 1].) Thus, Theorem 1.11 does imply Theorem 1.9. Further, in the case when A is density regular, [23, 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 [64]. 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 and also deals with the asymmetric case, namely, Theorem 4.1.

Additional note. Just before submitting the paper we were made aware of simultaneous and independent work of Spiegel [67]. In [67] the case $r = 1$ of Theorem 4.1 is proved. 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 [18] 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. There has also been interest in strengthening this result, e.g., in the case when H is bipartite; see, e.g., [22, 48]. Given any $k, r, n \in \mathbb{N}$ with $k \geq 2$ and k-uniform hypergraphs $H_1, \ldots, H_r$, define $\mathrm{Ram}(n; H_1, \ldots, H_r)$ to be the collection of all k-uniform hypergraphs on vertex set $[n]$ that are $(H_1, \ldots, H_r)$-Ramsey and $\overline{\mathrm{Ram}}(n; H_1, \ldots, H_r)$ to be all those k-uniform hypergraphs on $[n]$ that are not $(H_1, \ldots, H_r)$-Ramsey. A natural question is to determine the size of $\mathrm{Ram}(n; H_1, \ldots, H_r)$. 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 $H_1, \ldots, H_r$ be k-uniform hypergraphs. Then*

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

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 [18]. In fact, one can also obtain Theorem 1.12 by using the work from [50], a hypergraph analogue of the result in [18]; see section 5.5 for a proof of this. Similar results were obtained also using containers by Falgas-Ravry, O'Connell, and Uzzell in [21], and in [70] by Terry, who reproved a result of Ishigami [33].

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 [32] 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. Related results (in the 1-color case) were obtained by Green [27] and Saxton and Thomason [63].

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

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 use the convention that the set of natural numbers $\mathbb{N}$ does not include zero.

We will make use of the following Chernoff inequality (see, e.g., [34, 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 (see [3], Theorem 2.2). *For every $k \in \mathbb{N}$ and all positive c and ε, 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 $f : \mathcal{S} \to \overline{\mathcal{F}}$ and $g : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ 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 $(I_1,\dots,I_r) \cup (J_1,\dots,J_r) := (I_1 \cup J_1,\dots,I_r \cup J_r)$.

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 $1 \leq i \leq r$. So, for example, $\mathcal{P}([n])^r$ denotes the collection of all r-tuples $(X_1,\dots,X_r)$ so that $X_i \subseteq [n]$ 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 proposition is an analogous result for the r-tuples of disjoint independent sets of a hypergraph. It is a straightforward consequence of Theorem 3.1.

PROPOSITION 3.2. *For every $k, r \in \mathbb{N}$ and all positive c and ε, 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 $\mathcal{S}_r \subseteq \mathcal{I}_r(\mathcal{H})$ and functions $f : \mathcal{S}_r \to (\overline{\mathcal{F}})^r$ and $g : \mathcal{I}_r(\mathcal{H}) \to \mathcal{S}_r$ 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) *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, ε 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 Proposition 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 $f' : \mathcal{S} \to \overline{\mathcal{F}}$ and $g' : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ 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

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

Let $(S_1, \ldots, S_r) \in \mathcal{S}_r$. 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 $\mathcal{S}_r \subseteq \mathcal{I}_r(\mathcal{H})$.

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

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 $(f'(S_1), \ldots, f'(S_r)) \in (\overline{\mathcal{F}})^r$ and $(g'(I_1), \ldots, g'(I_r)) \in \mathcal{S}_r$.

Now for (ii), 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 $g(I_1, \ldots, I_r) = (g'(I_1), \ldots, g'(I_r)) \subseteq (I_1, \ldots, I_r)$. Since $f(g(I_1, \ldots, I_r)) = (f'(g'(I_1)), \ldots, f'(g'(I_r)))$ we also have $(I_1, \ldots, I_r) \subseteq f(g(I_1, \ldots, I_r)) \cup g(I_1, \ldots, I_r)$ as required. ∎

In all of our applications of the container method, we will in fact apply the following asymmetric version of Proposition 3.2. In particular, 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 $\mathcal{I}(\mathcal{H}_1, \ldots, \mathcal{H}_r)$ to be the set of all r-tuples $(I_1, \ldots, I_r) \in \prod_{i \in [r]} \mathcal{I}(\mathcal{H}_i)$ such that $I_i \cap I_j = \emptyset$ for all $1 \leq i < j \leq r$.

We omit the proof of Proposition 3.3 since it follows from Theorem 3.1 as in the proof of Proposition 3.2.

PROPOSITION 3.3. *For every $r, k_1, \ldots, k_r \in \mathbb{N}$ 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 $\mathcal{S}_r \subseteq \mathcal{I}(\mathcal{H}_1, \ldots, \mathcal{H}_r)$ and functions $f : \mathcal{S}_r \to \prod_{i \in [r]} \overline{\mathcal{F}_i}$ and $g : \mathcal{I}(\mathcal{H}_1, \ldots, \mathcal{H}_r) \to \mathcal{S}_r$ such that the following conditions hold:

(i) *if $(S_1, \ldots, S_r) \in \mathcal{S}_r$, then $\sum |S_i| \leq Cp|V|$;*

(ii) *for every $(I_1, \ldots, I_r) \in \mathcal{I}(\mathcal{H}_1, \ldots, \mathcal{H}_r)$, we have that $S \subseteq (I_1, \ldots, I_r) \subseteq S \cup f(S)$, where $S := g(I_1, \ldots, I_r)$.*

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. 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 Proposition 3.3. 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 nonzero rational entries.

Call an integer matrix A (and the corresponding system of linear equations $\mathcal{L}$) *r-regular* if all r-colorings 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. As outlined in the next subsection, given any $r \geq 2$, all irredundant r-regular matrices A satisfy $(*)$. We will in fact prove stronger versions of Theorems 1.11 and 1.13 that consider irredundant matrices with property $(*)$.

These general results also consider "asymmetric" Rado properties: Suppose that $\mathcal{L}_i$ is a system of linear equations for each $1 \leq i \leq r$ (and, here and elsewhere, A_i is the matrix such that $\mathcal{L}_i = \mathcal{L}(A_i)$). We say a set $X \subseteq \mathbb{N}$ is *$(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free* if there is an r-coloring of X such that there are no solutions to $\mathcal{L}_i$ in X in color i for every $i \in [r]$. Otherwise we say that X is *$(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-Rado*. We denote the size of the largest $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset of $[n]$ by $\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$.

In general it is not known which systems of linear equations $\mathcal{L}_1, \ldots, \mathcal{L}_r$ are such that $\mathbb{N}$ is $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-Rado. However, if each $\mathcal{L}_i$ is an r-regular homogenous linear equation, then $\mathbb{N}$ is $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-Rado (see [44, Theorem 9.19]).

We will prove the following strengthenings of Theorems 1.11 and 1.13.

THEOREM 4.1. *For all positive integers r and all irredundant full rank matrices $A_1, \ldots, A_r$ which satisfy $(*)$ with $m(A_1) \geq \cdots \geq m(A_r)$, and all $\delta > 0$, there exists a constant $C > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\mathrm{res}([n]_p, (\mathcal{L}_1, \ldots, \mathcal{L}_r)\text{-}Rado)}{|[n]_p|} = 1 - \frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} \pm \delta\right] = 1 \quad \text{if } p > Cn^{-1/m(A_1)}.$$

THEOREM 4.2. *For all positive integers r and all irredundant full rank matrices $A_1, \ldots, A_r$ which satisfy $(*)$, there are $2^{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r) + o(n)}$ $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-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 [27, 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 [67] independently proved the case $r = 1$ of Theorem 4.1. (Note in [67] this result is mentioned in terms of *abundant* matrices A. That is, every $\ell \times (k-2)$ submatrix of A has rank ℓ. 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 there exists a pair $ij \in \binom{[k]}{2}$ and nonzero rationals α, β such that for all solutions $(x_1, \dots, x_k)$ to $\mathcal{L}$ we have $\alpha x_i = \beta x_j$. 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 $\alpha > \beta > 0$, and devise the following 2-coloring of $\mathbb{N}$: greedily color the numbers $\{1, 2, 3, \dots\}$ so that when coloring x, we always give it a different color than $\beta x/\alpha$ (if $\beta x/\alpha \in \mathbb{N}$). Such a coloring 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, $[k] = D_1 \cup \cdots \cup D_t$, such that

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

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

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

Rado's theorem [53] states that a matrix is partition regular if and only if it satisfies the columns property. Now, for example, $A := \begin{pmatrix} 2 & 2 & -1 \end{pmatrix}$ is irredundant and clearly satisfies (∗), and additionally does not have the columns property, so is not partition regular.

The argument above actually implies that if an irredundant matrix A is 2-regular, then it satisfies (∗). So in the symmetric case, 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 proved for irredundant matrices which satisfy (∗), since in [57] 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 [57]). 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 ℓ which satisfies (∗). Then for every $W \subseteq [k]$, the following hold.*

(i) *If $|W| = 1$, then* $\operatorname{rank}(A_{\overline{W}}) = \ell$.

(ii) *If $|W| \geq 2$, then* $\ell - \operatorname{rank}(A_{\overline{W}}) + 2 \leq |W|$.

(iii) *If $|W| \geq 2$, then*

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

Furthermore,

(iv) $k \geq \ell + 2$;

(v) $m(A) > 1$.

Proof. For (i), suppose that $\operatorname{rank}(A_{\overline{W}}) = \ell - 1$ 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 nonzero entry. If there is a nonzero entry in this row, then there are no positive solutions to $\mathcal{L}$, which contradicts A being irredundant. If there are none, then A does not have rank ℓ, 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 nonzero entries. If there are two nonzero entries this contradicts A satisfying $(*)$. Otherwise we again get a contradiction to either A being irredundant or of rank ℓ. 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, and 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), by taking $W = [k]$ the result follows immediately from (ii). For (v), again take $W = [k]$. Then by definition (1.5), $m(A) \geq (k-1)/(k-\ell-1) > 1$, where the second inequality follows since the denominator is positive by (iv). ∎

The following supersaturation lemma follows easily from the (1-color) removal lemma proved for integer matrices by Král', Serra, and Vena (Theorem 2 in [43]).

LEMMA 4.4. *Fix $r \in \mathbb{N}$ and for each $i \in [r]$, let A_i be an $\ell_i \times k_i$ integer matrix of rank ℓ_i, and write $\mathcal{L}_i := \mathcal{L}(A_i)$. For every $\delta > 0$ there exist $n_0, \varepsilon > 0$ with the following property. Suppose $n \geq n_0$ is an integer and $X \subseteq [n]$ is r-colored, and $|X| \geq \mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r) + \delta n$. Then there exists an $i \in [r]$ such that there are more than $\varepsilon n^{k_i - \ell_i}$ k_i-distinct solutions to $\mathcal{L}_i$ in color i 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 ℓ-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. Use Gaussian elimination to turn A into echelon form. Now note that when picking a solution to $Ax = b$ in X (where $x = (x_1, \ldots, x_k)$), there are $|X|$ choices for $k - \text{rank}(A)$ of the x_i (the "free" variables), and the other $\text{rank}(A)$ of the x_i are immediately determined. Thus there are at most $|X|^{k-\text{rank}(A)}$ solutions as required. ∎

COROLLARY 4.6. *Consider an $\ell \times k$ matrix A of rank ℓ, a set $X \subseteq [n]$, and an integer $1 \leq t \leq k$. Fix distinct $y_1, \ldots, y_t \in X$ and consider any $W = \{s_1, \ldots, s_t\} \subseteq [k]$. The system $Ax = 0$ has at most $|X|^{k-t-\text{rank}(A_{\overline{W}})}$ solutions $(x_1, \ldots, x_k)$ 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 $(x_1, \ldots, x_k)$ 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 rth 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 $|X|^{k-t-\mathrm{rank}(A_{\overline{W}})}$ 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 $\mathrm{rank}(A_{\overline{W}}) = \ell$ and so the result follows. ∎

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 $\mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$ denote the set of all ordered r-tuples $(X_1, \dots, X_r) \in \mathcal{P}([n])^r$ so that each X_i is $\mathcal{L}_i$-free and $X_i \cap X_j = \emptyset$ for all distinct $i, j \in [r]$. Note that any $(\mathcal{L}_1, \dots, \mathcal{L}_r)$-free subset X of $[n]$ has a partition $X_1, \dots, X_r$ so that $(X_1, \dots, X_r) \in \mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$. We now prove a container theorem for the elements of $\mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$.

THEOREM 4.7. *Let $r \in \mathbb{N}$ and $0 < \delta < 1$. For each $i \in [r]$ let A_i be an $\ell_i \times k_i$ irredundant matrix of full rank ℓ_i which satisfies $(*)$, and suppose that $m(A_1) \geq \cdots \geq m(A_r)$. Then there exists $D > 0$ such that the following holds. For all $n \in \mathbb{N}$, there is a collection $\mathcal{S}_r \subseteq \mathcal{P}([n])^r$ and a function $f : \mathcal{S}_r \to \mathcal{P}([n])^r$ such that the following holds.*

(i) *For all $(I_1, \dots, I_r) \in \mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$, 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 $\sum_{i \in [r]} |S_i| \leq Dn^{\frac{m(A_1)-1}{m(A_1)}}$.*

(iii) *Every $S \in \mathcal{S}_r$ satisfies $S \in \mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$.*

(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) *for each $1 \leq i \leq r$, $f(S_i)$ contains at most $\delta n^{k_i - \ell_i}$ k_i-distinct solutions to $\mathcal{L}_i$; and*

  (b) *$|\cup_{i \in [r]} f(S_i)| \leq \mu(n, \mathcal{L}_1, \dots, \mathcal{L}_r) + \delta n$.*

We emphasize that (iv)(b) does not necessarily guarantee $\sum_{i \in [r]} |f(S_i)| \leq \mu(n, \mathcal{L}_1, \dots, \mathcal{L}_r) + \delta n$. Rather it ensures at most $\mu(n, \mathcal{L}_1, \dots, \mathcal{L}_r) + \delta n$ elements of $[n]$ appear in at least one of the co-ordinates of $f(S)$. This property is crucial for our applications.

Proof. First note that since each of the matrices A_i are irredundant, a result of Janson and Ruciński [35] implies that there exists a constant $d > 0$ such that, for each $i \in [r]$, there are at least $dn^{k_i - \ell_i}$ k_i-distinct solutions to $\mathcal{L}_i$ in $[n]$.

Note that it suffices to prove the theorem in the case when $0 < \delta < d$. Also, it suffices to prove the theorem when n is sufficiently large; otherwise we can set $\mathcal{S}_r$ to be $\mathcal{I}(n, \mathcal{L}_1, \dots, \mathcal{L}_r)$; set f to be the identity function and choose D to be large.

Let $0 < \delta < d$ 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 $\varepsilon \leq \delta$. Define $k := \max k_i$ and let

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

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

For each $i \in [r]$ let $\mathcal{H}_{n,i}$ be the hypergraph with $V(\mathcal{H}_{n,i}) := [n]$ and an edge set which consists of all k_i-distinct solutions to $\mathcal{L}_i$ in $[n]$. Observe that $\mathcal{H}_{n,i}$ is k_i-uniform and an independent set in $\mathcal{H}_{n,i}$ is an $\mathcal{L}_i$-free set.

For each $i \in [r]$ we define $\mathcal{F}_{n,i} := \{F \subseteq V(\mathcal{H}_{n,i}) : e(\mathcal{H}_{n,i}[F]) \geq \varepsilon' e(\mathcal{H}_{n,i})\}$. Note that since $\varepsilon' < d$, we have

$$\varepsilon' n^{k_i - \ell_i} \leq e(\mathcal{H}_{n,i}). \tag{4.1}$$

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

$$p = p(n) := n^{-1/m(A_1)}.$$

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

$$|F| \geq e(\mathcal{H}_{n,i}[F])^{\frac{1}{k_i-\ell_i}} \geq (\varepsilon' e(\mathcal{H}_{n,i}))^{\frac{1}{k_i-\ell_i}} \overset{(4.1)}{\geq} ((\varepsilon')^2 n^{k_i-\ell_i})^{\frac{1}{k_i-\ell_i}} \geq \varepsilon' n,$$

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

For each $j \in [k_i]$, we the number of hyperedges containing some $\{y_1, \ldots, y_j\} \subseteq V(\mathcal{H}_{n,i})$. Suppose $(x_1, \ldots, x_{k_i})$ is a k_i-distinct solution to $\mathcal{L}_i$ so that $\{y_1, \ldots, y_j\} \subseteq \{x_1, \ldots, x_{k_i}\}$. There are $k_i!/(k_i - j)!$ choices for picking the j roles the y_i play in $(x_1, \ldots, x_{k_i})$. Let W be one such choice for the set of indices of the x_a used by $\{y_1, \ldots, y_j\}$. In this case, Corollary 4.6 implies there are at most $n^{k_i-j-\mathrm{rank}((A_i)_{\overline{W}})}$ such solutions to $\mathcal{L}_i$, and if $j = 1$, there are at most $n^{k_i-\ell_i-1}$ such solutions. So for $j = 1$ this yields

$$\deg_{\mathcal{H}_{n,i}}(y_1) \leq k_i n^{k_i-\ell_i-1} \overset{(4.1)}{\leq} \frac{k_i}{\varepsilon'} \frac{e(\mathcal{H}_{n,i})}{v(\mathcal{H}_{n,i})} \leq c \frac{e(\mathcal{H}_{n,i})}{v(\mathcal{H}_{n,i})}.$$

For $j \geq 2$, by Proposition 4.3(iii) we have $k_i - j - \mathrm{rank}((A_i)_{\overline{W}}) \leq k_i - \ell_i - 1 - (j - 1)/m(A_i)$. Also $m(A_1) \geq m(A_i)$ for all $i \in [r]$ and hence we have

$$\begin{aligned}\deg_{\mathcal{H}_{n,i}}(\{y_1, \ldots, y_j\}) &\leq k_i! n^{k_i-\ell_i-1-\frac{j-1}{m(A_i)}} \leq k_i! n^{k_i-\ell_i-1-\frac{j-1}{m(A_1)}} \\ &\leq \frac{k_i!}{\varepsilon'} p^{j-1} \frac{e(\mathcal{H}_{n,i})}{v(\mathcal{H}_{n,i})} \leq c p^{j-1} \frac{e(\mathcal{H}_{n,i})}{v(\mathcal{H}_{n,i})}.\end{aligned}$$

Since $\{y_1, \ldots, y_j\}$ was arbitrary, we therefore have $\Delta_j(\mathcal{H}_{n,i}) \leq cp^{j-1} e(\mathcal{H}_{n,i})/v(\mathcal{H}_{n,i})$, as required. We have therefore shown that $\mathcal{H}_{n,i}$ and $\mathcal{F}_{n,i}$ satisfy the hypotheses of Proposition 3.3 for all $i \in [r]$.

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

(a) if $(S_1, \ldots, S_r) \in \mathcal{S}_r$, then $\sum_{i\in[r]} |S_i| \leq Dpn$;
(b) every $S \in \mathcal{S}_r$ satisfies $S \in \mathcal{I}(\mathcal{H}_{n,1}, \ldots, \mathcal{H}_{n,r})$;
(c) for every $(I_1, \ldots, I_r) \in \mathcal{I}(\mathcal{H}_{n,1}, \ldots, \mathcal{H}_{n,r})$, we have that $S \subseteq (I_1, \ldots, I_r) \subseteq S \cup f'(S)$, where $S := g(I_1, \ldots, I_r)$.

Note that $\mathcal{I}(\mathcal{H}_{n,1}, \ldots, \mathcal{H}_{n,r}) = \mathcal{I}(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$. For each $S \in \mathcal{S}_r$, define

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

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

Given any $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$ write $f(S) =: (f(S_1), \ldots, f(S_r))$ and $f'(S) =: (f'(S_1), \ldots, f'(S_r))$. (Note the slight abuse of the use of the f and f' notation here.) By definition of $\mathcal{F}_{n,i}$ any $F \in \overline{\mathcal{F}_{n,i}}$ contains at most $\varepsilon' n^{k_i-\ell_i}$ k_i-distinct solutions to $\mathcal{L}_i$. By Corollary 4.6, the number of k_i-distinct solutions to $\mathcal{L}_i$ in $[n]$ that use at least one element from S_i is at most $k_i n^{k_i-\ell_i-1}|S_i|$. Further,

$$k_i n^{k_i-\ell_i-1}|S_i| \leq k_i Dpn^{k_i-\ell_i} \leq \varepsilon' n^{k_i-\ell_i}.$$

Here, the first inequality holds by (a), and the second since $p = n^{-1/m(A_1)}$ and $m(A_1) > 0$ by Proposition 4.3(v). Thus, in total $f(S_i) = S_i \cup f'(S_i)$ contains at most $2\varepsilon' n^{k_i-\ell_i} \leq \delta n^{k_i-\ell_i}$ k_i-distinct solutions to $\mathcal{L}_i$, so (iv)(a) holds.

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

4.4. The number of $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subsets of $[n]$. Our first application of Theorem 4.7 yields an enumeration result (Theorem 4.2) for the number of $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subsets of $[n]$.

Proof of Theorem 4.2. By definition of $\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$ there are at least $2^{\mu(n,\mathcal{L}_1,\ldots,\mathcal{L}_r)}$ $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-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 $A_1, \ldots, A_r$ with parameter δ. We obtain a collection $\mathcal{S}_r$ and function f as in Theorem 4.7. Consider any $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset X of $[n]$. Note that X has a partition $X_1, \ldots, X_r$ so that $(X_1, \ldots, X_r) \in \mathcal{I}(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$. So by Theorem 4.7(i) this means there is some $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$ so that $X \subseteq \cup_{i\in[r]} f(S_i)$.

Further, given any $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$, we have that

$$|\cup_{i\in[r]} f(S_i)| \leq \mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r) + \delta n.$$

Thus, each such $\cup_{i\in[r]} f(S_i)$ contains at most $2^{\mu(n,\mathcal{L}_1,\ldots,\mathcal{L}_r)+\delta n}$ $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subsets of $[n]$. Note that, by Theorem 4.7(ii),

$$|\mathcal{S}_r| \leq \left(\sum_{s=0}^{Dn^{\frac{m(A_1)-1}{m(A_1)}}} \binom{n}{s} \right)^r < 2^{\delta n},$$

where the last inequality holds since n is sufficiently large.

Altogether, this implies that the number of $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subsets of $[n]$ is at most

$$2^{\delta n} \times 2^{\mu(n,\mathcal{L}_1,\ldots,\mathcal{L}_r)+\delta n} = 2^{\mu(n,\mathcal{L}_1,\ldots,\mathcal{L}_r)+2\delta n}.$$

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

4.5. The resilience of being $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-Rado. Recall that the resilience of S with respect to $\mathcal{P}$, $\mathrm{res}(S, \mathcal{P})$, 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 $\mathrm{res}([n]_p, (\mathcal{L}_1, \ldots, \mathcal{L}_r)\text{-Rado})$ for irredundant matrices $A_1, \ldots, A_r$ 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 $A_1, \ldots, A_r$ be matrices as in the statement of the theorem. Given n, if $p > n^{-1/m(A_1)}$, then since $m(A_1) > 1$ by Proposition 4.3(v), Proposition 2.1 implies that, w.h.p.,

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

We first show that

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\mathrm{res}([n]_p, (\mathcal{L}_1, \ldots, \mathcal{L}_r)\text{-Rado})}{|[n]_p|} \leq 1 - \frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta\right] = 1 \quad \text{if } p > n^{-1/m(A_1)}.$$

For this, we must show that the probability of the event that there exists a set $S \subseteq [n]_p$ such that $|S| \geq (\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)/n - \delta)|[n]_p|$ and S is $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free tends to one as n tends to infinity. This indeed follows: Let T be an $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset of $[n]$ of maximum size $\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$. Then, by Proposition 2.1, w.h.p. we have $|T \cap [n]_p| = (\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)/n \pm \delta)|[n]_p|$, and $T \cap [n]_p$ is $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-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 $p > Cn^{-1/m(A_1)}$,

$$\mathbb{P}\left[\text{res}([n]_p, (\mathcal{L}_1, \ldots, \mathcal{L}_r)\text{-Rado}) \geq \left(1 - \frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} - \delta\right)|[n]_p|\right] \to 1 \quad \text{as } n \to \infty. \tag{4.3}$$

Suppose n is sufficiently large. Apply Theorem 4.7 with parameters r, $\delta/8, A_1, \ldots, A_r$ 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 $p \geq Cn^{-1/m(A_1)}$.

Since (4.2) holds w.h.p., to prove (4.3) holds it suffices to show that the probability $[n]_p$ contains an $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset of size at least $(\frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta/2)np$ tends to zero as n tends to infinity.

Suppose that $[n]_p$ does contain an $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset I of size at least $(\frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta/2)np$. Note that I has a partition $I_1, \ldots, I_r$ so that $(I_1, \ldots, I_r) \in \mathcal{I}(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$. Further, there is some $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$ such that $S \subseteq (I_1, \ldots, I_r) \subseteq f(S)$. Thus, $[n]_p$ must contain $\cup_{i \in [r]} S_i$ as well as at least $(\frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta/4)np$ elements from $\left(\cup_{i \in [r]} f(S_i)\right) \setminus \left(\cup_{i \in [r]} S_i\right)$. (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 $|\left(\cup_{i \in [r]} f(S_i)\right) \setminus \left(\cup_{i \in [r]} S_i\right)| \leq \mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r) + \delta n/8$ by Theorem 4.7(iv)(b). So by the first part of Proposition 2.1, the probability $[n]_p$ contains at least $(\frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta/4)np$ elements from $\left(\cup_{i \in [r]} f(S_i)\right) \setminus \left(\cup_{i \in [r]} S_i\right)$ is at most $\exp(-\delta^2 np/256)$.

Write $N := n^{(m(A_1)-1)/m(A_1)}$ and $\gamma := \delta^2/256$. Given some $0 \leq s \leq DN$, there are at most $r^s \binom{n}{s}$ elements $(S_1, \ldots, S_r) \in \mathcal{S}_r$ such that $|\cup_{i \in [r]} S_i| = s$. Indeed, this follows since there are r^s ways to partition a set of size s into r classes. (Note we only need to consider $s \leq DN$ by Theorem 4.7(ii).) Thus, the probability $[n]_p$ does contain an $(\mathcal{L}_1, \ldots, \mathcal{L}_r)$-free subset I of size at least $(\frac{\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)}{n} + \delta/2)np$ is at most

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

which tends to zero as n tends to infinity. This completes the proof. ∎

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 $\mu(n, \mathcal{L}_1, \ldots, \mathcal{L}_r)$. For the rest of this section consider the symmetric case ($A := A_1 = \cdots = A_r$) and assume that A is a $1 \times k$ matrix, i.e., we are interested in solutions to a linear equation $a_1 x_1 + \cdots + a_k x_k = 0$. 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 [60]). Determining exact bounds remains open in many cases, famously including *progression-free sets* (where $\mathcal{L}$ is $x + y = 2z$). See [6, 15, 28] 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-coloring 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, $(x, \ldots, x)$ is a solution to $\mathcal{L}$ for all $x \in [n]$ (as observed by Frankl, Graham, and Rödl [23, 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 give 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., [43, Theorem 2], since such $\mathcal{L}$ have $o(n^{k-\ell})$ non-k-distinct solutions in $[n]$ (i.e., a solution $(x_1, \ldots, x_k)$ 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 nondistinct 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 $m + 1$.

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

THEOREM 4.8 (see [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 [32] proof that $\mu(n, 2) = n - \lfloor \frac{n}{5} \rfloor$. (To see why this is a lower bound, consider the set $\{x \in [n] : x \equiv 1 \text{ or } 4 \mod 5\} \cup \{y \in [n] : y \equiv 2 \text{ or } 3 \mod 5\}$.) 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). ∎

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

Note that Theorem 4.10 does indeed recover Hu's bound [32] 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 $S_1 \dot\cup \cdots \dot\cup S_r \dot\cup R = [n]$, 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 $\{j_1, \ldots, j_r\} = [r]$ and subsets $D_0, D_1, \ldots, D_r$ 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 $\{j_1, \ldots, j_i\} \subseteq [r]$ and $D_0, D_1, \ldots, D_i$ such that $P_1(t)$–$P_3(t)$ hold for all $t \in [i]$.

Suppose that $|D_i \cap \bigcup_{t \in [r] \setminus \{j_1, \ldots, j_i\}} S_t| \leq (\ell(r-i)-1)q$. Then we have that

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

a contradiction. So by averaging, there exists $j_{i+1} \in [r] \setminus \{j_1, \ldots, j_i\}$ 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 $D_i \cap S_{j_{i+1}} \supseteq \{s_{i,0} < \cdots < s_{i,\ell(r-i-1)q}\}$. 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. ∎

4.7. Open problem. We conclude the section with an open problem. Recall Hu [32] 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-colored analogue of the Cameron–Erdős conjecture [11], which was famously resolved by Green [26] and independently Sapozhenko [61].

Since our paper was submitted, Tran [71] has proved a slight variant of Conjecture 4.11; that is, he proves the result where one instead defines sum-free to also forbid nondistinct sums $x+x=z$ (as in the previous section). Note Tran's result does not quite imply Conjecture 4.11 directly.

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 (Proposition 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 do the following:

(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 $G^{(k)}_{n,p}$ 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 $G \subseteq G^{(k)}_{n,p}$ with $e(G) > t + \Omega(pn^k)$ is $(H_1, \dots, H_r)$-Ramsey. Further, there is some $G' \subseteq G^{(k)}_{n,p}$ with $e(G') > t - o(pn^k)$ which is *not* $(H_1, \dots, H_r)$-Ramsey.

(3) As a corollary of (2), we see that, whenever $p \geq Cn^{-1/m_k(H_1)}$, the random hypergraph $G^{(k)}_{n,p}$ is $(H_1, \dots, H_r)$-Ramsey with high probability.

The statements of (1)–(3) all involve a common parameter: the maximum size $\mathrm{ex}^r(n; H_1, \dots, H_r)$ of an n-vertex k-uniform hypergraph which is not $(H_1, \dots, H_r)$-Ramsey. For this reason, we generalize the classical supersaturation result of Erdős and Simonovits [19] to show that any n-vertex k-uniform hypergraph G with at least $\mathrm{ex}^r(n; H_1, \dots, H_r) + \Omega(n^k)$ edges is somehow "strongly" $(H_1, \dots, H_r)$-Ramsey. In the graph case, an old result of Burr, Erdős, and Lovász [9] allows us to quite accurately determine $\mathrm{ex}^r(n; H_1, \dots, H_r)$.

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-coloring* is a function $\sigma : E(G) \to [r]$. Given k-graphs $H_1, \dots, H_r$, we say that σ 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-coloring.

Given an integer $\ell \geq k$, denote by $K^{(k)}_\ell$ the complete k-graph on ℓ vertices. A k-graph H is *k-partite* if the vertices of H can be k-colored so that each edge contains one vertex of each color. 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 $\mathrm{ex}(n; H)$ the maximum size of an n-vertex H-free k-graph. Define the *Turán density* $\pi(H)$ of H by

$$\pi(H) := \lim_{n \to \infty} \frac{\mathrm{ex}(n; H)}{\binom{n}{k}} \tag{5.1}$$

(which exists by a simple averaging argument; see [36]). The so-called supersaturation phenomenon discovered by Erdős and Simonovits [19] 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. Note supersaturation problems date back to a result of Rademacher (see [16]).

THEOREM 5.1 (see [19]). *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 [20] says that for all graphs H, the value of $\pi(H)$ is determined by the chromatic number $\chi(H)$ of H, via

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

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 nontrivial 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 [37].

In this section, we generalize Theorem 5.1 from H-free hypergraphs to non$(H_1, \ldots, H_r)$-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 *ε-weakly $(H_1, \ldots, H_r)$-Ramsey* if there exists an r-coloring σ 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 *ε-strongly $(H_1, \ldots, H_r)$-Ramsey*. Note that ε-weakly $(H_1, \ldots, H_r)$-Ramsey graphs may not in fact be $(H_1, \ldots, H_r)$-Ramsey.

Using a well-known averaging argument of Katona, Nemetz, and Simonovits [36], we can show that $\binom{n}{k}^{-1}\mathrm{ex}^r(n; H_1, \ldots, H_r)$ converges as n tends to infinity. Indeed, let G be an n-vertex non-$(H_1, \ldots, H_r)$-Ramsey graph with $e(G) = \mathrm{ex}^r(n; H_1, \ldots, H_r)$. 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 $\binom{n-1}{k}^{-1} \cdot \mathrm{ex}^r(n-1; H_1, \ldots, H_r)$; otherwise G would contain an $(n-1)$-vertex subgraph which is $(H_1, \ldots, H_r)$-Ramsey, violating the choice of G. We have shown that

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

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

$$\pi(H_1, \ldots, H_r) := \lim_{n \to \infty} \frac{\mathrm{ex}^r(n; H_1, \ldots, H_r)}{\binom{n}{k}}.$$

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

that $K_m^{(k)}$ is $(H_1, \ldots, H_r)$-Ramsey. Thus

$$\pi(H_1, \ldots, H_r) \leq \pi(K_s^{(k)}), \tag{5.3}$$

which is at most $1 - \binom{s-1}{k-1}^{-1}$ (de Caen [10]). An interesting question is for which $H_1, \ldots, H_r$ 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 generalizes Theorem 5.1 to $r \geq 1$. The proof follows a standard approach to proving supersaturation results.

THEOREM 5.2. *For all $\delta > 0$, integers $r \geq 1$ and $k \geq 2$, and k-graphs $H_1, \ldots, H_r$, there exist $n_0, \varepsilon > 0$ such that for all integers $n \geq n_0$, every n-vertex k-graph G with $e(G) \geq (\pi(H_1, \ldots, H_r) + \delta)\binom{n}{k}$ is ε-strongly $(H_1, \ldots, H_r)$-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 $m \geq m_0$,

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

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 $e(G) = (\pi(H_1, \ldots, H_r) + \delta)\binom{n}{k}$. We need to show that, for every r-coloring σ 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 σ.

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

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

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

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

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 color 1 under σ. 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 ε-strongly $(H_1,\ldots,H_r)$-Ramsey, as required. ∎

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 $V_1,\ldots,V_s$ 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 $\mathcal{H}_1,\ldots,\mathcal{H}_r$ of graphs, the* Ramsey number $R(\mathcal{H}_1,\ldots,\mathcal{H}_r)$ *is the least m such that any r-coloring of K_m contains an i-colored copy of H_j for some $i\in[r]$ and some $H_j\in\mathcal{H}_i$. If $\mathcal{H}_i=\{K_{\ell_i}\}$ for all $i\in[r]$, then we instead write $R(\ell_1,\ldots,\ell_r)$, and simply $R^r(\ell)$ in the case when $\ell_1=\cdots=\ell_r=:\ell$.*

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

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

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

Thus

$$\pi(H_1,\ldots,H_r)=1-\frac{1}{R_\chi(H_1,\ldots,H_r)-1}=\pi\left(K_{R_\chi(H_1,\ldots,H_r)}\right). \tag{5.5}$$

The first inequality in (5.4) follows by definition of $\mathrm{ex}^r(n;H_1,\ldots,H_r)$; the second from (5.2) applied with a graph H which is $(H_1,\ldots,H_r)$-Ramsey and has $\chi(H)=R_\chi(H_1,\ldots,H_r)$. Clearly, then, $\pi(H_1,\ldots,H_r)=\pi(J_1,\ldots,J_r)$ if and only if $R_\chi(J_1,\ldots,J_r)=R_\chi(H_1,\ldots,H_r)$. 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 [5], one can determine $\mathrm{ex}^r(n;H_1,\ldots,H_r)$ exactly in the case when $H_1,\ldots,H_r$ are cliques of equal size.

Theorem 5.4 (see [5]). *For all positive integers $\ell,n\geq 3$ and $r\geq 1$, we have* $\mathrm{ex}^r(n;K_\ell,\ldots,K_\ell)=t_{R^r(\ell)-1}(n)$.

Thus in this case (5.3) is tight. The chromatic Ramsey number was introduced by Burr, Erdős, and Lovász [9], who showed that, in principle, one can determine R_χ given the usual Ramsey number R. A *graph homomorphism* from a graph H to a graph K is a function $\phi: V(H)\to V(K)$ such that $\phi(x)\phi(y)\in E(K)$ whenever

$xy \in E(H)$. Let $\mathrm{Hom}(H)$ denote the set of all graphs K such that there exists a graph homomorphism ϕ for which $K = \phi(H)$. Since there exists a homomorphism from H into K_ℓ if and only if $\chi(H) \leq \ell$, we also have that $R(\mathrm{Hom}(H)) = \chi(H)$. Thus $R(\mathrm{Hom}(H)) = R_\chi(H)$. In fact this relationship extends to all $r \geq 1$.

LEMMA 5.5 (see [9, 12, 46]). *For all integers $r \in \mathbb{N}$ and graphs $H_1, \ldots, H_r$,*

$$R_\chi(H_1, \ldots, H_r) = R(\mathrm{Hom}(H_1), \ldots, \mathrm{Hom}(H_r)).$$

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

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

The second statement is a corollary of the first since $\mathrm{Hom}(K_\ell) = \{K_\ell\}$. Another observation (see [9]) 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, $\mathrm{Hom}(H) \supseteq \mathrm{Hom}(G)$ and so $R_\chi(H, \ldots, H) = R_\chi(G, \ldots, G)$. Let F be an n-vertex graph with $e(F) = \mathrm{ex}^r(n; G, \ldots, G)$ which is not (G, r)-Ramsey. Obtain a graph T by adding an edge e to F. Then there exists an r-coloring of T in which every monochromatic copy of G contains e (the monochromatic-G-free coloring of F, with e arbitrarily colored). Hence T is not (H, r)-Ramsey and so

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

We say that a graph H is *(weakly) color-critical* if there exists $e \in E(H)$ for which $\chi(H - e) < \chi(H)$. Complete graphs and odd cycles are examples of color-critical graphs. The following conjecture would generalize 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 color-critical graph. Then, whenever n is sufficiently large,*

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

If true, this conjecture would also generalize a well-known result of Simonovits [66], which extends Turán's theorem to color-critical graphs. It would also determine $\mathrm{ex}^r(n; H, \ldots, H)$ 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 [42] 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 [52].

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

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

The next result of Prömel and Steger [52] immediately tells us the typical structure of non-Ramsey graphs in certain cases.

THEOREM 5.7 (see [52]). *For every graph H, the following holds. Almost all H-free graphs are $(\chi(H)-1)$-partite if and only if H is color-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 color-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. ∎

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 $\overline{\mathrm{Ram}}(n;H_1,\dots,H_r)$ is the set of n-vertex k-graphs which are not $(H_1,\dots,H_r)$-Ramsey and $\mathrm{Ram}(H_1,\dots,H_r)$ 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 a non-$(H,1)$-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 $G\in\overline{\mathrm{Ram}}(n;H_1,\dots,H_r)$, 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 Proposition 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 $k\ge 2$, 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 (see [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 $p=n^{-1/m_k(H)}$,*

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

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 $k\ge 2$ and $\delta>0$. Let $H_1,\dots,H_r$ be k-graphs such that $m_k(H_1)\ge\dots\ge m_k(H_r)$ and $\Delta_1(H_i)\ge 2$ 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 $f:\mathcal{S}_r\to(\mathcal{G}_k(n))^r$ such that the following hold.*

(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 $\sum_{i\in[r]}e(S_i)\le Dn^{k-1/m_k(H_1)}$.*
(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 δ-weakly $(H_1,\dots,H_r)$-Ramsey; and*
  (b) *$e\left(\bigcup_{i\in[r]}f(S_i)\right)\le\mathrm{ex}^r(n;H_1,\dots,H_r)+\delta\binom{n}{k}$.*

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 δ) 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 Proposition 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,

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

where $\mathrm{Aut}(H_i)$ 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 Proposition 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

$$v_i!\binom{n-k}{v_i-k}. \tag{5.7}$$

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

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

where, in the final inequality, we used the fact that $|\mathrm{Aut}(H_i)| \leq v_i!$. 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 Proposition 3.3 for all $i \in [r]$.

Then Proposition 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 $f' : \mathcal{S}_r \to \prod_{i\in[r]} \overline{\mathcal{F}_{n,i}}$ and $g : \mathcal{I}(\mathcal{H}_{n,1}, \ldots, \mathcal{H}_{n,r}) \to \mathcal{S}_r$ such that the following conditions hold:

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

Note that $(G_1, \ldots, G_r) \in \mathcal{I}(\mathcal{H}_{n,1}, \ldots, \mathcal{H}_{n,r})$ if and only if $(G_1, \ldots, G_r) \in \mathcal{I}_r(n; H_1, \ldots, H_r)$ (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 $f : \mathcal{S}_r \to \mathcal{P}(\binom{[n]}{k})^r$. (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 $(S_1, \ldots, S_r) \in \mathcal{S}_r$ 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 $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$ write $f(S) =: (f(S_1), \ldots, f(S_r))$ and $f'(S) =: (f'(S_1), \ldots, f'(S_r))$. 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-coloring σ 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-coloring σ 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 colored 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

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

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 ε-weakly $(H_1, \ldots, H_r)$-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. ∎

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 $(S_1, \ldots, S_r) \in \mathcal{S}_r$ 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 $|\overline{\mathrm{Ram}}(n; H_1, \dots, H_r)|$.

Proof of Theorem 1.12. Let $0 < \delta < 1$ be arbitrary, and let $n \in \mathbb{N}$ be sufficiently large. Clearly, $|\overline{\mathrm{Ram}}(n; H_1, \dots, H_r)| \geq 2^{\mathrm{ex}^r(n; H_1, \dots, H_r)}$ since no subhypergraph of an n-vertex non-$(H_1, \dots, H_r)$-Ramsey k-graph with $\mathrm{ex}^r(n; H_1, \dots, H_r)$ 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 δ. We obtain a collection $\mathcal{S}_r$ and a function f as in Theorem 5.11. Consider any $G \in \overline{\mathrm{Ram}}(n; H_1, \dots, H_r)$. 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

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

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

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

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 $H \in \overline{\mathrm{Ram}}(n; H_1, \dots, H_r)$. Then there exists an r-coloring σ 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 $J \in \overline{\mathrm{Ram}}(n; H_2, \dots, H_r)$ 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 [17] (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,

$$\begin{aligned}
|\overline{\mathrm{Ram}}(n; H_1, \dots, H_r)| &\leq \sum_{J \in \overline{\mathrm{Ram}}(n; H_2, \dots, H_r)} \sum_{e(J')=0}^{(h-1)\binom{n-1}{k-1}} \binom{\binom{n}{k}}{e(J')} \\
&= |\overline{\mathrm{Ram}}(n; H_2, \dots, H_r)| \sum_{e(J')=0}^{\frac{k(h-1)}{n}\binom{n}{k}} \binom{\binom{n}{k}}{e(J')} \\
&\leq |\overline{\mathrm{Ram}}(n; H_2, \dots, H_r)| \cdot 2^{\delta \binom{n}{k}}.
\end{aligned}$$

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. ∎

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 $\mathrm{Forb}(n;\mathcal{F})$ 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 [50] asymptotically determines the logarithm of $|\mathrm{Forb}(n;\mathcal{F})|$. (This generalizes the corresponding result of Erdős, Frankl, and Rödl [18] for graphs.) Let

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

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

THEOREM 5.12 (Theorem 2.3, [50]). *Let $k \geq 2$ be a positive integer and $\mathcal{F}$ be a (possibly infinite) family of k-graphs. Then, for all $n \in \mathbb{N}$,*

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

Since $G \in \overline{\mathrm{Ram}}(n;H_1,\ldots,H_r)$ if and only if G is an n-vertex k-graph without a copy of any $F \in \mathrm{Ram}(H_1,\ldots,H_r)$ as a subhypergraph, Theorem 5.12 immediately implies Theorem 1.12.

5.6. The resilience of being $(H_1,\ldots,H_r)$-Ramsey. Recall that $G^{(k)}_{n,p}$ 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 $\mathrm{res}(G^{(k)}_{n,p},(H_1,\ldots,H_r)\text{-Ramsey})$ for given fixed k-graphs $H_1,\ldots,H_r$. Explicitly, $\mathrm{res}(G^{(k)}_{n,p},(H_1,\ldots,H_r)\text{-Ramsey})$ is the minimum integer t such that one can remove t edges from $G^{(k)}_{n,p}$ to obtain a k-graph H which has an $(H_1,\ldots,H_r)$-free r-coloring.

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 $H_1,\ldots,H_r$ with $m_k(H_1) \geq \cdots \geq m_k(H_r)$ and $\Delta_1(H_i) \geq 2$ for all $i \in [r]$, there exists $C > 0$ such that*

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

In the case when $m_k(H_1) = m_k(H_2)$, Corollary 5.13 generalizes 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 $H_1,\ldots,H_r$ 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.,

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

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

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

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

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

$$\mathbb{P}\left[\mathrm{res}(G^{(k)}_{n,p}, (H_1, \ldots, H_r)\text{-Ramsey}) \geq (1 - \pi - \delta)e(G^{(k)}_{n,p})\right] \to 1 \quad \text{as} \quad n \to \infty. \tag{5.10}$$

Suppose n is sufficiently large. Apply Theorem 5.11 with parameters $r, k, \delta/16$, $(H_1, \ldots, H_r)$ 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 $G^{(k)}_{n,p}$ contains a non-$(H_1, \ldots, H_r)$-Ramsey k-graph with at least $(\pi + \delta/2)p\binom{n}{k}$ edges tends to zero as n tends to infinity.

Suppose that $G^{(k)}_{n,p}$ does contain a non-$(H_1, \ldots, H_r)$-Ramsey k-graph I with at least $(\pi + \delta/2)p\binom{n}{k}$ edges. Then there exist pairwise edge-disjoint k-graphs $I_1, \ldots, I_r$ such that $\bigcup_{i\in[r]} I_i = I$ and $(I_1, \ldots, I_r) \in \mathcal{I}_r(n; H_1, \ldots, H_r)$. Further, there is some $S = (S_1, \ldots, S_r) \in \mathcal{S}_r$ such that $S \subseteq (I_1, \ldots, I_r) \subseteq f(S)$. Thus, $G^{(k)}_{n,p}$ 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 $G^{(k)}_{n,p}$ 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 $G^{(k)}_{n,p}$ 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 $\exp(-\delta^2 p\binom{n}{k}/256) \leq \exp(-\delta^2 pn^k/256k^k)$.

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

$$\begin{aligned}
\sum_{s=0}^{DN} r^s \binom{\binom{n}{k}}{s} \cdot p^s \cdot e^{-\gamma n^k p} &\leq (DN+1)(rp)^{DN}\binom{\binom{n}{k}}{DN} e^{-\gamma n^k p} \\
&\leq (DN+1)\left(\frac{re^{k+1}pn^k}{DNk^k}\right)^{DN} e^{-\gamma n^k p} \\
&\leq (DN+1)\left(\frac{re^{k+1}C}{Dk^k}\right)^{DN} e^{-\gamma CN} \\
&\leq e^{\gamma CN/2} e^{-\gamma CN} = e^{-\gamma CN/2},
\end{aligned}$$

which tends to zero as n tends to infinity. This completes the proof. ∎

Acknowledgments. The authors are grateful to Victor Falgas-Ravry for a helpful conversation on [21], to Hong Liu for a helpful conversation on the graph Ramsey problems considered in this paper, and to Yury Person for a helpful conversation on Theorem 1.7. The authors are also grateful to the reviewers for their helpful and careful reviews.

REFERENCES

[1] H.L. Abbott and E.T.H. Wang, *Sum-free sets of integers*, Proc. Amer. Math. Soc., 67 (1977), pp. 11–16.
[2] J. Balogh, B. Csaba, and W. Samotij, *Local resilience of almost spanning trees in random graphs*, Random Structures Algorithms, 38 (2011), pp. 121–139.
[3] J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc., 28 (2015), pp. 669–709.
[4] J. Balogh, R. Morris, and W. Samotij, *The method of hypergraph containers*, submitted.
[5] A. Bialostocki, Y. Caro, and Y. Roditty, *On zero sum Turan numbers*, Ars Combin., 29 (1990), pp. 117–127.
[6] T.F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc., 93 (2016), pp. 643–663.
[7] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica, 7 (1987), pp. 35–38.
[8] J. Böttcher, Y. Kohayakawa, and A. Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, Combin. Probab. Comput., 22 (2013), pp. 639–683.
[9] S.A. Burr, P. Erdős, and L. Lovász, *On graphs of Ramsey type*, Ars Combin., 1 (1976), pp. 167–190.
[10] D. de Caen, *Extension of a theorem of Moon and Moser on complete subgraphs*, Ars Combin., 16 (1983), pp. 5–10.
[11] P. Cameron and P. Erdős, *On the number of sets of integers with various properties*, in Number Theory, R.A. Mollin, ed., Walter de Gruyter, Berlin, 1990, pp. 61–79.
[12] V. Chvátal, *private communication to the authors of* [9].
[13] D. Conlon, *Combinatorial theorems relative to a random set*, in Proceedings of the International Congress of Mathematicians 2014, Vol. 4, pp. 303–328.
[14] D. Conlon and W.T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math., 184 (2016), pp. 367–454.
[15] M. Elkin, *An improved construction of progression-free sets*, Israel J. Math., 184 (2011), pp. 93–128.
[16] P. Erdős, *On a theorem of Rademacher–Turán*, Illinois J. Math., 6 (1962), pp. 122–127.
[17] P. Erdős, *A problem on independent r-tuples*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., 8 (1965), pp. 93–95.
[18] 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), pp. 113–121.
[19] P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica, 3 (1983), pp. 181–192.
[20] P. Erdős and A.H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc., 52 (1946), pp. 1087–1091.
[21] V. Falgas-Ravry, K. O'Connell, and A. Uzzell, *Multicolour containers, extremal entropy and counting*, Random Structures Algorithms, 2018, https://doi.org/10.1002/rsa.20777.
[22] A. Ferber, G. McKinley, and W. Samotij, *Supersaturated sparse graphs and hypergraphs*, Int. Math. Res. Not. IMRN, to appear.
[23] P. Frankl, R.L. Graham, and V. Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Theory Ser. A, 47 (1988), pp. 246–261.
[24] E. Friedgut, V. Rödl, and M. Schacht, *Ramsey properties of random discrete structures*, Random Structures Algorithms, 37 (2010), pp. 407–436.
[25] R. Graham, V. Rödl, and A. Ruciński, *On Schur properties of random subsets of integers*, J. Number Theory, 61 (1996), pp. 388–408.
[26] B. Green, *The Cameron–Erdős conjecture*, Bull. Lond. Math. Soc., 36 (2004), pp. 769–778.
[27] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal., 15 (2005), pp. 340–376.
[28] B. Green and J. Wolf, *A note on Elkin's improvement of Behrend's construction*, in Additive Number Theory: Festschrift in Honor of the Sixtieth Birthday of Melvyn B. Nathanson, 1st ed., Springer-Verlag, New York, 2010, pp. 141–144.

[29] L. Gugelmann, R. Nenadov, Y. Person, A. Steger, N. Škorić, and H. Thomas, *Symmetric and asymmetric Ramsey properties in random hypergraphs*, Forum Math. Sigma, 5 (2017), E28.

[30] P.E. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: Forbidding even cycles*, J. Combin. Theory B, 64 (1995), pp. 273–287.

[31] P.E. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: Forbidding odd cycles*, Combinatorica, 16 (1996), pp. 107–122.

[32] M.C. Hu, *A note on sum-free sets of integers*, Proc. Amer. Math. Soc., 80 (1980), pp. 711–712.

[33] Y. Ishigami, *The Number of Hypergraphs and Colored Hypergraphs with Hereditary Properties*, preprint.

[34] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[35] S. Janson and A. Ruciński, *Upper tails for counting objects in randomly induced subhypergraphs and rooted random graphs*, Ark. Mat., 49 (2011), pp. 79–96.

[36] G. Katona, T. Nemetz, and M. Simonovits, *On a problem of Turán in the theory of graphs*, Mat. Lapok, 15 (1964), pp. 228–238.

[37] P. Keevash, *Hypergraph Turán problems*, in Surveys in Combinatorics, Cambridge University Press, Cambridge, 2011, pp. 83–140.

[38] D. Kleitman and K.J. Winston, *The asymptotic number of lattices*, Combinatorial Mathematics, Optimal Designs and Their Applications, Ann. Discrete Math. 6, Elsevier, New York, 1980, pp. 243–249.

[39] D. Kleitman and K.J. Winston, *On the number of graphs without* 4*-cycles*, Discrete Math., 6 (1982), pp. 167–172.

[40] Y. Kohayakawa and B. Kreuter, *Threshold functions for asymmetric Ramsey properties involving cycles*, Random Structures Algorithms, 11 (1997), pp. 245–276.

[41] Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K_4-free subgraphs of random graphs*, Combinatorica, 17 (1997), pp. 173–213.

[42] 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), pp. 637–671.

[43] D. Král', O. Serra, and L. Vena, *A removal lemma for systems of linear equations over finite fields*, Israel J. Math., 187 (2012), pp. 193–207.

[44] B.M. Landman and A. Robertson, *Ramsey Theory on the Integers*, Stud. Math. Libr. 24, AMS, Providence, RI, 2004.

[45] C. Lee and B. Sudakov, *Dirac's theorem for random graphs*, Random Structures Algorithms, 41 (2012), pp. 293–305.

[46] S. Lin, *On Ramsey numbers and K_r-coloring of graphs*, J. Combin. Theory, 12 (1972), pp. 82–92.

[47] M. Marciniszyn, J. Skokan, R. Spöhel, and A. Steger, *Asymmetric Ramsey properties of random graphs involving cliques*, Random Structures Algorithms, 34 (2009), pp. 419–453.

[48] R. Morris and D. Saxton, *The number of $C_{2\ell}$-free graphs*, Adv. Math., 298 (2016), pp. 534–580.

[49] F. Mousset, R. Nenadov, and W. Samotij, *Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties*, submitted.

[50] B. Nagle, V. Rödl, and M. Schacht, *Extremal hypergraph problems and the regularity method*, in Topics in Discrete Mathematics, Algorithms Combin. 26, 2006, pp. 247–278.

[51] R. Nenadov and A. Steger, *A short proof of the random Ramsey theorem*, Combin. Probab. Comput., 25 (2016), pp. 130–144.

[52] H. Prömel and A. Steger, *The asymptotic number of graphs not containing a fixed color-critical subgraph*, Combinatorica, 12 (1992), pp. 463–473.

[53] R. Rado, *Studien zur kombinatorik*, Math. Z., 36 (1933), pp. 424–470.

[54] V. Rödl and A. Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, in Combinatorics, Paul Erdős Is Eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Mathematical Society, Budapest, 1993, pp. 317–346.

[55] V. Rödl and A. Ruciński, *Random graphs with monochromatic triangles in every edge coloring*, Random Structures Algorithms, 5 (1994), pp. 253–270.

[56] V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc., 8 (1995), pp. 917–942.

[57] V. Rödl and A. Ruciński, *Rado partition theorem for random subsets of integers*, Proc. Lond. Math. Soc., 74 (1997), pp. 481–502.

[58] V. Rödl and A. Ruciński, *Ramsey properties of random hypergraphs*, J. Combin. Theory Ser. A, 81 (1998), pp. 1–33.

[59] V. Rödl, A. Ruciński, and M. Schacht, *Ramsey properties of random k-partite, k-uniform hypergraphs*, SIAM J. Discrete Math., 21 (2007), pp. 442–460.

[60] I.Z. Ruzsa, *Solving a linear equation in a set of integers* I, Acta Arith., 65 (1993), pp. 259–282.

[61] A.A. Sapozhenko, *The Cameron–Erdős conjecture*, Dokl. Akad. Nauk, 393 (2003), pp. 749–752 (Russian).

[62] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math., 201 (2015), pp. 925–992.

[63] D. Saxton and A. Thomason, *Online containers for hypergraphs, with applications to linear equations*, J. Combin. Theory Ser. B, 121 (2016), pp. 248–283.

[64] M. Schacht, *Extremal results for random discrete structures*, Ann. of Math., 184 (2016), pp. 331–363.

[65] I. Schur, *Über die Kongruenz $x^m + y^m \equiv z^m$ (mod p)*, Jber. Deutsch. Math. Verein., 25 (1916), pp. 114–117.

[66] M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, in Theory of Graphs, Academic Press, New York, 1968, pp. 279–319.

[67] C. Spiegel, *A note on sparse supersaturation and extremal results for linear homogeneous systems*, Electron. J. Combin., 24 (2017), P3.38.

[68] B. Sudakov, *Robustness of graph properties*, in Surveys in Combinatorics, Cambridge University Press, Cambridge, 2017, pp. 372–408.

[69] B. Sudakov and V.H. Vu, *Local resilience of graphs*, Random Structures Algorithms, 33 (2008), pp. 409–433.

[70] C. Terry, *Structure and enumeration theorems for hereditary properties in finite relational languages*, Ann. Pure Appl. Logic, 169 (2018), pp. 413–449.

[71] T. Tran, *On the structure of large sum-free sets of integers*, Israel J. Math., 228 (2018), pp. 249–292.

[72] B.L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk., 15 (1927), pp. 212–216.