

TYPICAL RAMSEY PROPERTIES OF THE PRIMES, ABELIAN GROUPS AND OTHER DISCRETE STRUCTURES

ANDREA FRESCHI, ROBERT HANCOCK AND ANDREW TREGLOWN

ABSTRACT. Given a matrix A with integer entries, a subset S of an abelian group and $r \in \mathbb{N}$, we say that S is (A, r) -Rado if any r -colouring of S yields a monochromatic solution to the system of equations $Ax = 0$. A classical result of Rado characterises all those matrices A such that \mathbb{N} is (A, r) -Rado for all $r \in \mathbb{N}$. Rödl and Ruciński [Proceedings of the London Mathematical Society, 1997] and Friedgut, Rödl and Schacht [Random Structures & Algorithms, 2010] proved a random version of Rado’s theorem where one considers a random subset of $[n] := \{1, \dots, n\}$ instead of \mathbb{N} .

In this paper, we investigate the analogous random Ramsey problem in the more general setting of abelian groups. Given a sequence $(S_n)_{n \in \mathbb{N}}$ of finite subsets of abelian groups, let $S_{n,p}$ be a random subset of S_n obtained by including each element of S_n independently with probability p . We are interested in determining the probability threshold $\hat{p} := \hat{p}(n)$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p = o(\hat{p}); \\ 1 & \text{if } p = \omega(\hat{p}). \end{cases}$$

Our main result, which we coin the *random Rado lemma*, is a general black box to tackle problems of this type. Using this tool in conjunction with a series of supersaturation results, we determine the probability threshold for a number of different cases. A consequence of the Green–Tao theorem [Annals of Mathematics, 2008] is the *van der Waerden theorem for the primes*: every finite colouring of the primes contains arbitrarily long monochromatic arithmetic progressions. Using our machinery, we obtain a random version of this result. We also prove a novel supersaturation result for $S_n := [n]^d$ and use it to prove an integer lattice generalisation of the random version of Rado’s theorem. Various threshold results for abelian groups are also given.

Furthermore, we prove a 1-statement (the $p = \omega(\hat{p})$ regime) and a 0-statement (the $p = o(\hat{p})$ regime) for hypergraphs that imply several of the previously known 1- and 0-statements in various settings, as well as our random Rado lemma.

1. INTRODUCTION

Ramsey theory concerns the study of partitions of mathematical objects, and in particular, what structures one can guarantee in such partitions. For example, Ramsey’s original theorem asserts that, given any $r, t \in \mathbb{N}$, if $n \in \mathbb{N}$ is sufficiently large then however one partitions the edge set of the complete graph K_n into r colour classes, at least one of these colour classes must contain a copy of K_t . Analogous Ramsey-type behaviour is also exhibited in arithmetic settings and Ramsey-type results for the integers have been studied since the end of the 19th Century, with early progress due to Hilbert [34] (1892), Schur [56] (1916) and van der Waerden [61] (1927). In 1933, Rado [45] characterised all those homogeneous systems of linear equations \mathcal{L} for which every finite colouring of \mathbb{N} yields a monochromatic solution to \mathcal{L} . In 1975, Deuber [19] resolved the analogous problem

Date: June 4, 2024.

AF: School of Mathematics, University of Birmingham, United Kingdom, axf079@bham.ac.uk.

RH: Mathematical Institute, University of Oxford, United Kingdom, robert.hancock@maths.ox.ac.uk. Research supported by ERC Advanced Grant 883810 and by a Humboldt Research Fellowship at Heidelberg University.

AT: School of Mathematics, University of Birmingham, United Kingdom, a.c.treglown@bham.ac.uk. Research supported by EPSRC grant EP/V048287/1.

where now one works in the setting of abelian groups rather than \mathbb{N} . See, e.g., [7, 12, 52, 57, 59] for further Ramsey-type results for groups and other arithmetic structures.

Similarly, it is natural to investigate *how many* copies of a given structure one can guarantee within a partition of a mathematical object. This is the topic of a branch of Ramsey theory known as *Ramsey multiplicity*, in which researchers initially focused on graphs. For example, a seminal result of Goodman [27] from 1959 determines the minimum number of monochromatic triangles in a 2-edge-colouring of the complete graph K_n ; see, e.g., the survey [10] for further results on the topic. More recently, there has been interest in studying Ramsey multiplicity in arithmetic structures, such as the integers (see, e.g., [17, 18, 28, 46, 55]) and abelian groups. In particular, after earlier work of Cameron, Cilleruelo and Serra [12], a 2017 paper of Saad and Wolf [52] initiated the systematic study of Ramsey multiplicity for systems of linear equations in abelian groups. A number of subsequent papers have made significant progress on the topic (see, e.g., [2, 20, 21, 37, 51, 63]).

In this paper, we consider yet another Ramsey-type question, this time concerning the *typical Ramsey properties* of mathematical objects. Broadly speaking, once it is established that any finite partition of a set S yields a copy of a given structure, we are interested in whether we typically expect to witness the same Ramsey-type behaviour in subsets of S of a given size. As we will see in the following sections, this problem is formalised by considering a random subset within S , and so we shall refer to it as the *random Ramsey problem*. This direction of research is part of the wider study of transferring (combinatorial) theorems into the random setting; see, e.g., the ICM survey of Conlon [14]. In particular, the random Ramsey problem for graphs and the integers has been well-studied, through the *random Ramsey theorem* [47, 48, 49] (see Theorem 4.1) and the *random Rado theorem* [26, 50] (see Theorems 4.2 and 4.3), respectively. The statements of both theorems, and of random Ramsey-type results in general, break down into two parts: the so-called 1-statement, asserting that random subsets of size above a certain threshold exhibit a Ramsey-type behaviour with high probability, and the 0-statement, asserting that random subsets of size below the threshold do not exhibit a Ramsey-type behaviour with high probability.

The main goal of this paper is to construct a general framework to study the random Ramsey problem for abelian groups. Our main result, which we call *the random Rado lemma* (Lemma 6.8), provides a machine that, on input of a homogeneous system of linear equations \mathcal{L} , a sequence of ‘well-behaved’ subsets S_n of abelian groups and a Ramsey-type supersaturation result¹ for the number of monochromatic solutions of \mathcal{L} in S_n , outputs an appropriate 1-statement and 0-statement. In order to rigorously state the random Rado lemma, we need to introduce various terminology as well as recall some fundamental algebraic concepts. As such, we defer the statement of Lemma 6.8 to Section 6. We give several highlight applications of the random Rado lemma. For example, we prove a random version of the *van der Waerden theorem for the primes* (Theorem 2.7) as well as a *random Rado theorem for integer lattices* (Theorem 5.1).²

Building towards the aforementioned theory for (subsets of) abelian groups, we also survey the random Ramsey problem for graphs and the integers. As explained in more detail in Sections 4–7, there is a well-known heuristic for what the threshold in the 1- and 0-statements of a random Ramsey-type result should be. Indeed, this intuition turns out to be correct in the setting of graphs and the integers as well as for the various new results we prove. Therefore, it would be highly desirable to obtain a unifying theory which implies all previously known random Ramsey-type results. In this direction, we prove a general 1-statement and a general 0-statement for hypergraphs (Theorem 7.7) which imply several of the previously known 1- and 0-statements for discrete structures. Furthermore, the random Rado lemma arises as a direct corollary of this

¹A Ramsey-type supersaturation result states that in any r -colouring of a given mathematical object, a *linear proportion* of a certain class of substructures is monochromatic.

²In fact, we deduce Theorem 5.1 from a simplified version of the random Rado lemma (Lemma 6.33).

hypergraph result. The proof of Theorem 7.7 makes use of the *hypergraph container method* as well as a recent argument of the second and third authors [33].

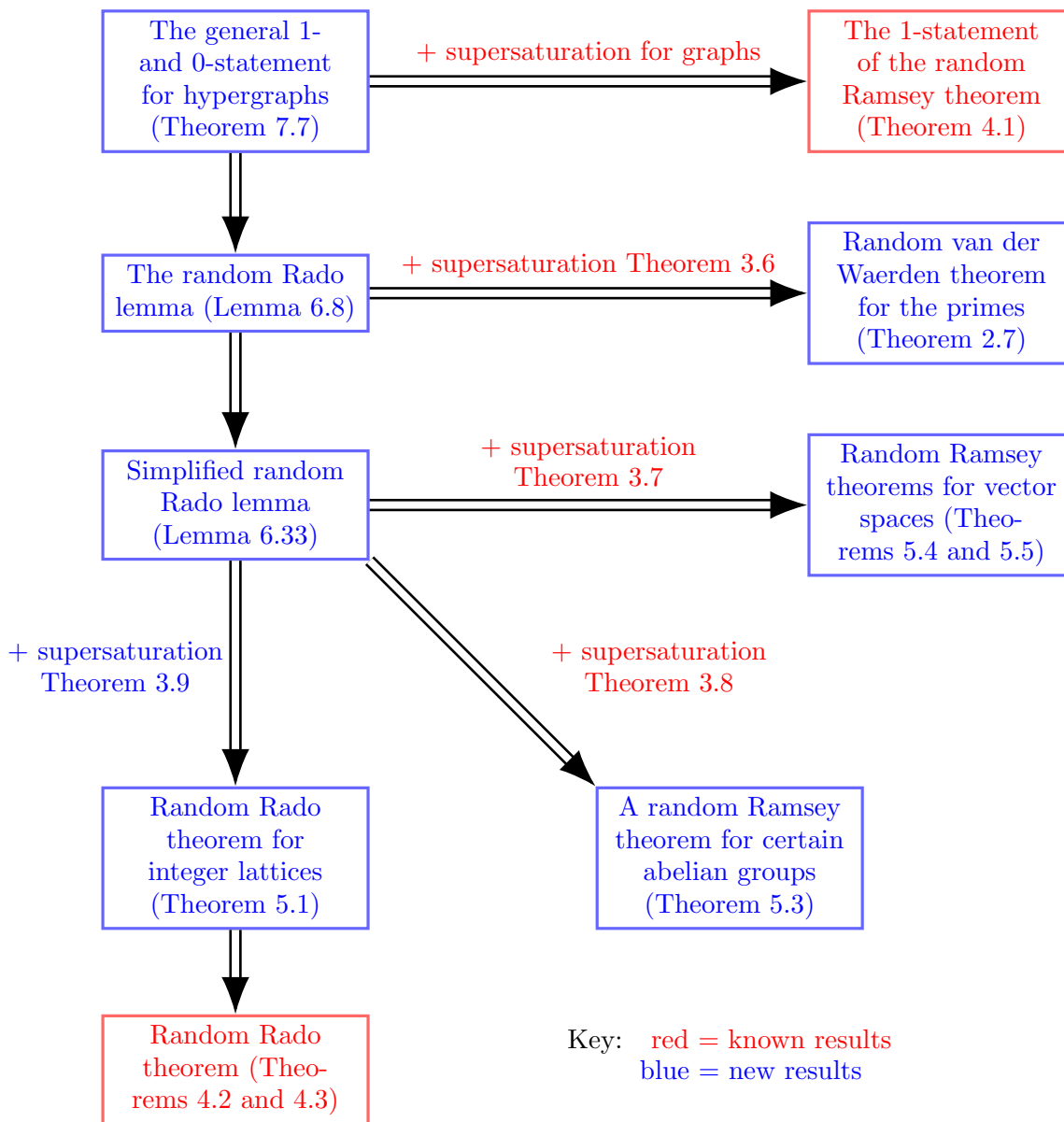


FIGURE 1. A visualisation of the main new (blue) and previously known (red) results covered in this paper.

1.1. Organisation of the paper. One of the most significant applications of our theory is to arithmetic progressions in the primes; we present our main contribution to this topic (Theorem 2.7) in Section 2. Sections 3 and 4 serve as a gentle introduction to some of the key results in the area. Namely, in Section 3 we briefly cover some of the classical themes in Ramsey theory, specifically partition regularity and supersaturation; in Section 4 we discuss the random Ramsey problem for graphs and the integers. In Section 5, we highlight the power of the random Rado lemma by presenting various novel results which we obtain from this black box. In Sections 6 and 7,

we formally state the random Rado lemma and our general 1- and 0-statements for hypergraphs, respectively. Note that Figure 1 illustrates how these black box results are connected, and some of their implications.

The rest of the paper is devoted to the technical proofs of our results. In Section 8, we prove the 1- and 0-statements for hypergraphs and in Section 9 we explain how to deduce the random Rado lemma from them. In Section 11, we give the proofs of all the applications of the random Rado lemma that appear in Sections 2 and 5. For these proofs, we apply a couple of new results that may be of independent interest: (i) a Ramsey-type supersaturation result for $[n]^d$ (Theorem 3.9) and (ii) an upper bound on the number of k -term arithmetic progressions in the primes that contain a given prime number (Lemma 2.3). The former generalises a well-known supersaturation result of Frankl, Graham and Rödl [22]. We prove these two results in Section 10.

In Section 12, we explain how, with the random Rado lemma at hand, one can attack further related problems. Indeed, we state further consequences of the method as well as discuss potential future research directions.

2. ARITHMETIC PROGRESSIONS IN THE PRIMES

In this section, we highlight an application of the random Rado lemma to arithmetic progressions in the primes. We defer stating our other applications of the random Rado lemma until Section 5 as they require the introduction of a little more notation and a few definitions.

2.1. The Green–Tao theorem. One of the most celebrated results of the 21st Century is the following theorem of Green and Tao [30]. Here we write \mathbf{P} for the set of prime numbers and, given $n \in \mathbb{N}$, we write \mathbf{P}_n for the set of all prime numbers less than or equal to n . Recall the prime number theorem asserts that $|\mathbf{P}_n| \sim n/\log n$, where throughout this paper we write \log for the natural logarithm. We abbreviate k -term arithmetic progressions by k -APs.

Theorem 2.1 (Green and Tao [30]). *Let $k \geq 2$. The set of primes \mathbf{P} contains arithmetic progressions of arbitrary length. Moreover, given any subset $S \subseteq \mathbf{P}_n$ of size $|S| = \Omega(|\mathbf{P}_n|)$, the number of k -APs in S is $\Theta(n^2/\log^k n)$.*³

The Green–Tao theorem therefore tells us that any linear size subset of \mathbf{P}_n is rich in arithmetic progressions, for any large $n \in \mathbb{N}$. It is natural to ask about sparser subsets of \mathbf{P}_n : given some $p = p(n)$, does a typical subset S of \mathbf{P}_n of size approximately $p|\mathbf{P}_n|$ contain k -APs? Let $\mathbf{P}_{n,p}$ denote the subset of \mathbf{P}_n obtained by retaining each element of \mathbf{P}_n with probability p , independently of all other elements.

Theorem 2.2 (Random Green–Tao theorem). *Let $k \geq 3$.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{P}_{n,p} \text{ contains a } k\text{-AP}] = \begin{cases} 0 & \text{if } p = o(n^{-2/k} \log n); \\ 1 & \text{if } p = \omega(n^{-2/k} \log n). \end{cases}$$

Theorem 2.2 tells us that typical subsets of \mathbf{P}_n of size significantly above $|\mathbf{P}_n|n^{-2/k} \log n \sim n^{(k-2)/k}$ contain k -APs; conversely, typical subsets of \mathbf{P}_n of size significantly smaller than $n^{(k-2)/k}$ do not. Note that, since Theorem 2.2 is not a Ramsey-type result, we do not use the random Rado lemma to prove it. Instead, we make use of the following result.

Lemma 2.3. *Let $k \in \mathbb{N}$ with $k \geq 2$ and let $\ell \in [k]$. For every prime $q \in \mathbf{P}_n$, the number of k -APs in \mathbf{P}_n such that the ℓ th term is equal to q is $O(n/\log^{k-1} n)$.*

³Note that the moreover part of Theorem 2.1 is implicit in [30]. An explicit proof of (a strengthening of) this part of Theorem 2.1 can be found in, e.g., [16, Theorem 1.2].

We prove Lemma 2.3 in Section 10.1 via a sieve theory argument. With Lemma 2.3 and Theorem 2.1 at hand, the proof of Theorem 2.2 is just a simple application of the second moment method. For completeness, a full proof is given in Appendix A.2.

Given $k \geq 3$ and $\delta > 0$, we say a finite set $X \subseteq \mathbb{N}$ is (δ, k) -Szemerédi if every subset of X of size at least $\delta|X|$ contains a k -AP. The following result tells us that not too sparse subsets of \mathbf{P}_n are typically (δ, k) -Szemerédi.

Theorem 2.4 (Balogh, Liu and Sharifzadeh [3]). *For any $\delta, \gamma > 0$ and $k \geq 3$, if $p \geq n^{-\frac{1}{k-1}+\gamma}$ then*

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

Note that the bound on p in Theorem 2.4 is significantly higher than the corresponding bound in the 1-statement of Theorem 2.2. Further, Theorem 2.4 is tight up to the γ term. Indeed, it is not hard to see that if $p = o(n^{-\frac{1}{k-1}} \log n)$ then

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{P}_{n,p} \text{ is } (\delta, k)\text{-Szemerédi}] = 0.$$

This follows as for such p (suitably bounded away from 0), the expected number of k -APs is much less than the expected number of elements in $\mathbf{P}_{n,p}$. Hence, with high probability, greedily deleting a number from each k -AP produces a linear size subset of $\mathbf{P}_{n,p}$ containing no k -AP. In Section 12.2 we give a significant strengthening of Theorem 2.4; see Theorem 12.4.

2.2. Monochromatic arithmetic progressions in the primes. The following seminal result is one of the first proven in Ramsey theory.

Theorem 2.5 (van der Waerden [61]). *Any finite colouring of \mathbb{N} yields a monochromatic arithmetic progression of arbitrary length.*

One beautiful consequence of the Green–Tao theorem is the following generalisation of van der Waerden’s theorem.⁴

Theorem 2.6 (van der Waerden theorem for the primes [30]). *Any finite colouring of \mathbf{P} yields a monochromatic arithmetic progression of arbitrary length.*

Given $k \geq 3$ and $r \geq 2$, we say a subset $X \subseteq \mathbb{N}$ is (r, k) -van der Waerden if whenever X is r -coloured there is a monochromatic k -AP. One of the main applications of our machinery is the following random analogue of Theorem 2.6.

Theorem 2.7 (Random van der Waerden theorem for the primes). *Let $k \geq 3$ and $r \geq 2$. There are constants $c, C > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{P}_{n,p} \text{ is } (r, k)\text{-van der Waerden}] = \begin{cases} 0 & \text{if } p \leq cn^{-\frac{1}{k-1}} \log n; \\ 1 & \text{if } p \geq Cn^{-\frac{1}{k-1}} \log n. \end{cases}$$

Theorem 2.7 tells us that typical subsets of \mathbf{P}_n of size significantly above $|\mathbf{P}_n|n^{-\frac{1}{k-1}} \log n \sim n^{\frac{k-2}{k-1}}$ are (r, k) -van der Waerden whilst typical subsets of \mathbf{P}_n of size significantly smaller than $n^{\frac{k-2}{k-1}}$ are not.

Theorem 2.7 follows from the random Rado lemma combined with Theorem 2.1 and Lemma 2.3. The proof is postponed to Section 11.

3. PARTITION REGULARITY AND SUPERSATURATION

In this section, we give a brief survey of Ramsey-type results for homogeneous systems of linear equations. Partition regularity is covered in the first subsection while supersaturation, including a new result for integer lattices, is covered in the second.

⁴In fact, the Green–Tao theorem implies a supersaturation version of Theorem 2.6, see Theorem 3.6.

3.1. Partition regularity. An integer matrix A is called *partition regular* if every finite colouring of \mathbb{N} yields a monochromatic solution to $Ax = 0$. A classical result of Rado [45] characterises all those matrices that are partition regular. In order to state Rado's theorem, and other results in this section, we require the following definition.

Definition 3.1 (The columns condition over R). Let R be a commutative ring and A be a matrix with entries in R . We say A satisfies the *columns condition over R* if there exists a partition C_0, C_1, \dots, C_m of the columns of A such that the following holds. Let c_j denote the j th column of A and set

$$s_i := \sum_{c_j \in C_i} c_j$$

for every $0 \leq i \leq m$. Then

- $s_0 = 0$ and
- s_i can be written as a linear combination of the elements $\{c_j : c_j \in C_k, k < i\}$ over R , for every $1 \leq i \leq m$.

Theorem 3.2 (Rado [45]). *An integer matrix A is partition regular if and only if A satisfies the columns condition over \mathbb{Q} .*

There have been several works investigating the analogous problem for other algebraic structures. In general, given a ring R acting on an abelian group G , a subset $X \subseteq G$ and $r \in \mathbb{N}$, we say an $\ell \times k$ matrix A with entries in R is *r -partition regular in X* if every r -colouring of X yields a monochromatic solution to $Ax = 0$, where 0 denotes the identity of G^ℓ . Furthermore, A is *partition regular in X* if it is r -partition regular in X for every $r \in \mathbb{N}$. Bergelson, Deuber, Hindman and Lefmann [7] gave some sufficient conditions for partition regularity in the case where $G = R$ and R is a commutative ring that acts on itself via ring multiplication. In a recent paper, Byszewski and Krawczyk [11] proved several interesting extensions of Rado's theorem for integral domains, noetherian rings and modules.

In this paper, we are interested in monochromatic solutions within abelian groups. We mainly focus on matrices with integer entries since there is an obvious action from \mathbb{Z} to abelian groups: given an arbitrary abelian group $(G, +)$, for $g \in G$ and $n \in \mathbb{N}$ we write ng to denote $g + \dots + g$. If n is a negative integer, we write ng to denote the inverse of $|n|g$ in G . For the rest of the paper, \mathbb{Z} will act on abelian groups as we just described.

Given any abelian group G , Deuber [19] provided a necessary and sufficient condition for an integer matrix A to be partition regular in $G \setminus \{0\}$. For *finite* G though it is not interesting as it simply states that A is partition regular in $G \setminus \{0\}$ if and only if there exists $x \in G \setminus \{0\}$ such that $A(x, \dots, x)^T = 0$, i.e., there is a *trivial* solution. Indeed, if the number of colours r is larger than $|G \setminus \{0\}|$ then colouring every element of $G \setminus \{0\}$ differently prevents any monochromatic solution other than trivial ones.

Therefore, in order to ensure a non-trivial monochromatic solution, one must insist that the size of the finite arithmetic structure is large compared to the number of colours. In this direction, Bergelson, Deuber and Hindman [6] proved the following analogue of Rado's theorem for finite vector spaces.

Theorem 3.3 (Bergelson, Deuber and Hindman [6]). *Let \mathbb{F} be a finite field and A be an $\ell \times k$ matrix with entries from \mathbb{F} . The following statements are equivalent:*

- for all $r \in \mathbb{N}$, there exists $n_0 = n_0(r, |\mathbb{F}|, k) \in \mathbb{N}$ such that for all $n \geq n_0$ every r -colouring of $\mathbb{F}^n \setminus \{0\}$ yields a monochromatic solution to $Ax = 0$ in $\mathbb{F}^n \setminus \{0\}$;
- the matrix A is partition regular in $\mathbb{F}^{\mathbb{N}} \setminus \{0\}$;
- the matrix A satisfies the columns condition over \mathbb{F} .

Note that Rado’s theorem itself may be restated in a style similar to Theorem 3.3 by using the following well-known fact.

Fact 3.4. *For an integer matrix A , the following statements are equivalent:*

- (A) *A is partition regular;*
- (B) *for every $r \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ every r -colouring of $[n] := \{1, 2, \dots, n\}$ yields a monochromatic solution to $Ax = 0$ in $[n]$. \square*

Statement (B) trivially implies statement (A) while the converse can be proved using a standard compactness argument.

3.2. Supersaturation. As mentioned in the introduction, the topic of Ramsey multiplicity concerns how many copies of a given structure one can guarantee within a partition of a mathematical object. We are interested in a particular class of results in this area which we refer to as *supersaturation* results. In the context of this paper, a supersaturation result essentially states that within any finite colouring of an arithmetic structure, a *linear proportion* of the total number of solutions of a system of linear equations are monochromatic. As a clarifying example, consider the following theorem of Frankl, Graham and Rödl [22].

Theorem 3.5 (Frankl, Graham and Rödl [22]). *Let $r \in \mathbb{N}$ and let A be a partition regular $\ell \times k$ integer matrix of rank ℓ . There exist $\delta, n_0 > 0$ such that the following holds: for all $n \geq n_0$, every r -colouring of $[n]$ yields at least $\delta n^{k-\ell}$ monochromatic solutions to $Ax = 0$ in $[n]$.*

It is easy to see that there are at most $n^{k-\ell}$ solutions to $Ax = 0$ in $[n]$ (see, e.g., case (ii) of Lemma 6.16). Thus, Theorem 3.5 states that in any r -colouring of $[n]$, a linear proportion of the total number of solutions are monochromatic.

Our motivation for considering supersaturation results is that they are a prerequisite for understanding the typical Ramsey properties of subsets of a given size within an arithmetic structure. Specifically, supersaturation results can be used in combination with the hypergraph container method to give an upper bound on the probability that a Ramsey property is not satisfied, thereby yielding a proof for a 1-statement. See Section 8.1.1 for a more detailed sketch of how this calculation works. Note that supersaturation results were employed to prove 1-statements of Ramsey-type results even before the introduction of the hypergraph container method. In particular, in [26], Theorem 3.5 was used in the proof of the 1-statement of the random Rado theorem (Theorem 4.3 in the next section).

As we will see in Section 6, supersaturation is one of the key conditions to check in the random Rado lemma (Lemma 6.8). In some cases, supersaturation can be easily obtained from existing results. For example, Theorem 2.1 immediately implies the following supersaturation result for the primes.

Theorem 3.6. *Given any $r, k \in \mathbb{N}$ with $k \geq 3$, any r -colouring of \mathbf{P}_n yields $\Theta(n^2/\log^k n)$ monochromatic k -APs in \mathbf{P}_n . \square*

One can similarly obtain a supersaturation result for finite abelian groups and translation-invariant matrices. We say an $\ell \times k$ integer matrix A is *translation-invariant* with respect to a group G if $A(g, \dots, g)^T = 0$ for every $g \in G$.⁵ Consider such a translation-invariant matrix A with respect to an abelian group G . As observed by Saad and Wolf [52, page 2], a generalised version of Szemerédi’s theorem for finite abelian groups [29] yields that for any linear sized subset S of G , provided $|G|$ is sufficiently large, a constant proportion of all solutions to $Ax = 0$ in G are fully contained in S . This immediately implies that, given any $r \in \mathbb{N}$, any r -colouring of G is

⁵For example, any integer matrix A whose columns sum to 0 is translation-invariant with respect to any abelian group. This includes integer matrices for which the solutions to $Ax = 0$ in \mathbb{Z} correspond to k -APs, for fixed $k \geq 3$.

such that a linear proportion of all solutions to $Ax = 0$ in G are monochromatic, provided $|G|$ is sufficiently large.

We finish this section by highlighting three further supersaturation results that we require for our applications of the random Rado lemma appearing in Section 5. First, Serra and Vena [57, Theorem 2.1] proved an analogue of Theorem 3.5 for finite vector spaces. This can be seen as a quantitative analogue of Theorem 3.3.

Theorem 3.7 (Serra and Vena [57]). *Let $r \in \mathbb{N}$, let \mathbb{F} be a finite field and let A be an $\ell \times k$ matrix with entries in \mathbb{F} and rank ℓ which satisfies the columns condition over \mathbb{F} . There exist $\delta, n_0 > 0$ such that, for all $n \geq n_0$, every r -colouring of $\mathbb{F}^n \setminus \{0\}$ yields at least $\delta |\mathbb{F}|^{n(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $\mathbb{F}^n \setminus \{0\}$.*

One can show that there are at most $|\mathbb{F}|^{n(k-\ell)}$ solutions to $Ax = 0$ in \mathbb{F}^n and so the lower bound in Theorem 3.7 is linear in the total number of solutions.

The works of Serra and Vena [57, Theorem 1.3] and Vena [62] extend Theorem 3.7 to finite abelian groups. To state (a special case of) their result, we need a couple of definitions. We say an abelian group $(G, +)$ has *exponent* s if s is the smallest positive integer such that $sg = g + \dots + g = 0$ for every $g \in G$. We say an integer matrix A satisfies the *s -columns condition* if A satisfies the columns condition over the ring \mathbb{Z}_s (where the entries of A are treated as elements of \mathbb{Z}_s , i.e., residues modulo s). Finally, $S(A, G)$ denotes the set of solutions to $Ax = 0$ in G .

Theorem 3.8 (Serra and Vena [57], Vena [62]). *Let $r \in \mathbb{N}$, let G be a finite abelian group with exponent s and let A be an $\ell \times k$ integer matrix which satisfies the s -columns condition. There exist $\delta, n_0 > 0$ such that, for all $n \geq n_0$, every r -colouring of $G^n \setminus \{0\}$ yields at least $\delta |S(A, G^n)|$ monochromatic solutions to $Ax = 0$ in $G^n \setminus \{0\}$.*

Note that in [57], a slightly weaker quantitative bound was obtained for Theorem 3.8. However, the result used a previous weaker removal lemma of Král', Serra and Vena [39]. Vena [62] observed that by using his own removal lemma, one can obtain the stronger bound as stated in Theorem 3.8.

We conclude this section with a new supersaturation result for integer lattices.

Theorem 3.9. *Let $r, d \in \mathbb{N}$ and let A be a partition regular $\ell \times k$ integer matrix of rank ℓ . There exist $\delta, n_0 > 0$ such that the following holds: for all $n \geq n_0$, every r -colouring of $[n]^d$ yields at least $\delta n^{d(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^d$.*

Again, the lower bound in Theorem 3.9 is a linear proportion of the total number of solutions. Observe that the case $d = 1$ of Theorem 3.9 is exactly Theorem 3.5. The proof of Theorem 3.9 is given in Section 10.

4. THE RANDOM RAMSEY PROBLEM FOR GRAPHS AND THE INTEGERS

In this section we cover two settings where the random Ramsey problem has been extensively studied: edge-coloured graphs and the integers. Further background on these two problems can be found, e.g., in [14] and [32, Section 1]. We also discuss a few known random Ramsey-type results for \mathbb{Z}_n .

4.1. Ramsey properties of random graphs. The *random graph* model provides a framework for studying the typical Ramsey properties of graphs of a given density. The *random graph* $G_{n,p}$ has vertex set $[n]$ where each possible edge is present with probability p , independently of all other edges. An event occurs in $G_{n,p}$ *with high probability* (w.h.p.) if its probability tends to 1 as $n \rightarrow \infty$. For many properties \mathcal{P} of $G_{n,p}$ (including Ramsey properties), the probability that $G_{n,p}$ has the property exhibits a phase transition, that is, there is a *threshold* for \mathcal{P} : a function $\hat{p} := \hat{p}(n)$ such that $G_{n,p}$ has \mathcal{P} w.h.p. when $p = \omega(\hat{p})$ (the *1-statement*), while $G_{n,p}$ does not have \mathcal{P} w.h.p. when

$p = o(\hat{p})$ (the 0-statement). Bollobás and Thomason [8] proved that every *monotone* property \mathcal{P} has a threshold.

Rödl and Ruciński [47, 48, 49] proved a general Ramsey-type result for the random graph. To state their result we need a few definitions. Recall that we write $V(H)$ and $E(H)$ for the vertex and edge sets of a graph H , respectively, and define $v(H) := |V(H)|$ and $e(H) := |E(H)|$. Given an $r \in \mathbb{N}$ and a graph G , an r -edge-colouring of G is a function $\sigma : E(G) \rightarrow [r]$. Given a graph H , we say that G is (H, r) -Ramsey if every r -edge-colouring of G yields a monochromatic copy of H in G . 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 $m_2(H) := \max_{H' \subset H} d_2(H')$ to be the 2-density of H . For example, if we consider the triangle K_3 , then $m_2(K_3) = (e(K_3) - 1)/(v(K_3) - 2) = 2$. On the other hand, if we consider the disjoint union of K_3 and an edge, i.e., $H := K_3 \cup K_2$, then we also have that $m_2(H) = 2$, even though $d_2(H) = 1$, as the ‘densest’ part of H is the copy of K_3 .

Theorem 4.1 (The random Ramsey theorem [47, 48, 49]). *Let $r \geq 2$ be a positive integer and let H be a graph that is not a forest consisting of stars and paths of length 3. There are positive constants c, C such that*

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

Thus $n^{-1/m_2(H)}$ is the threshold for the (H, r) -Ramsey property. Roughly speaking, the random Ramsey theorem tells us that a typical n -vertex graph of edge density significantly greater than $n^{-1/m_2(H)}$ is (H, r) -Ramsey, whilst a typical n -vertex graph of edge density significantly less than $n^{-1/m_2(H)}$ is not (H, r) -Ramsey.

An intuition for the threshold in Theorem 4.1 is as follows: for small $c > 0$, if $p < cn^{-1/m_2(H)}$ then there is a subgraph H' of H for which the expected number of copies of H' in $G_{n,p}$ is much smaller than the expected number of edges in $G_{n,p}$. So intuitively this suggests the copies of H' (and therefore H) are ‘spread out’ in the random graph and so one may be able to colour the edges to avoid a monochromatic copy of H . On the other hand, for large $C > 0$, if $p > Cn^{-1/m_2(H)}$ then the expected number of copies of any subgraph H'' of H in $G_{n,p}$ is much larger than the expected number of edges; so in this sense, the copies of H in $G_{n,p}$ are more ‘densely distributed’ making it harder to avoid a monochromatic copy of H .

In more recent work, Nenadov and Steger [44] gave a short proof of Theorem 4.1 using the hypergraph container method. There has also been significant effort to generalise Theorem 4.1 to the setting of *random hypergraphs* (see, e.g., [15, 26, 31]) and also to asymmetric Ramsey properties; that is, now the monochromatic graph one seeks can be different for each colour class (see, e.g., [9, 13, 32, 35, 38, 40, 41, 42, 43]).

4.2. Ramsey properties of random sets of integers. Suppose that A is an $\ell \times k$ integer matrix, and let S be a set of elements of an abelian group (e.g., a set of integers). If a vector $x = (x_1, \dots, x_k) \in S^k$ satisfies $Ax = 0$ and the x_i s are pairwise distinct, we call x a *k-distinct solution* to $Ax = 0$ in S . We say that S is (A, r) -Rado if given any r -colouring of S , there is a monochromatic k -distinct solution $x = (x_1, \dots, x_k)$ to $Ax = 0$ in S . Note that in the study of random versions of Rado’s theorem authors have (implicitly) considered the (A, r) -Rado property, rather than seeking a monochromatic solution that is not necessarily k -distinct (as in the original theorem of Rado); see [33, Section 4] for a discussion on why this has been the case.

An $\ell \times k$ integer matrix A is *irredundant* if there exists a k -distinct solution to $Ax = 0$ 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 for which $Ax = 0$ has solutions in \mathbb{N} , there exists an irredundant $\ell' \times k'$ matrix A' for some $\ell' \leq \ell$ and $k' \leq k$ with the same family of solutions (viewed as sets). See [50, Section 1] for a full explanation.

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

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

The definition of $m(A)$ was introduced in [50]. As noted there, the denominator of $m(A)$ is strictly positive provided that A is irredundant and partition regular, and so $m(A)$ is well-defined in this case. We say A is *strictly balanced* if the expression in (1) is maximised precisely when $W = [k]$. If A has precisely one row consisting of non-zero entries (i.e., it corresponds to a single linear equation) then A is strictly balanced and so

$$m(A) = \frac{k - 1}{k - 2}.$$

For example, if $Ax = 0$ corresponds to the equation $x + y = z$, then $m(A) = 2$.

We write $[n]_p$ for the random subset of $[n] = \{1, \dots, n\}$ where each element in $[n]$ is included with probability p independently of all other elements. Rödl and Ruciński [50] showed that $m(A)$ is an important parameter for determining whether $[n]_p$ is (A, r) -Rado or not. Indeed, one can view $m(A)$ as the arithmetic analogue of the 2-density $m_2(H)$ which we considered for the graph case in the previous subsection.

Theorem 4.2 (Rödl and Ruciński [50]). *For all irredundant partition regular full rank matrices A and all positive integers $r \geq 2$, there exists a constant $c > 0$ such that*

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

Roughly speaking, Theorem 4.2 implies that a typical subset of $[n]$ with significantly fewer than $n^{1-1/m(A)}$ elements is not (A, r) -Rado for any irredundant partition regular matrix A . The following theorem of Friedgut, Rödl and Schacht [26] complements this result, implying that a typical subset of $[n]$ with significantly more than $n^{1-1/m(A)}$ elements is (A, r) -Rado.

Theorem 4.3 (Friedgut, Rödl and Schacht [26]). *For all irredundant partition regular full rank matrices A and all positive integers r , there exists a constant $C > 0$ such that*

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

So together, Theorems 4.2 and 4.3 form the *random Rado theorem* and show that the threshold for the property of being (A, r) -Rado is $p = n^{-1/m(A)}$. The intuition for this threshold is similar to that for the threshold in Theorem 4.1: in the case when A is strictly balanced, when $p = n^{-1/m(A)}$, in expectation there are roughly the same number of elements in $[n]_p$ as there are solutions to $Ax = 0$ in $[n]_p$. So if $p \leq cn^{-1/m(A)}$ for some small $c > 0$, then one would expect the solutions to $Ax = 0$ to be ‘spread out’ (i.e., one would expect most elements of $[n]_p$ lie in very few solutions to $Ax = 0$). Meanwhile, if $p \geq Cn^{-1/m(A)}$ for some large $C > 0$, then one would expect some clustering of solutions to $Ax = 0$, making it harder to avoid monochromatic solutions.

Note that Theorem 4.3 was confirmed earlier by Graham, Rödl and Ruciński [28] in the case where $r = 2$ and $Ax = 0$ corresponds to $x + y = z$, and then by Rödl and Ruciński [50] in the case when A is so-called density regular. We remark that both Theorem 4.2 and Theorem 4.3 are straightforward corollaries of the random Rado lemma (Lemma 6.8), which in turns follows from our general 1- and 0-statement for hypergraphs (Theorem 7.7). Furthermore, the 1-statement of the random Ramsey theorem (Theorem 4.1) also follows directly from the general 1-statement for hypergraphs (Theorem 7.7).

In [32, 58], Theorem 4.3 was generalised to allow one to consider matrices A that are not partition regular, but for which \mathbb{N} is (A, r) -Rado for some choices of r . In particular, the condition that A is partition regular can be replaced with one of the following two equivalent definitions, subject to a suitable Ramsey-type supersaturation result holding. We say that an integer matrix A *satisfies property* $(*)$ if under Gaussian elimination A does not have any row which consists of precisely two non-zero rational entries. The definition of *abundant*, introduced in [54] and used in [58] is that, for an $\ell \times k$ matrix A , every $\ell \times (k - 2)$ submatrix of A has the same rank as A . Clearly irredundant full rank matrices which satisfy $(*)$ are equivalent to irredundant full rank abundant matrices.

4.3. Ramsey properties of random sets of \mathbb{Z}_n . Recently there have been a few random Ramsey-type results proven for \mathbb{Z}_n . The first explicit result in this area is the following *sharp threshold* version of van der Waerden’s theorem for random subsets of \mathbb{Z}_n [24]. Recalling the definition of (r, k) -van der Waerden from earlier, note that, for example, $(r, 3)$ -van der Waerden and (A, r) -Rado are equivalent notions if $A = (1, 1, -2)$. We write $\mathbb{Z}_{n,p}$ to denote the subset of \mathbb{Z}_n obtained by including each element of \mathbb{Z}_n independently and with probability p .

Theorem 4.4 (Friedgut, Hàn, Person and Schacht [24]). *For all $k \geq 3$, there exist constants $c_1, c_0 > 0$ and a function $c(n)$ satisfying $c_0 \leq c(n) \leq c_1$ such that for every $\varepsilon > 0$ we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{Z}_{n,p} \text{ is } (2, k)\text{-van der Waerden}] = \begin{cases} 0 & \text{if } p \leq (1 - \varepsilon)c(n)n^{-1/(k-1)}; \\ 1 & \text{if } p \geq (1 + \varepsilon)c(n)n^{-1/(k-1)}. \end{cases}$$

Note that the threshold in Theorem 4.4 is sharp compared to, e.g., the threshold given by Theorems 4.2 and 4.3. See, e.g., [14, Section 5.1] for further discussion on sharp thresholds.

Recently, Theorem 4.4 has been generalised to the (r, k) -van der Waerden property (i.e., r does not have to be 2 now) and an analogous sharp threshold version of the random Schur theorem for \mathbb{Z}_n (where n is prime) has been obtained; see [25].

5. FURTHER APPLICATIONS OF THE RANDOM RADO LEMMA

In addition to the random van der Waerden theorem for the primes (Theorem 2.7), we prove several other novel random Ramsey-type results via the random Rado lemma (Lemma 6.8). The proofs of all the results presented in this section are postponed to Section 11. Given a subset X of an abelian group, we write X_p to denote the subset of X obtained by including each element of X independently and with probability p .

Recall that Theorems 4.2 and 4.3 determine the threshold for the property that $[n]_p$ is (A, r) -Rado. The following result is a generalisation of this statement to $[n]_p^d$ for any fixed $d \in \mathbb{N}$.

Theorem 5.1 (Random Rado theorem for integer lattices). *For all irredundant partition regular full rank matrices A and all positive integers $r \geq 2$ and $d \geq 1$, there exist constants $C, c > 0$ such that the following holds.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[[n]_p^d \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq cn^{-d/m(A)}; \\ 1 & \text{if } p \geq Cn^{-d/m(A)}. \end{cases}$$

Observe that the $d = 1$ case corresponds precisely to Theorems 4.2 and 4.3. To deduce Theorem 5.1 from the random Rado lemma, we require a corresponding supersaturation result for $[n]^d$, namely Theorem 3.9.

Note that the probability threshold in Theorem 5.1 is independent of d in the following sense. As we are interested in how ‘sparse’ a subset can be while typically retaining its Ramsey properties, the probability threshold should always be compared to the size of the initial set we are considering. In this case, the size of $[n]^d$ is n^d and so we may rewrite the probability threshold $n^{-d/m(A)}$ as

$|[n]^d|^{-1/m(A)}$. Indeed, this shows that the probability threshold depends exclusively on the matrix A . However, this phenomenon does not occur in general: for a fixed matrix A , the probability threshold may change depending on the underlying arithmetic structure. For instance, consider the equation $2x + 2y = 2z$ and the product group $(\mathbb{Z}_4)^n$.

Theorem 5.2. *Let $A = \begin{pmatrix} 2 & 2 & -2 \end{pmatrix}$ and $r \geq 2$. There exist constants $c, C > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[(\mathbb{Z}_4)_p^n \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c 4^{-3n/4}; \\ 1 & \text{if } p \geq C 4^{-3n/4}. \end{cases}$$

As the size of $(\mathbb{Z}_4)^n$ is 4^n , we may rewrite $4^{-3n/4}$ as $|(\mathbb{Z}_4)^n|^{-3/4}$. On the other hand, $m(A) = 2$ and so the corresponding threshold when considering solutions to $Ax = 0$ in $[n]_p$ is $n^{-1/2}$. Thus, we can have different thresholds depending on the underlying abelian group.

Note that the probability threshold $p = 4^{-3n/4}$ is consistent with the heuristic that the expected size of $(\mathbb{Z}_4)_p^n$ should ‘match’ the expected number of solutions to $Ax = 0$ in $(\mathbb{Z}_4)_p^n$. Indeed, for $p = 4^{-3n/4}$, the expected size of $(\mathbb{Z}_4)_p^n$ is $4^n p = 4^{n/4}$. The total number of solutions of $2x + 2y = 2z$ in $(\mathbb{Z}_4)^n$ is $4^n \cdot 4^n \cdot 2^n = 4^{5n/2}$ and so the expected number of solutions in $(\mathbb{Z}_4)_p^n$ is $4^{5n/2} p^3 = 4^{n/4}$.

Observe that the total number of solutions to $2x + 2y = 2z$ in $[n]$ is at most $n \cdot n \cdot 1 = n^2$ since for every choice of x and y there is at most one choice of z . The change of threshold in Theorem 5.2 (and more generally in Theorem 5.3 below) depends on the fact that we have many more choices for z in $(\mathbb{Z}_4)^n$.

Theorem 5.2 follows easily from the following application of the random Rado lemma to abelian groups. The definitions of $\text{rank}_G(A)$ and $m_G(A)$ appearing in the statement require further motivation so we defer them to the next section; see Definitions 6.11 and 6.27. However, one may think of $m_G(A)$ as the natural generalisation of $m(A)$ which captures the intuition described for Theorem 5.2. Indeed, $m_{\mathbb{Z}_4}(A) = 4/3$ for $A = \begin{pmatrix} 2 & 2 & -2 \end{pmatrix}$. For an $\ell \times k$ matrix A , we say that (A, G) is *abundant* if every $\ell \times (k - 2)$ submatrix A' of A satisfies $\text{rank}_G(A') = \text{rank}_G(A)$. This naturally generalises the previous definition of abundant; see also Definition 6.22.

Theorem 5.3. *Let G be a finite abelian group with exponent $s \in \mathbb{N}$. Let A be an integer matrix which satisfies the s -columns condition, with $\text{rank}_G(A) > 0$ and such that (A, G) is abundant. For all $r \geq 2$, there exist constants $C, c > 0$ such that the following holds.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_p^n \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c|G|^{-n/m_G(A)}; \\ 1 & \text{if } p \geq C|G|^{-n/m_G(A)}. \end{cases}$$

Note that the condition $\text{rank}_G(A) > 0$ in Theorem 5.3 is very mild: if A has dimension $\ell \times k$, $\text{rank}_G(A) > 0$ is equivalent to stating that not every ordered k -tuple over G is a solution to $Ax = 0$.

Given a finite subset S of an abelian group, we say that an $\ell \times k$ matrix A is *irredundant with respect to S* if there exists a k -distinct solution to $Ax = 0$ in S . So a matrix A is irredundant precisely if it is irredundant with respect to \mathbb{N} . Another application of the random Rado lemma is the following random version of the Ramsey-type result of Bergelson, Deuber and Hindman for vector spaces \mathbb{F}^n (Theorem 3.3). For an $\ell \times k$ matrix A with entries from a field \mathbb{F} , we define $m_{\mathbb{F}}(A)$ analogously to (1), that is

$$(2) \quad m_{\mathbb{F}}(A) := \max_{\substack{W \cup \overline{W} = [k] \\ |W| \geq 2}} \frac{|W| - 1}{|W| - 1 + \text{rank}_{\mathbb{F}}(A_{\overline{W}}) - \text{rank}_{\mathbb{F}}(A)},$$

where we write $\text{rank}_{\mathbb{F}}$ to denote that we calculate rank with respect to the field \mathbb{F} .

Theorem 5.4. *Let \mathbb{F} be a non-trivial finite field. Consider any full rank matrix A with entries from \mathbb{F} so that A satisfies the columns condition over \mathbb{F} and A is also irredundant with respect to*

\mathbb{F}^n for all $n \in \mathbb{N}$ sufficiently large. Given any positive integer $r \geq 2$, there exist constants $C, c > 0$ such that the following holds.

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{F}_p^n \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c|\mathbb{F}|^{-n/m_{\mathbb{F}}(A)}; \\ 1 & \text{if } p \geq C|\mathbb{F}|^{-n/m_{\mathbb{F}}(A)}. \end{cases}$$

We also obtain the analogue of Theorem 5.4 for matrices A which have integer entries. Here we must be slightly careful: as we are interested in solutions to $Ax = 0$ in \mathbb{F}^n , one must compute the rank of the integer matrix A with respect to the underlying field \mathbb{F} (see Definition 6.14) and hence the threshold depends on $m_{\mathbb{F}}(A)$ rather than $m(A)$.⁶ For example, the matrix $A = \begin{pmatrix} 3 & 3 & -3 \end{pmatrix}$ usually has $\text{rank}(A) = 1$, however, in \mathbb{Z}_3 , we have $3 \equiv 0 \pmod{3}$ and so we have $\text{rank}_{\mathbb{Z}_3}(A) = 0$.

Theorem 5.5. *Let \mathbb{F} be a finite field of order q^k , for some prime q and $k \in \mathbb{N}$. Consider any full rank integer matrix A so that A satisfies the q -columns condition and A is also irredundant with respect to \mathbb{F}^n for all $n \in \mathbb{N}$ sufficiently large. Given any positive integer $r \geq 2$, there exist constants $C, c > 0$ such that the following holds.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{F}_p^n \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c|\mathbb{F}|^{-n/m_{\mathbb{F}}(A)}; \\ 1 & \text{if } p \geq C|\mathbb{F}|^{-n/m_{\mathbb{F}}(A)}. \end{cases}$$

Theorem 5.5 is just a simple application of Theorem 5.4; its proof is given in Section 11.

6. THE RANDOM RADO LEMMA

In this section, we rigorously state the random Rado lemma (Lemma 6.8). As previously mentioned, the random Rado lemma serves as a general black box that outputs a random Ramsey-type result for a given sequence of finite subsets of abelian groups provided that certain conditions, most notably supersaturation, are satisfied.

The content of this section is divided into four subsections. The first subsection contains the statement of Lemma 6.8 preceded by a list of all the required definitions. In the second subsection, we prove a general bound for the number of solutions to $Ax = 0$ in S when S is either a finite abelian group or a (power of a) finite subset of a field. This bound is used in the third subsection to prove that several technical requirements of the random Rado lemma trivially hold, or can be relaxed, in these specific settings. In light of these results, in the fourth subsection we formulate a simplified version of the random Rado lemma (Lemma 6.33) for finite abelian groups and (powers of) finite subsets of fields.

We stress once more that the random Rado lemma can be straightforwardly deduced from our general 1-statement and 0-statement for hypergraphs (Theorem 7.7). The proof of Lemma 6.8 is deferred to Section 9.

6.1. Definitions and main statement. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups. We write $S_{n,p}$ to denote the subset of S_n obtained by including each element of S_n independently and with probability p . We say an event occurs in $S_{n,p}$ *with high probability* (w.h.p.) if its probability tends to 1 as $n \rightarrow \infty$. Let A be an $\ell \times k$ integer matrix. We are interested in the probability threshold $\hat{p} := \hat{p}(n)$ such that $S_{n,p}$ is (A, r) -Rado w.h.p. if $p = \omega(\hat{p})$ and $S_{n,p}$ is not (A, r) -Rado w.h.p. if $p = o(\hat{p})$.

Of course, as each subset S_n can lie in a different abelian group, it is easy to come up with a sequence $(S_n)_{n \in \mathbb{N}}$ for which no such probability threshold \hat{p} exists. On the other hand, in Definition 6.2 we provide a candidate for \hat{p} for ‘well-behaved’ sequences $(S_n)_{n \in \mathbb{N}}$. Given a k -tuple $(x_i)_{i \in [k]}$ and $W \subseteq [k]$, we write x_W for the $|W|$ -tuple $(x_i)_{i \in W}$.

⁶Note that the definition of $m_{\mathbb{F}}(A)$ for an integer matrix A is also consistent with putting $G = \mathbb{F}$ into the definition of $m_G(A)$ for groups (see Definition 6.27 and Fact 6.15).

Definition 6.1 (Projected solutions). Let S be a finite subset of an abelian group and A be an $\ell \times k$ integer matrix. Let $W \subseteq Y \subseteq [k]$ and $w_0 \in S^{|W|}$. The set of *projected solutions* $\text{Sol}_S^A(w_0, W, Y)$ is the set of vectors $y \in S^{|Y|}$ such that there exists a solution $x \in S^k$ to $Ax = 0$ with $x_Y = y$ and $x_W = w_0$. For brevity, we write $\text{Sol}_S^A(Y) := \text{Sol}_S^A(\emptyset, \emptyset, Y)$, i.e., $\text{Sol}_S^A(Y)$ is the set of vectors $y \in S^{|Y|}$ such that there exists a solution $x \in S^k$ to $Ax = 0$ with $x_Y = y$. In particular, $\text{Sol}_S^A([k])$ is the set of all solutions to $Ax = 0$ in S .

Definition 6.2 (Probability threshold for abelian groups). Let S be a finite subset of an abelian group and A be an $\ell \times k$ integer matrix. For every $W \subseteq [k]$, we let

$$p_W(A, S) := \left(\frac{|\text{Sol}_S^A(W)|}{|S|} \right)^{-\frac{1}{|W|-1}}.$$

Additionally, we let

$$\hat{p}(A, S) := \max_{\substack{W \subseteq [k] \\ |W| \geq 2}} p_W(A, S).$$

The intuition for the choice of $\hat{p}(A, S)$ in Definition 6.2 aligns with the heuristic we discussed in Section 4, where we considered the random Ramsey problem for the integers. Indeed, it is easy to check that for $p := p_W(A, S)$ the expected number of projected solutions $|\text{Sol}_S^A(W)|p^{|W|}$ equals the expected size $|S|p$ of the ground set. The probability threshold $\hat{p}(A, S)$ is obtained by taking the largest of these $p_W(A, S)$.

Remark 6.3. *As the definition of (A, r) -Rado involves k -distinct solutions, the reader may wonder why, in Definition 6.1, we do not insist that the solutions to $Ax = 0$ must be k -distinct. This choice is designed to simplify later calculations. We will be able to restrict our attention to k -distinct solutions by imposing additional assumptions over the sequence $(S_n)_{n \in \mathbb{N}}$ (namely, conditions (A3) and (A6) of Lemma 6.8).*

Next, we introduce four properties that the sequence $(S_n)_{n \in \mathbb{N}}$ must satisfy in the hypothesis of the random Rado lemma. The first one is supersaturation, which we introduced and gave motivation for in Section 3.2. We will use the following formal definition.

Definition 6.4 (Supersaturation). Let $r \in \mathbb{N}$, $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups and A be an $\ell \times k$ integer matrix. We say that $(S_n)_{n \in \mathbb{N}}$ is (A, r) -supersaturated if there exist an $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that the following holds: if $n \geq n_0$ then whenever S_n is r -coloured there are at least $\varepsilon |\text{Sol}_{S_n}^A([k])|$ monochromatic solutions to $Ax = 0$ in S_n .

The second property involves an upper bound on the number of projected solutions. Intuitively, this forces the solutions to be uniformly distributed over the ground set. This condition is rather natural since the heuristic argument for the probability threshold relies on the assumption that the solutions to $Ax = 0$ are ‘spread out’.

Definition 6.5 (Extendability). Let S be a finite subset of an abelian group, let A be an $\ell \times k$ integer matrix and let $B \geq 1$ be a real number. We say the pair (A, S) is B -extendable if for all non-empty $W \subseteq Y \subseteq [k]$ and any $w_0 \in S^{|W|}$ we have

$$|\text{Sol}_S^A(w_0, W, Y)| \leq B \cdot \frac{|\text{Sol}_S^A(Y)|}{|\text{Sol}_S^A(W)|}.$$

The third property is more involved and perhaps the most artificial out of all requirements of the random Rado lemma. This property is tied to the calculations appearing in the proof of the general 0-statement for hypergraphs (Theorem 7.7) in Section 8.

Definition 6.6 (Compatibility). Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups and A be an $\ell \times k$ integer matrix with $k \geq 3$. We say $(S_n)_{n \in \mathbb{N}}$ is *compatible* with respect to A if there exists a sequence $(X_n)_{n \in \mathbb{N}}$ of subsets $X_n \subseteq [k]$ with $|X_n| \geq 3$ for all $n \in \mathbb{N}$ such that the following holds. We have $p_{X_n}(A, S_n) = \Omega(\hat{p}(A, S_n))$ and, for all sequences $(W_n)_{n \in \mathbb{N}}$ and $(W'_n)_{n \in \mathbb{N}}$ with $W_n, W'_n \subset X_n$, $|W_n| = 2$ and $|W'_n| \geq 2$,

$$(3) \quad \frac{|S_n|^2}{|\text{Sol}_{S_n}^A(W_n)|} \left(\frac{p_{W'_n}(A, S_n)}{p_{X_n}(A, S_n)} \right)^{|W'_n|-1} \rightarrow 0$$

as $n \rightarrow \infty$.

The fourth property is a mild technical requirement that will be needed in the proof of Lemma 6.8 to allow us to restrict our attention to k -distinct solutions (see Lemma 6.9).

Definition 6.7 (Weak compatibility). Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups and A be an $\ell \times k$ integer matrix with $k \geq 2$. We say $(S_n)_{n \in \mathbb{N}}$ is *weakly compatible* with respect to A if for every $W \subseteq [k]$ with $|W| = 2$ we have

$$\frac{|S_n|}{|\text{Sol}_{S_n}^A(W)|} \rightarrow 0$$

as $n \rightarrow \infty$.

We are now ready to formally state the random Rado lemma.

Lemma 6.8 (The random Rado lemma). *Let $r \geq 2$, $k \geq 3$ and $\ell \geq 1$ be integers. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups, A be an $\ell \times k$ integer matrix and set $\hat{p}_n := \hat{p}(A, S_n)$.*

Suppose that

- (A1) $\hat{p}_n \rightarrow 0$ and $|S_n| \hat{p}_n \rightarrow \infty$ as $n \rightarrow \infty$;
- (A2) $(S_n)_{n \in \mathbb{N}}$ is (A, r) -supersaturated;
- (A3) there exists $B \geq 1$ such that the pair (A, S_n) is B -extendable for all sufficiently large $n \in \mathbb{N}$;
- (A4) there exists $D > 0$ such that for every $W \subset Y \subseteq [k]$ with $|W| = 1$ and $w_0 \in S_n$, we have $|\text{Sol}_{S_n}^A(w_0, W, Y)| \leq D \frac{|\text{Sol}_{S_n}^A(Y)|}{|S_n|}$ for all sufficiently large $n \in \mathbb{N}$;
- (A5) $(S_n)_{n \in \mathbb{N}}$ is compatible with respect to A ;
- (A6) $(S_n)_{n \in \mathbb{N}}$ is weakly compatible with respect to A .

Then there exist constants $c, C > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c\hat{p}_n; \\ 1 & \text{if } p \geq C\hat{p}_n. \end{cases}$$

We find the generality of Lemma 6.8 very interesting - it confirms that for ‘well-behaved’ sequences $(S_n)_{n \in \mathbb{N}}$, the probability threshold essentially depends on the number of (projected) solutions to $Ax = 0$ in S_n .

Note that (A4) is a similar condition to the $|W| = 1$ case of the definition of B -extendable; we have written (A3) and (A4) as separate conditions as they correspond to conditions (P3) and (P4) in our general 1-statement and 0-statement for hypergraphs (Theorem 7.7). In fact, the random Rado lemma is a simple corollary of Theorem 7.7 and each of conditions (A1)–(A5) are adaptations of conditions (P1)–(P5) from Theorem 7.7. Condition (A6) is a mild requirement that we impose in order to restrict our attention to k -distinct solutions (see Lemma 6.9 below).

For most natural applications, conditions (A3)–(A6) can be further simplified. Thus, in practice, condition (A2) is the crucial assumption that one needs to verify. The goal of the next three subsections is to state a simplified version of the random Rado lemma (Lemma 6.33) which holds for sequences of finite abelian groups and sequences of (powers of) finite subsets of fields. Nonetheless, there are interesting examples where the generality of the random Rado lemma as stated in

Lemma 6.8 is required. The most notable such example is the random van der Waerden theorem for the primes (Theorem 2.7), proved in Section 11.

It would be interesting to see if one can relax some of the assumptions in the random Rado lemma; if so, condition (A5) is the most likely candidate for where the hypothesis can be weakened. See Section 12.3 for a brief discussion related to this.

The next result states that conditions (A3) and (A6) imply a linear proportion of projected solutions are induced by k -distinct solutions. The proof is postponed to Section 9 but we state the result here in order to give the reader a better intuition why we did not impose in Definition 6.1 that the solutions are k -distinct.

Lemma 6.9. *Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups and let A be an $\ell \times k$ integer matrix with $k \geq 2$. Suppose that conditions (A3) and (A6) from Lemma 6.8 hold; so there is a $B \geq 1$ such that (A, S_n) is B -extendable for every $n \in \mathbb{N}$ sufficiently large.*

For each non-empty $Y \subseteq [k]$, let $k\text{-Sol}_{S_n}^A(Y)$ be the set of $y \in S_n^{|Y|}$ such that there exists a k -distinct solution $x \in S^k$ to $Ax = 0$ with $x_Y = y$. For n sufficiently large, we have

$$\frac{1}{2B} \leq \frac{|k\text{-Sol}_{S_n}^A(Y)|}{|\text{Sol}_{S_n}^A(Y)|} \leq 1.$$

Furthermore, for $Y = [k]$ we have

$$\lim_{n \rightarrow \infty} \frac{|k\text{-Sol}_{S_n}^A([k])|}{|\text{Sol}_{S_n}^A([k])|} = 1.$$

Remark 6.10. *Recall that given a finite subset S of an abelian group, we say that an $\ell \times k$ matrix A is irredundant with respect to S if there exists a k -distinct solution to $Ax = 0$ in S . In many of the results mentioned previously in this paper (e.g., Theorem 5.1) we require that the matrix considered is irredundant. However, we do not explicitly state that A is irredundant with respect to S_n in Lemma 6.8 as conditions (A3) and (A6) combined with Lemma 6.9 already imply that many solutions are k -distinct.*

6.2. Rank and counting solutions. The notion of extendability and (weak) compatibility introduced in the previous subsection are tied to the quantity $|\text{Sol}_S^A(W)|$. In many natural cases we are able to approximate, or at least bound, this number using standard methods from group theory and linear algebra. For example, it is well known that the size of $\text{Sol}_{[n]}^A([k])$, i.e., the number of solutions to $Ax = 0$ in $[n]$, is at most $n^{k-\text{rank}(A)}$ (see, e.g., case (ii) of Lemma 6.16).

The notion of rank is useful here. For instance, the parameter $m(A)$ which governs the probability threshold for the random Rado theorem for the integers is expressed in terms of $\text{rank}(A_W)$ for $W \subseteq [k]$. We now provide an appropriate generalisation of the notion of rank for the cases where S is either a finite abelian group or (a power of) a finite subset of a field.

Definition 6.11 (Rank for finite abelian groups). Let S be a finite abelian group. Let A be an $\ell \times k$ integer matrix and let $f_A : S^k \rightarrow S^\ell$ be the function defined by $f_A : x \mapsto Ax$. We define

$$\text{rank}_S(A) := \begin{cases} \log_{|S|} |\text{im}(f_A)| & \text{if } |S| > 1; \\ 0 & \text{if } |S| = 1. \end{cases}$$

By the first isomorphism theorem for groups, we have $|S|^k = |\text{im}(f_A)| \cdot |\ker(f_A)|$. In particular, the number of solutions to $Ax = b$ is equal to

$$(4) \quad \begin{cases} |\ker(f_A)| = \frac{|S|^k}{|\text{im}(f_A)|} = |S|^{k-\text{rank}_S(A)} & \text{if } b \in \text{im}(f_A); \\ 0 & \text{otherwise.} \end{cases}$$

A motivation for the choice of $\text{rank}_S(A)$ in Definition 6.11 is that it will allow us to rewrite $p_W(A, S)$ from Definition 6.2 as a power of $|S|$. Indeed, in all applications of the random Rado lemma discussed in Sections 2 and 5, the probability threshold is expressed in such form. For example, it follows immediately from equation (4) that if A is an $\ell \times k$ integer matrix and S is a finite abelian group then $|\text{Sol}_S^A([k])| = |S|^{k - \text{rank}_S(A)}$ and so $p_{[k]}(A, S) = |S|^{-(k-1 - \text{rank}_S(A))/(k-1)}$.

Next, we consider the case where $S = L^d$ for some finite subset L of a field \mathbb{F} and $d \in \mathbb{N}$. We are mostly interested in integer matrices, but it will be convenient to first consider matrices with entries in \mathbb{F} . We write $0_{\mathbb{F}}$ and $1_{\mathbb{F}}$ for the additive and multiplicative identities of \mathbb{F} , respectively.

Definition 6.12. Let $S = L^d$ where L is a finite subset of a field \mathbb{F} and $d \in \mathbb{N}$. Let A be a matrix with entries from \mathbb{F} . If $L = \{0_{\mathbb{F}}\}$, let $\text{rank}_S(A) := 0$. Otherwise, let $\text{rank}_S(A)$ be the rank of A with respect to \mathbb{F} .

The following simple fact shows that $\text{rank}_S(A)$ in Definition 6.12 is well-defined: the value of $\text{rank}_S(A)$ is independent of the choice of \mathbb{F} .

Fact 6.13. Let $\mathbb{F} \subseteq \mathbb{F}'$ be two fields. Let A be a matrix whose entries lie in \mathbb{F} . Then, the rank of A with respect to \mathbb{F} is the same as the rank of A with respect to \mathbb{F}' .

Proof. The rank of A is the size of the largest square submatrix of A with non-zero determinant. The determinant of a submatrix is computed by performing multiplications and additions using the entries of the submatrix. Let \mathbb{F}_0 be the field generated by the entries of A . Since the entries of A lie in $\mathbb{F}_0 \subseteq \mathbb{F} \subseteq \mathbb{F}'$ and \mathbb{F}_0 is closed under addition and multiplication, the determinant of any submatrix of A is the same whether we are working in \mathbb{F} or \mathbb{F}' . \square

Definition 6.12 easily extends to matrices with integer entries by treating the entries of such matrices as elements of the field \mathbb{F} .

Definition 6.14 (Rank for powers of finite subsets of fields). Let $S = L^d$ where L is a finite subset of a field \mathbb{F} and $d \in \mathbb{N}$. Let $A = \{a_{ij}\}$ be an $\ell \times k$ matrix with integer entries. If $L = \{0_{\mathbb{F}}\}$, let $\text{rank}_S(A) := 0$. Otherwise, let $\text{rank}_S(A)$ be the rank of B with respect to \mathbb{F} , where $B := \{b_{ij}\}$ is the $\ell \times k$ matrix with $b_{ij} := a_{ij} \cdot 1_{\mathbb{F}} \in \mathbb{F}$.

Note that the entries of the matrix B in Definition 6.14 lie in the subfield $\langle 1_{\mathbb{F}} \rangle \subseteq \mathbb{F}$ generated by $1_{\mathbb{F}}$. Thus, by Fact 6.13, the rank of B is the same with respect to \mathbb{F} and $\langle 1_{\mathbb{F}} \rangle$. If \mathbb{F} and \mathbb{F}' are non-trivial fields with $L \subseteq \mathbb{F}$ and $L \subseteq \mathbb{F}'$, then $1_{\mathbb{F}} = 1_{\mathbb{F}'}$ and in particular $\langle 1_{\mathbb{F}} \rangle = \langle 1_{\mathbb{F}'} \rangle$. Thus $\text{rank}_S(A)$ from Definition 6.14 is independent from different choices of \mathbb{F} and so it is well-defined.

The following fact states that Definition 6.14 is consistent with Definition 6.11.

Fact 6.15. Let $A = \{a_{ij}\}$ be an $\ell \times k$ integer matrix and $S = L^d$ where L is a finite subset of a field \mathbb{F} and $d \in \mathbb{N}$. If S is a finite abelian group, then $\text{rank}_S(A)$ from Definition 6.14 is the same as $\text{rank}_S(A)$ from Definition 6.11.

Proof. As S is a finite abelian group, L is a finite abelian group too. If $L = \{0_{\mathbb{F}}\}$ then $\text{rank}_S(A) = 0$ in both Definitions 6.11 and 6.14; so suppose $L \neq \{0_{\mathbb{F}}\}$.

It is well-known that every element in $(\mathbb{F}, +)$ other than $0_{\mathbb{F}}$ has (additive) order exactly p where either $p = \infty$ or p is a prime number. Since L is a finite abelian group, we must have $L \cong ((\mathbb{Z}_p)^t, +)$ for some prime p and $t \in \mathbb{N}$. Furthermore, we have $\langle 1_{\mathbb{F}} \rangle \cong (\mathbb{Z}_p, +, \times)$.

Let $B := \{b_{ij}\}$ where $b_{ij} := a_{ij} \cdot 1_{\mathbb{F}} \in \langle 1_{\mathbb{F}} \rangle \subseteq \mathbb{F}$ and let r be the rank of B with respect to \mathbb{F} . By Fact 6.13, r equals the rank of B with respect to $\langle 1_{\mathbb{F}} \rangle$. This in turn is equal to the rank of A with respect to \mathbb{Z}_p , where we treat the entries of A modulo p , by the definition of B and since $\langle 1_{\mathbb{F}} \rangle \cong (\mathbb{Z}_p, +, \times)$.

Let $f_A : S^k \rightarrow S^\ell$, $g_A : L^k \rightarrow L^\ell$, $h_A : (\mathbb{Z}_p)^k \rightarrow (\mathbb{Z}_p)^\ell$ be defined by $f_A, g_A, h_A : x \mapsto Ax$. Note that the rank of h_A is precisely r by the previous observations, and so $|\text{im}(h_A)| = |\mathbb{Z}_p|^r = p^r$. It is easy to see that

$$|\text{im}(f_A)| = |\text{im}(g_A)|^d = |\text{im}(h_A)|^{dt}.$$

In particular, $|\text{im}(f_A)| = p^{rdt} = |S|^r$ and so $r = \log_{|S|} |\text{im}(f_A)|$, as required. \square

Next, we prove a general upper bound for the number of projected solutions for finite abelian groups and powers of finite subsets of fields. We will use this tool in Section 6.3 to simplify conditions (A3)–(A6) for these settings.

Lemma 6.16 (Key bounding result). *Let A be an $\ell \times k$ integer matrix. Suppose that either*

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

Then for all $W \subseteq Y \subseteq [k]$ and $w_0 \in S^{|W|}$ we have

$$(5) \quad |\text{Sol}_S^A(w_0, W, Y)| \leq |S|^{|Y| - |W| - \text{rank}_S(A_{\overline{W}}) + \text{rank}_S(A_{\overline{Y}})}.$$

Note that here we define $\text{rank}_S(A_{\overline{W}}) := 0$ if $\overline{W} = \emptyset$.

Proof. Fix $W \subseteq Y \subseteq [k]$ and $w_0 \in S^{|W|}$. If $W = Y$ then $|\text{Sol}_S^A(w_0, W, Y)| = 1$ and (5) holds. So we may assume that $W \subset Y$.

Case (i): S is a finite abelian group. We double count the quantity $|\text{Sol}_S^A(w_0, W, [k])|$. This is equal to the number of solutions x to $A_{\overline{W}}x = -A_W w_0$ which in turn is equal to either 0 or $|S|^{k - |W| - \text{rank}_S(A_{\overline{W}})}$ by (4). In the former case, $|\text{Sol}_S^A(w_0, W, [k])| = 0$ implies $|\text{Sol}_S^A(w_0, W, Y)| = 0$ and so (5) holds. Hence, suppose that

$$(6) \quad |\text{Sol}_S^A(w_0, W, [k])| = |S|^{k - |W| - \text{rank}_S(A_{\overline{W}})}.$$

In particular, we have $|\text{Sol}_S^A(w_0, W, Y)| > 0$. Let $y_0 \in \text{Sol}_S^A(w_0, W, Y)$. By the same argument as above, we have $|\text{Sol}_S^A(y_0, Y, [k])| = |S|^{k - |Y| - \text{rank}_S(A_{\overline{Y}})}$. Thus,

$$|\text{Sol}_S^A(w_0, W, [k])| = \sum_{y_0 \in \text{Sol}_S^A(w_0, W, Y)} |\text{Sol}_S^A(y_0, Y, [k])| = |\text{Sol}_S^A(w_0, W, Y)| \cdot |S|^{k - |Y| - \text{rank}_S(A_{\overline{Y}})}.$$

Combining the above with (6) yields

$$|\text{Sol}_S^A(w_0, W, Y)| = |S|^{|Y| - |W| - \text{rank}_S(A_{\overline{W}}) + \text{rank}_S(A_{\overline{Y}})},$$

as required.

Case (ii): $S = L^d$ for some finite subset L of a field \mathbb{F} and $d \in \mathbb{N}$. If $L = \{0_{\mathbb{F}}\}$ then $|\text{Sol}_S^A(w_0, W, Y)| \leq 1$ and so (5) holds. Hence, assume $L \neq \{0_{\mathbb{F}}\}$. For the rest of this case, we treat the entries of A as elements of \mathbb{F} (namely, we view $n \in \mathbb{N}$ as $n \cdot 1_{\mathbb{F}}$). Let $I_0 \subseteq [k]$ be the indices of a maximal set of linearly independent columns in $A_{\overline{Y}}$ and I_1 be the indices of a maximal set of columns in $A_{Y \setminus W}$ such that the columns with indices $I_0 \cup I_1$ are linearly independent. Then recalling $W \subseteq Y$, we have $\text{rank}_S(A_{\overline{W}}) = |I_0 \cup I_1|$.

Claim 6.17. *Let u, v be two solutions $Au = Av = 0$ in \mathbb{F} . If $u_{Y \setminus I_1} = v_{Y \setminus I_1}$ then $u_{I_1} = v_{I_1}$.*

Proof of the claim. Since $A(u - v) = 0$ and $u_{Y \setminus I_1} = v_{Y \setminus I_1}$

$$A_{I_1}(u - v)_{I_1} + A_{\overline{Y}}(u - v)_{\overline{Y}} = 0,$$

and thus $A_{I_1}(u - v)_{I_1} \in \langle A_{\overline{Y}} \rangle$, where $\langle A_{\overline{Y}} \rangle$ denotes the vector space spanned by the columns of $A_{\overline{Y}}$. By the maximality of I_0 we have $\langle A_{\overline{Y}} \rangle = \langle A_{I_0} \rangle$ and so $A_{I_1}(u - v)_{I_1} \in \langle A_{I_0} \rangle$. Since the columns with indices $I_0 \cup I_1$ are linearly independent, this implies that $(u - v)_{I_1} = 0$.

■

Now we bound $|\text{Sol}_S^A(w_0, W, Y)|$ from above. Recall $w_0 \in S^{|W|}$ and $S = L^d$. We can therefore write w_0 in the form $w_0 = (t_1, \dots, t_d)$ where $t_i \in L^{|W|}$. Note that

$$|\text{Sol}_S^A(w_0, W, Y)| = \prod_{i=1}^d |\text{Sol}_L^A(t_i, W, Y)|,$$

so it suffices to prove that $|\text{Sol}_L^A(t_i, W, Y)| \leq |L|^{|Y|-|W|-\text{rank}_S(A_{\overline{W}})+\text{rank}_S(A_{\overline{Y}})}$ for every i . By Claim 6.17, there is at most one projected solution for every choice of the entries with indices $Y \setminus I_1$. Note that the entries with indices in W are already fixed and W and I_1 are disjoint, thus

$$|\text{Sol}_L^A(t_i, W, Y)| \leq |L|^{|Y|-|W|-|I_1|} = |L|^{|Y|-|W|-|I_0 \cup I_1|+|I_0|} = |L|^{|Y|-|W|-\text{rank}_S(A_{\overline{W}})+\text{rank}_S(A_{\overline{Y}})}.$$

□

Applying inequality (5) from Lemma 6.16 with $W = \emptyset$ and $Y = [k]$ immediately implies that there are at most $|S|^{k-\text{rank}_S(A)}$ solutions to $Ax = 0$ in S . Informally, we say (A, S) is *rich* if this bound is close to tight.

Definition 6.18 (Richness). Let A be an $\ell \times k$ integer matrix, let $\varepsilon > 0$ and let S be either a finite abelian group or a power of a finite subset of a field. We say (A, S) is ε -*rich* if there are at least $\varepsilon|S|^{k-\text{rank}_S(A)}$ solutions to $Ax = 0$ in S .

Intuitively, richness is a measure of how strongly the elements of S are related within the arithmetic structure of S . For example, if S is a finite abelian group then (A, S) is 1-rich by (4) for any integer matrix A . If A is partition regular then Theorem 3.5 implies (in a very strong form) that $S = [n]$ is δ -rich for some $\delta > 0$ and $n \in \mathbb{N}$ sufficiently large. On the other hand, this is not the case for $S = \mathbf{P}_n$ (see, e.g., Theorem 2.1), and indeed, the primes have a very loose arithmetic structure.

The next corollary states that if (A, S) is rich then the bound given by Lemma 6.16 is close to tight for certain projected solutions.

Corollary 6.19. *Let A be an $\ell \times k$ integer matrix and let $0 < \varepsilon \leq 1$. Suppose that either*

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

If (A, S) is ε -rich, then for every non-empty $Z \subseteq [k]$ we have

$$(7) \quad \varepsilon|S|^{|Z|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{Z}})} \leq |\text{Sol}_S^A(Z)| \leq |S|^{|Z|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{Z}})}.$$

Proof. Fix $Z \subseteq [k]$. Applying (5) from Lemma 6.16 with $Y := Z$ and $W := \emptyset$ yields the upper bound

$$|\text{Sol}_S^A(Z)| \leq |S|^{|Z|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{Z}})}.$$

Fix $z_0 \in \text{Sol}_S^A(Z)$. By applying (5) with $Y := [k]$ and $W := Z$, we obtain that

$$(8) \quad |\text{Sol}_S^A(z_0, Z, [k])| \leq |S|^{k-|Z|-\text{rank}_S(A_{\overline{Z}})}.$$

The total number of solutions to $Ax = 0$ in S is

$$|\text{Sol}_S^A([k])| = \sum_{z_0 \in \text{Sol}_S^A(Z)} |\text{Sol}_S^A(z_0, Z, [k])| \stackrel{(8)}{\leq} |\text{Sol}_S^A(Z)| \cdot |S|^{k-|Z|-\text{rank}_S(A_{\overline{Z}})}.$$

Combining the above with $|\text{Sol}_S^A([k])| \geq \varepsilon|S|^{k-\text{rank}_S(A)}$ yields

$$|\text{Sol}_S^A(Z)| \geq \varepsilon|S|^{k-\text{rank}_S(A)} \cdot |S|^{-k+|Z|+\text{rank}_S(A_{\overline{Z}})} = \varepsilon|S|^{|Z|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{Z}})}.$$

□

6.3. Conditions (A3)–(A6) for finite groups and powers of finite subsets of fields. In this subsection, we use the key bounding results (Lemma 6.16 and Corollary 6.19) from the previous subsection to investigate conditions (A3)–(A6) when the sets S_n in the sequence $(S_n)_{n \in \mathbb{N}}$ are either finite abelian groups or powers of finite subsets of fields. We start by showing that in this case, richness implies condition (A3).

Lemma 6.20. *Let A be an $\ell \times k$ integer matrix. Suppose that either*

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

If there exists $B > 0$ such that (A, S) is $(1/B)$ -rich, then (A, S) is B -extendable.

Proof. Let $W \subseteq Y \subseteq [k]$ and $w_0 \in S^{|W|}$. It follows from Lemma 6.16 and Corollary 6.19 that

$$\frac{B |\text{Sol}_S^A(Y)|}{|\text{Sol}_S^A(W)|} \stackrel{(7)}{\geq} \frac{|S|^{|Y| - \text{rank}_S(A) + \text{rank}_S(A_{\overline{Y}})}|}{|S|^{|W| - \text{rank}_S(A) + \text{rank}_S(A_{\overline{W}})}|} = |S|^{|Y| - |W| - \text{rank}_S(A_{\overline{W}}) + \text{rank}_S(A_{\overline{Y}})} \stackrel{(5)}{\geq} |\text{Sol}_S^A(w_0, W, Y)|.$$

Hence (A, S) is B -extendable. □

For finite abelian groups, we deduce the following corollary.

Corollary 6.21. *Let A be an $\ell \times k$ integer matrix and S be a finite abelian group. Then (A, S) is 1-extendable.*

Proof. By (4), (A, S) is 1-rich. Lemma 6.20 implies that (A, S) is 1-extendable. □

Next, we consider conditions (A4), (A5) and (A6). First, we give the definition of *abundance*. This notion appeared in Section 4 when we considered the random Ramsey problem for the integers.

Definition 6.22 (Abundance). Let A be an $\ell \times k$ integer matrix and S be either a finite abelian group or a power of a finite subset of a field. We say (A, S) is *abundant* if $\text{rank}_S(A) = \text{rank}_S(A_{\overline{W}})$ for every $W \subseteq [k]$ with $|W| = 2$.

Before we show how abundance can help simplify conditions (A4), (A5) and (A6), we show it is a mild requirement in the case of powers of finite subsets of a field \mathbb{F} . For brevity, we will write 0 and 1 for the additive and multiplicative identities of \mathbb{F} , respectively. The next lemma establishes that irredundancy and 3-partition regularity imply abundance. We remark that the field \mathbb{F} in the statement of the lemma may be finite or infinite.

Lemma 6.23. *Let A be an $\ell \times k$ integer matrix and $S = L^d$ for some finite subset L of a field \mathbb{F} and $d \in \mathbb{N}$. If A is irredundant with respect to \mathbb{F}^d and 3-partition regular in $S \setminus \{0\}^d$, then (A, S) is abundant.*

Proof. If $L = \{0\}$ then $\text{rank}_S(A') = 0$ for any integer matrix A' , and we are done. Thus, suppose that $L \neq \{0\}$. As before, we can view A as a matrix with entries in \mathbb{F} . Let C_1, \dots, C_k denote the columns of A . To prove that (A, S) is abundant it suffices to show that C_1 and C_2 are both spanned by C_3, \dots, C_k in \mathbb{F} . For the rest of the proof, given a solution x to $Ax = 0$, we write x_i for the i th entry of x .

If C_1 cannot be expressed as a linear combination of C_2, \dots, C_k , then a solution x to $Ax = 0$ in \mathbb{F} satisfies $x_1 = 0$. In turn, this implies that a solution x to $Ax = 0$ in \mathbb{F}^d satisfies $x_1 = \{0\}^d$. This contradicts the assumption that A is 3-partition regular in $S \setminus \{0\}^d$.

Hence, C_1 is spanned by C_2, \dots, C_k . In particular, there exists $\alpha \in \mathbb{F}$ such that $C_1 + \alpha C_2$ is spanned by C_3, \dots, C_k . If α is not unique then one can write C_2 as a linear combination of C_3, \dots, C_k and thus we are done.

Suppose therefore that α is unique. Thus, any solution x to $Ax = 0$ in \mathbb{F}^d must satisfy $x_2 = \alpha \cdot x_1$. If $\alpha = 1$ then $x_1 = x_2$ and A is therefore not irredundant with respect to \mathbb{F}^d , a contradiction. If

$\alpha = 0$ then $x_2 = \{0\}^d$, and so A is not 3-partition regular in $S \setminus \{0\}^d$, a contradiction. Thus, $\alpha \neq 0, 1$.

Consider the graph G with vertex set $V(G) = S$ and edge set $E(G) = \{ss' : s' = \alpha \cdot s\}$ (here the edges are unordered pairs). Since $\alpha \neq 0$, each $s \in V(G)$ is incident to at most two edges, namely $\alpha \cdot s$ and $\alpha^{-1} \cdot s$. Thus $\Delta(G) \leq 2$. In particular, there exists a proper 3-vertex-colouring of G which induces a 3-colouring of S .

Recall that for any solution x of $Ax = 0$ in \mathbb{F}^d we have $x_2 = \alpha \cdot x_1$. Since $\alpha \neq 1$, x_1 and x_2 are distinct unless $x_1 = x_2 = \{0\}^d$. In the former case, x_1 and x_2 are adjacent in G and so they are coloured differently. It follows that there is no monochromatic solution to $Ax = 0$ in $S \setminus \{0\}^d$. Hence A is not 3-partition regular in $S \setminus \{0\}^d$, a contradiction. \square

Note that the statement of Lemma 6.23 does not necessarily hold for finite abelian groups. Indeed, consider the example $A = \begin{pmatrix} 2 & 2 & 1 & 1 \end{pmatrix}$ and $S = \mathbb{Z}_6^n$. The columns of A sum to zero (mod 6), so A satisfies the 6-columns condition. By Theorem 3.8, for fixed r and n sufficiently large, every r -colouring of $\mathbb{Z}_6^n \setminus \{0\}^n$ yields (many) monochromatic solutions to $Ax = 0$. Additionally A is irredundant with respect to \mathbb{Z}_6^n . This follows easily by observing that $(2, 4, 1, 5)$ is a 4-distinct solution to $Ax = 0$ in \mathbb{Z}_6 . However (A, \mathbb{Z}_6^n) is not abundant, since the rank decreases after deleting the last two columns of A .

Next, we show how abundancy can help simplify conditions (A4), (A5) and (A6). First, if (A, S) is abundant and rich then condition (A4) trivially holds.

Lemma 6.24. *Let A be an $\ell \times k$ integer matrix. Suppose that either*

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

Furthermore, assume that (A, S) is abundant and ε -rich for some $0 < \varepsilon \leq 1$. Then, for every $W \subseteq Y \subseteq [k]$ with $|W| = 1$ and $w_0 \in S$ we have $|\text{Sol}_S^A(w_0, W, Y)| \leq |\text{Sol}_S^A(Y)|/(\varepsilon|S|)$.

Proof. By Lemma 6.16 we have

$$|\text{Sol}_S^A(w_0, W, Y)| \leq |S|^{|Y|-1-\text{rank}_S(A_{\overline{W}})+\text{rank}_S(A_{\overline{Y}})},$$

while by Corollary 6.19 we have

$$|\text{Sol}_S^A(Y)| \geq \varepsilon|S|^{|Y|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{Y}})}.$$

Combining the two inequalities above yields

$$(9) \quad |\text{Sol}_S^A(w_0, W, Y)| \leq \frac{1}{\varepsilon} \cdot \frac{|\text{Sol}_S^A(Y)|}{|S|} \cdot |S|^{\text{rank}_S(A)-\text{rank}_S(A_{\overline{W}})}.$$

Let $W \subseteq W'$ with $|W'| = 2$ and note that $\text{rank}_S(A_{\overline{W'}}) \leq \text{rank}_S(A_{\overline{W}}) \leq \text{rank}_S(A)$. We have $\text{rank}_S(A_{\overline{W'}}) = \text{rank}_S(A)$ since $|W'| = 2$ and (A, S) is abundant. Thus, $\text{rank}_S(A_{\overline{W}}) = \text{rank}_S(A)$ and so (9) implies that $|\text{Sol}_S^A(w_0, W, Y)| \leq |\text{Sol}_S^A(Y)|/(\varepsilon|S|)$. \square

Next, we consider conditions (A5) and (A6). If (A, S) is abundant and rich then the following bound holds.

Lemma 6.25. *Let A be an $\ell \times k$ integer matrix. Suppose that either*

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

Furthermore, assume that (A, S) is abundant and ε -rich for some $0 < \varepsilon \leq 1$. Then for any $W \subseteq [k]$ with $|W| = 2$, we have

$$\frac{|S|^2}{|\text{Sol}_S^A(W)|} \leq \frac{1}{\varepsilon}.$$

Proof. We have

$$\frac{|S|^2}{|\text{Sol}_S^A(W)|} \stackrel{(7)}{\leq} \frac{|S|^2}{\varepsilon|S|^{|W|-\text{rank}_S(A)+\text{rank}_S(A_{\overline{W}})}} = \frac{|S|^2}{\varepsilon|S|^2} = \frac{1}{\varepsilon},$$

where the equality follows from the fact that $|W| = 2$ and $\text{rank}_S(A) = \text{rank}_S(A_{\overline{W}})$ since (A, S) is abundant. \square

Observe that if A and $(S_n)_{n \in \mathbb{N}}$ satisfy the hypothesis of Lemma 6.25 with the same $\varepsilon > 0$ for all $n \in \mathbb{N}$, and $|S_n| \rightarrow \infty$ as $n \rightarrow \infty$, then condition (A6) trivially holds. Condition (A5) is also close to being satisfied: it would suffice to additionally prove that there exists some sequence $(X_n)_{n \in \mathbb{N}}$ of subsets $X_n \subseteq [k]$ with $|X_n| \geq 3$ for all $n \in \mathbb{N}$ where $p_{X_n}(A, S_n) = \Omega(\hat{p}(A, S_n))$ and such that

$$\frac{p_{W'_n}(A, S_n)}{p_{X_n}(A, S_n)} \rightarrow 0$$

as $n \rightarrow \infty$ for any sequence $(W'_n)_{n \in \mathbb{N}}$ with $W'_n \subset X_n$ and $|W'_n| \geq 2$. The next result states that this is the case for powers of subsets of fields as long as we have abundancy and richness.

Lemma 6.26. *Let $0 < \varepsilon \leq 1$ and let A be an $\ell \times k$ integer matrix and $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$. Suppose $\text{rank}_S(A) > 0$. If (A, S) is ε -rich and abundant, then there exists $X \subseteq [k]$ such that $|X| \geq 3$, $p_X(A, S) \geq \varepsilon \hat{p}(A, S)$ and for every $W' \subset X$ with $|W'| \geq 2$ we have*

$$\frac{p_{W'}(A, S)}{p_X(A, S)} \leq (1/\varepsilon)|S|^{-1/k^2}.$$

Proof. For every $Z \subseteq [k]$ with $|Z| \geq 2$ set

$$D(Z) := \frac{|Z| - 1 - \text{rank}_S(A) + \text{rank}_S(A_{\overline{Z}})}{|Z| - 1}.$$

As $0 < \varepsilon \leq 1$, by definition of $p_Z(A, S)$ and Corollary 6.19, we have that

$$(10) \quad |S|^{-D(Z)} \stackrel{(7)}{\leq} p_Z(A, S) \stackrel{(7)}{\leq} (1/\varepsilon)|S|^{-D(Z)}.$$

Pick $X \subseteq [k]$ with $|X| \geq 2$ so that the function $D(X)$ is minimised; additionally choose X so that $|X|$ is as small as possible under this assumption. We have

$$p_X(A, S) \stackrel{(10)}{\geq} |S|^{-D(X)} \geq |S|^{-D(Z)} \stackrel{(10)}{\geq} \varepsilon p_Z(A, S)$$

for every $Z \subseteq [k]$ with $|Z| \geq 2$. By maximising over Z , it follows that $p_X(A, S) \geq \varepsilon \hat{p}(A, S)$, as required. Moreover, $|X| \geq 3$. Indeed, for any $Z \subseteq [k]$ with $|Z| = 2$, we have $\text{rank}_S(A) = \text{rank}_S(A_{\overline{Z}})$ since (A, S) is abundant. In particular, we have $D(Z) = 1$. On the other hand, $D([k]) = (k - 1 - \text{rank}_S(A))/(k - 1) < 1$ since $\text{rank}_S(A) > 0$. Thus, $|X| \neq 2$.

Let $W' \subset X$ with $|W'| \geq 2$. We have

$$\frac{p_{W'}(A, S)}{p_X(A, S)} \stackrel{(10)}{\leq} (1/\varepsilon)|S|^{-D(W')+D(X)}.$$

By the minimality of X , we have $D(X) < D(W')$. Since $S = L^d$ for some finite subset L of a field, $\text{rank}_S(\cdot)$ is always an integer and so

$$D(W') - D(X) \geq \frac{1}{(|X| - 1)(|W'| - 1)} \geq \frac{1}{k^2}.$$

It follows that

$$\frac{p_{W'}(A, S)}{p_X(A, S)} \leq (1/\varepsilon)|S|^{-1/k^2}.$$

\square

6.4. Simplified random Rado lemma. In this subsection we state a simplified version of the random Rado lemma for finite abelian groups and powers of finite subsets of fields. This version of the random Rado lemma will be used to prove most of our applications (see Section 11). First, we introduce a generalisation of the parameter $m(A)$ to this setting.

Definition 6.27. Let S be either a finite abelian group or a power of a finite subset of a field. Let A be an $\ell \times k$ integer matrix. We set

$$(11) \quad m_S(A) := \max_{\substack{W \subseteq [k] \\ |W| \geq 2}} \frac{|W| - 1}{|W| - 1 + \text{rank}_S(A_{\overline{W}}) - \text{rank}_S(A)}.$$

Here we recall that $\text{rank}_S(A_{\overline{W}}) := 0$ if $\overline{W} = \emptyset$. We call A *strictly balanced with respect to S* if the expression in (11) is maximised precisely when $W = [k]$.

Note that if $S := [n]$ then as S is a subset of the field \mathbb{Q} , we have that $m_S(A) = m(A)$. Indeed, in this setting the function rank_S is the same as the rank function appearing in the definition of $m(A)$.

Remark 6.28. Observe that $m_S(A)$ is not always well-defined since the denominator in (11) might be zero. However, if (A, S) is abundant then for all $|W| \geq 2$ we have

$$(12) \quad \text{rank}_S(A_{\overline{W}}) \geq \text{rank}_S(A) - |W| + 2,$$

since the rank of a matrix can decrease by at most one by deleting a column.⁷ If (12) is satisfied then the denominator in (11) is always strictly positive and so $m_S(A)$ is well-defined in this case.

Before proceeding further, we provide several examples and facts about the parameter $m_S(A)$. Firstly, we show that $m_{\mathbb{Z}_4}(A) = 4/3$ for $A = \begin{pmatrix} 2 & 2 & -2 \end{pmatrix}$, as stated in Section 5.

Example 6.29. Let $A = \begin{pmatrix} 2 & 2 & -2 \end{pmatrix}$. The number of solutions $x = (x_1, x_2, x_3)$ to $Ax = 0$ in \mathbb{Z}_4 is precisely $2 \cdot 4^2 = 4^{3-1/2}$: we have 4^2 choices for x_1, x_2 and, for each fixed choice, we have 2 choices for x_3 . It follows from (4) that $\text{rank}_{\mathbb{Z}_4}(A) = 1/2$. Now, let $W \subseteq [3]$. If $|W| = 3$ we trivially have $\text{rank}_{\mathbb{Z}_4}(A_{\overline{W}}) = 0$. If $|W| = 2$ then $A_{\overline{W}}x = 0$ is the same as $2x = 0$, which has $2 = 4^{1-1/2}$ solutions in \mathbb{Z}_4 . So we have $\text{rank}_{\mathbb{Z}_4}(A_{\overline{W}}) = 1/2$ again. Thus,

$$m_{\mathbb{Z}_4}(A) = \max \left\{ \frac{3-1}{3-1+0-1/2}, \frac{2-1}{2-1+1/2-1/2} \right\} = 4/3.$$

Next, we highlight certain similarities and differences between the parameters $m(A)$ and $m_S(A)$. Recall that if A is a $1 \times k$ integer matrix with non-zero entries then A is strictly balanced and $m(A) = \frac{k-1}{k-2}$. This is not always the case for $m_S(A)$, as shown in the next example.

Example 6.30. Let $m \in \mathbb{N}$ and $A = \begin{pmatrix} 1 & m & m \end{pmatrix}$. The number of solutions $x = (x_1, x_2, x_3)$ to $Ax = 0$ in \mathbb{Z}_{2m} is precisely $(2m)^2 = (2m)^{3-1}$. It follows that $\text{rank}_{\mathbb{Z}_{2m}}(A) = 1$. If $W = [3]$ we trivially have $\text{rank}_{\mathbb{Z}_4}(A_{\overline{W}}) = 0$. If $W = \{2, 3\}$ then $A_{\overline{W}}x = 0$ is the same as $x = 0$, which has $1 = (2m)^{1-1}$ solution in \mathbb{Z}_{2m} , and so $\text{rank}_{\mathbb{Z}_{2m}}(A_{\overline{W}}) = 1$. If $W = \{1, 2\}$ or $W = \{1, 3\}$ then $A_{\overline{W}}x = 0$ is the same as $mx = 0$, which has $m = (2m)^{1-\log_{2m} 2}$ solutions in \mathbb{Z}_{2m} , and so $\text{rank}_{\mathbb{Z}_{2m}}(A_{\overline{W}}) = \log_{2m} 2$. Thus,

$$m_{\mathbb{Z}_{2m}}(A) = \max \left\{ \frac{3-1}{3-1+0-1}, \frac{2-1}{2-1+1-1}, \frac{2-1}{2-1+\log_{2m} 2-1} \right\} = \max\{2, 1, (\log_{2m} 2)^{-1}\}.$$

In particular, for $m \geq 2$ we have $m_{\mathbb{Z}_{2m}}(A) = (\log_{2m} 2)^{-1}$. Furthermore, for $m > 2$, the value of $m_{\mathbb{Z}_{2m}}(A)$ is achieved precisely by $W = \{1, 2\}$ and $W = \{1, 3\}$, and so A is not strictly balanced with respect to \mathbb{Z}_{2m} .

⁷This fact follows immediately from the definition of $\text{rank}_S(A)$ when S is the power of a subset of a field F . When S is a finite abelian group, it can be proved using equation (4).

Despite the example above, by imposing some additional restrictions on the $1 \times k$ matrix A , one can infer that A is strictly balanced with respect to S and, furthermore, $m_S(A) = m(A) = \frac{k-1}{k-2}$.

Example 6.31. Let A be a $1 \times k$ integer matrix with non-zero entries. Let S be a finite subset of an abelian group G . Suppose that $S \neq \{0_G\}$ and that either

- S is a finite abelian group or
- S is a power of a finite subset of a field \mathbb{F} .

Furthermore, suppose that

(α) for every entry a of A and $s \in S \setminus \{0_G\}$ we have $a \cdot s \neq 0$.

Then A is strictly balanced with respect to S and $m_S(A) = m(A) = \frac{k-1}{k-2}$.

Indeed, let $W \subseteq [k]$ be non-empty. If S is a power of a finite subset of a field \mathbb{F} then (α) implies A_W is not the $1 \times |W|$ zero vector in \mathbb{F} and thus A_W has rank 1 with respect to \mathbb{F} . In particular, $\text{rank}_S(A_W) = 1$. Similarly, if S is a finite abelian group then the number of solutions to $A_W x = 0$ in S is precisely $|S|^{|W|-1}$: we can pick the first $|W| - 1$ entries of x arbitrarily and for each such choice we have 1 choice for the last entry (this follows from (α)). Thus, we have $\text{rank}_S(A_W) = 1$. In both cases, we have

$$m_S(A) = \max_{2 \leq w \leq k-1} \left\{ \frac{k-1}{k-1+0-1}, \frac{w-1}{w-1+1-1} \right\} = \frac{k-1}{k-2}.$$

In particular, the value of $m_S(A)$ is achieved precisely when $W = [k]$ and so A is strictly balanced with respect to S .

Note that if a $1 \times k$ integer matrix A is strictly balanced then $m(A) = \frac{k-1}{k-2}$. Conversely, if A is strictly balanced with respect to S , it may not be the case that $m_S(A) = \frac{k-1}{k-2}$. For instance, in Example 6.29, we saw $A = (2 \ 2 \ -2)$ is strictly balanced with respect to \mathbb{Z}_4 but $m_{\mathbb{Z}_4}(A) = 4/3 < 2$.

Another observation is that, for A fixed, $m_S(A)$ may change depending on S . For example, consider $A = (1 \ 3 \ 3)$. From Example 6.30 we have that $m_{\mathbb{Z}_6}(A) = (\log_6 2)^{-1} \approx 2.58$, while from Example 6.31 we have $m_{\mathbb{Z}_2}(A) = \frac{3-1}{2-1} = 2$ (since A and \mathbb{Z}_2 satisfy property (α)).

There are cases where $m_S(A)$ is independent of S , as long as S is non-trivial. Indeed, let A be a $1 \times k$ matrix whose entries are all from $\{1, -1\}$. Then (α) from Example 6.31 holds for every S which is either a finite abelian group or a power of a finite subset of a field, provided that S includes a non-identity element. In particular, we have $m_S(A) = \frac{k-1}{k-2}$ for any such S .

The next result states that $\hat{p}(A, S)$ and $|S|^{-1/m_S(A)}$ are within a constant factor of each other, as long as we are in a rich setting.

Lemma 6.32. Let A be an $\ell \times k$ integer matrix and $0 < \varepsilon \leq 1$. Suppose that either

- (i) S is a finite abelian group or
- (ii) $S = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$.

If $m_S(A)$ is well-defined and (A, S) is ε -rich then

$$|S|^{-\frac{1}{m_S(A)}} \leq \hat{p}(A, S) \leq (1/\varepsilon^2) |S|^{-\frac{1}{m_S(A)}}.$$

Proof. For every $Z \subseteq [k]$ with $|Z| \geq 2$ let

$$D(Z) := \frac{|Z| - 1 - \text{rank}_S(A) + \text{rank}_S(A_Z)}{|Z| - 1}.$$

As $0 < \varepsilon \leq 1$, by definition of $p_Z(A, S)$ and Corollary 6.19, we have that

$$(13) \quad |S|^{-D(Z)} \stackrel{(7)}{\leq} p_Z(A, S) \stackrel{(7)}{\leq} (1/\varepsilon) |S|^{-D(Z)}.$$

Pick $X \subseteq [k]$ with $|X| \geq 2$ so that the function $D(X)$ is minimised; so $m_S(A) = 1/D(X)$. For any $Z \subseteq [k]$ with $|Z| \geq 2$, we have

$$p_X(A, S) \stackrel{(13)}{\geq} |S|^{-D(X)} \geq |S|^{-D(Z)} \stackrel{(13)}{\geq} \varepsilon p_Z(A, S).$$

As Z was arbitrary, it follows that

$$p_X(A, S) \leq \hat{p}(A, S) \leq (1/\varepsilon)p_X(A, S).$$

Combining the above with (13) yields

$$|S|^{-\frac{1}{m_S(A)}} = |S|^{-D(X)} \stackrel{(13)}{\leq} \hat{p}(A, S) \stackrel{(13)}{\leq} (1/\varepsilon^2)|S|^{-D(X)} = (1/\varepsilon^2)|S|^{-\frac{1}{m_S(A)}}.$$

□

We can now state the simplified version of the random Rado lemma.

Lemma 6.33 (Simplified random Rado lemma). *Let $r \geq 2$, $k \geq 3$ and $\ell \geq 1$ be integers. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence where S_n is either a finite abelian group or $S_n = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$, for every $n \in \mathbb{N}$. Let A be an $\ell \times k$ integer matrix and set $\hat{p}_n := \hat{p}(A, S_n)$. Suppose that*

- (C1) $\hat{p}_n \rightarrow 0$ and $|S_n|\hat{p}_n \rightarrow \infty$ as $n \rightarrow \infty$;
- (C2) $(S_n)_{n \in \mathbb{N}}$ is (A, r) -supersaturated;
- (C3) there exists $0 < \varepsilon \leq 1$ such that (A, S_n) is ε -rich for every sufficiently large $n \in \mathbb{N}$;
- (C4) (A, S_n) is abundant for every sufficiently large $n \in \mathbb{N}$;
- (C5) given any $n \in \mathbb{N}$ there exists $X_n \subseteq [k]$ with $|X_n| \geq 3$ such that $p_{X_n}(A, S_n) = \Omega(\hat{p}(A, S_n))$ and for every sequence $(W'_n)_{n \in \mathbb{N}}$ with $W'_n \subset X_n$ and $|W'_n| \geq 2$ we have

$$\frac{p_{W'_n}(A, S_n)}{p_{X_n}(A, S_n)} \rightarrow 0$$

as $n \rightarrow \infty$.

Then there exist constants $c, C > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c|S_n|^{-\frac{1}{m_{S_n}(A)}}; \\ 1 & \text{if } p \geq C|S_n|^{-\frac{1}{m_{S_n}(A)}}. \end{cases}$$

Remark 6.34. *Note that Lemma 6.33 can be further simplified if we consider finite abelian groups and powers of finite subsets of fields separately. Namely, if all sets in $(S_n)_{n \in \mathbb{N}}$ are finite abelian groups then condition (C3) holds since by (4) (A, S_n) is 1-rich. On the other hand, if all sets in $(S_n)_{n \in \mathbb{N}}$ are powers of finite subsets of fields and $\text{rank}_{S_n}(A) > 0$ for all $n \in \mathbb{N}$, then condition (C5) follows from conditions (C1), (C3) and (C4) combined with Lemma 6.26. Furthermore, in this case, condition (C4) can be replaced by irredundancy and 3-partition regularity; see Lemma 6.23.*

If we set $S_n := [n]$ and let A be an irredundant partition regular matrix, then it is easy to check that (C1) and (C3)–(C5) hold. Theorem 3.5 gives (C2), and therefore Lemma 6.33 recovers the random Rado theorem (i.e., Theorems 4.2 and 4.3).⁸

Recall that Lemma 6.8 confirms the probability threshold essentially depends on the number of (projected) solutions to $Ax = 0$ in S_n . For sequences that satisfy the hypothesis of Lemma 6.33, we see that we can express this phenomenon cleanly in terms of the parameter $m_{S_n}(A)$. In particular, if $(S_n)_{n \in \mathbb{N}}$ is a sequence as in Lemma 6.33 then, for large $n \in \mathbb{N}$, a typical subset of S_n of size significantly greater than $|S_n|^{1-1/m_{S_n}(A)}$ will be (A, r) -Rado, whilst a typical subset of size significantly smaller than $|S_n|^{1-1/m_{S_n}(A)}$ will not be (A, r) -Rado.

⁸We formally verify this in the more general setting of integer lattices in the proof of Theorem 5.1 in Section 11.

The proof of Lemma 6.33 follows easily from our original random Rado lemma, Lemma 6.8, combined with the results from the previous subsection.

Proof of Lemma 6.33. Let $r \geq 2$, $k \geq 3$ and $\ell \geq 1$ be integers. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence where S_n is either a finite abelian group or $S_n = L^d$ for some finite subset L of a field and $d \in \mathbb{N}$, for every $n \in \mathbb{N}$. Suppose that conditions (C1)–(C5) hold. First we show that conditions (A1)–(A6) hold too.

Conditions (A1) and (A2) are identical to (C1) and (C2) respectively. Condition (A3) follows from condition (C3) and Lemma 6.20.

Pick arbitrary $W \subset Y \subseteq [k]$ with $|W| = 1$. As (C3) and (C4) hold, we can apply Lemma 6.24 and deduce that $|\text{Sol}_{S_n}^A(w_0, W, Y)| \leq |\text{Sol}_{S_n}^A(Y)|/(\varepsilon|S_n|)$ for every sufficiently large $n \in \mathbb{N}$ and every $w_0 \in S_n$. This immediately implies condition (A4).

As (C3) and (C4) hold, we can apply Lemma 6.25 to A and S_n , for any sufficiently large $n \in \mathbb{N}$, to obtain the following: for any $W \subseteq [k]$ with $|W| = 2$, we have

$$\frac{|S_n|^2}{|\text{Sol}_{S_n}^A(W)|} \leq \frac{1}{\varepsilon}.$$

Combining the above with conditions (C1) and (C5) yields conditions (A6) and (A5) respectively.

As conditions (A1)–(A6) hold, by Lemma 6.8 there exist constants $c_0, C_0 > 0$ such that

$$(14) \quad \lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c_0 \cdot \hat{p}(A, S_n); \\ 1 & \text{if } p \geq C_0 \cdot \hat{p}(A, S_n). \end{cases}$$

Condition (C4) implies $m_{S_n}(A)$ is well-defined for every sufficiently large $n \in \mathbb{N}$ (see, e.g., Remark 6.28). Condition (C3) combined with Lemma 6.32 implies that, for sufficiently large $n \in \mathbb{N}$,

$$(15) \quad |S_n|^{-\frac{1}{m_{S_n}(A)}} \leq \hat{p}(A, S_n) \leq (1/\varepsilon^2)|S_n|^{-\frac{1}{m_{S_n}(A)}}.$$

Let $C := (1/\varepsilon^2)C_0$ and $c := c_0$. Combining (14) and (15) yields

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c|S_n|^{-\frac{1}{m_{S_n}(A)}}; \\ 1 & \text{if } p \geq C|S_n|^{-\frac{1}{m_{S_n}(A)}}. \end{cases}$$

□

7. THE GENERAL 1-STATEMENT AND 0-STATEMENT FOR HYPERGRAPHS

In this section, we formulate our general random Ramsey 1-statement and 0-statement for hypergraphs. We combine both statements in a single result (Theorem 7.7), but first we provide some definitions and motivation.

7.1. Basic definitions for hypergraphs.

Definition 7.1 (Hypergraphs). Let $k \in \mathbb{N}$. A k -uniform hypergraph \mathcal{H} is a pair $(V(\mathcal{H}), E(\mathcal{H}))$ where $V(\mathcal{H})$ is a finite set and $E(\mathcal{H})$ is a set of *unordered* k -tuples $\{v_1, \dots, v_k\} \in V(\mathcal{H})^k$ with $v_i \neq v_j$ for $i \neq j$. A k -uniform ordered hypergraph \mathcal{H} is a pair $(V(\mathcal{H}), E(\mathcal{H}))$ where $V(\mathcal{H})$ is a finite set and $E(\mathcal{H})$ is a set of *ordered* k -tuples $(v_1, \dots, v_k) \in V(\mathcal{H})^k$ with $v_i \neq v_j$ for $i \neq j$.

$V(\mathcal{H})$ and $E(\mathcal{H})$ are respectively the *vertex set* and the *edge set* of \mathcal{H} . We write $v(\mathcal{H}) := |V(\mathcal{H})|$ and $e(\mathcal{H}) := |E(\mathcal{H})|$. Next, we define the analogues of the notions of *partition regularity* and *supersaturation* for hypergraphs and ordered hypergraphs.

Definition 7.2 (r -Ramsey, (r, ε) -supersaturated and r -supersaturated). Let \mathcal{H} be an (ordered) hypergraph, $r \in \mathbb{N}$ and $\varepsilon > 0$. We say \mathcal{H} is r -Ramsey if every r -colouring of $V(\mathcal{H})$ yields a monochromatic edge in $E(\mathcal{H})$. Similarly, \mathcal{H} is (r, ε) -supersaturated if every r -colouring of $V(\mathcal{H})$ yields at least $\varepsilon e(\mathcal{H})$ monochromatic edges. We say that a sequence $(\mathcal{H}_n)_{n \in \mathbb{N}}$ of (ordered) hypergraphs is r -supersaturated if there exists $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, \mathcal{H}_n is (r, ε) -supersaturated.

Before proceeding further, we give an informal explanation of how a Ramsey problem can be rephrased in terms of (ordered) hypergraphs. Say \mathcal{F} is some discrete object (e.g., a graph, an abelian group etc.) over a ground set of elements. Given a colouring of the elements in \mathcal{F} , we are interested in finding a monochromatic copy of a certain structure F in \mathcal{F} . We consider the (ordered) hypergraph \mathcal{H} whose vertex set is the set of elements of \mathcal{F} and the edge set is the set of copies of F in \mathcal{F} . Clearly, the colouring of the elements of \mathcal{F} induces a colouring of the vertices of \mathcal{H} . In particular, every monochromatic copy of F in \mathcal{F} corresponds to a monochromatic edge in \mathcal{H} and vice versa. Consider the following example, where $\mathcal{F} := [n]$ and F is a 3-distinct solution to the equation $x + y = z$.

Example 7.3. Any 2-colouring of $[n]$ yields a monochromatic 3-distinct solution to the equation $x + y = z$ as long as n is sufficiently large (A). In fact, Theorem 3.5 implies that there exists $c > 0$ so that if n is sufficiently large, then any 2-colouring of $[n]$ yields at least cn^2 3-distinct monochromatic solutions⁹ (B).

Consider the 3-uniform ordered hypergraph \mathcal{H} with $V(\mathcal{H}) = [n]$ and $E(\mathcal{H}) = \{(x, y, z) : x + y = z\}$. Statement (A) is equivalent to \mathcal{H} being 2-Ramsey as long as $v(\mathcal{H})$ is sufficiently large while statement (B) is equivalent to \mathcal{H} being $(2, c')$ -supersaturated for some constant $c' > 0$, provided $v(\mathcal{H})$ is sufficiently large.

In practice, we can usually work in the setting of *unordered* hypergraphs.

Definition 7.4. Given an ordered hypergraph \mathcal{H} , the *unordered restriction* of \mathcal{H} is the hypergraph \mathcal{H}' obtained from \mathcal{H} by ignoring the order of the edges in $E(\mathcal{H})$. Formally, $V(\mathcal{H}') := V(\mathcal{H})$ and

$$E(\mathcal{H}') := \{\{v_1, \dots, v_k\} : (v_1, \dots, v_k) \in E(\mathcal{H})\}.$$

Note that we do not allow multiple edges here; that is, if two ordered edges e_1, e_2 become the same unordered edge e , then we only keep one copy of e .

Remark 7.5. By definition, a k -uniform ordered hypergraph \mathcal{H} is r -Ramsey if and only if the unordered restriction \mathcal{H}' of \mathcal{H} is also r -Ramsey. Similarly, a sequence of k -uniform ordered hypergraphs $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is r -supersaturated if and only if the sequence $(\mathcal{H}'_n)_{n \in \mathbb{N}}$ is r -supersaturated, where \mathcal{H}'_n is the unordered restriction of \mathcal{H}_n .

Next, we define notions of maximum degree for (ordered) hypergraphs. Let \mathcal{H} be a k -uniform ordered hypergraph and let $W \subseteq [k]$, say $W = \{i_1, i_2, \dots, i_{|W|}\}$ where $i_1 < i_2 < \dots < i_{|W|}$. The *restriction* \mathcal{H}_W of \mathcal{H} is the $|W|$ -uniform ordered hypergraph with $V(\mathcal{H}_W) := V(\mathcal{H})$ and

$$E(\mathcal{H}_W) := \{(x_{i_1}, \dots, x_{i_{|W|}}) : \exists (y_1, \dots, y_k) \in E(\mathcal{H}) \text{ with } y_{i_s} = x_{i_s} \text{ for each } s \in \{1, \dots, |W|\}\}.$$

Again, we do not allow multiple edges. Given sets $W \subseteq Y \subseteq [k]$, we define $\Delta_W(\mathcal{H}_Y)$ to be the maximum number of edges in \mathcal{H}_Y that restrict to the same edge in \mathcal{H}_W . More precisely, given any ordered tuple $f \in E(\mathcal{H}_W)$, let $E_{\mathcal{H}}(f, Y)$ be the set of ordered tuples $e \in E(\mathcal{H}_Y)$ so that f equals

⁹While the statement of Theorem 3.5 does not insist that the solutions are 3-distinct, since there are at most $O(n)$ solutions to $x + y = z$ which are not 3-distinct, it is easy to obtain the result as stated here.

the ordered tuple obtained from e by deleting the entries of e in the positions corresponding to $Y \setminus W$. Then we define

$$\Delta_W(\mathcal{H}_Y) := \max_{f \in E(\mathcal{H}_W)} |E_{\mathcal{H}}(f, Y)|.$$

Let \mathcal{H}' be a k -uniform hypergraph. For $i \in [k]$, we set

$$\Delta_i(\mathcal{H}') := \max_{f \subseteq V(\mathcal{H}'), |f|=i} |\{e \in E(\mathcal{H}') : f \subseteq e\}|.$$

Note that if \mathcal{H}' is the unordered restriction of an ordered hypergraph \mathcal{H} then

$$(16) \quad \frac{1}{k!} \cdot \max_{\substack{W \subseteq [k] \\ |W|=i}} \Delta_W(\mathcal{H}) \leq \frac{1}{(k-i)!} \cdot \max_{\substack{W \subseteq [k] \\ |W|=i}} \Delta_W(\mathcal{H}) \leq \Delta_i(\mathcal{H}') \leq \binom{k}{i} \cdot i! \cdot \max_{\substack{W \subseteq [k] \\ |W|=i}} \Delta_W(\mathcal{H}) \leq k! \cdot \max_{\substack{W \subseteq [k] \\ |W|=i}} \Delta_W(\mathcal{H}).$$

The second inequality in (16) uses that, given any edge e in \mathcal{H}' , there are $(k-i)!$ ordered k -tuples that have e as their underlying (unordered) set and have the entries in i of their positions fixed.

For an (ordered) hypergraph \mathcal{H} , we write \mathcal{H}_p^v to denote the hypergraph obtained from \mathcal{H} by including each *vertex* of \mathcal{H} with probability p independently of all other vertices, while including each edge of \mathcal{H} if and only if all of its vertices are kept. We use the superscript v to remind the reader that it is the vertices that we keep with probability p rather than edges (as in the binomial random hypergraph).

7.2. Heuristic for the probability threshold. The random Ramsey-type results we covered in Section 2 (for the primes), in Section 4 (for graphs and integers) and in Section 5 (for abelian groups) are formulated as follows: given a sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$ of objects that exhibit a Ramsey property in a robust way (i.e., supersaturation), there is probability threshold $\hat{p} = \hat{p}(n)$ such that, as n tends to infinity, the random binomial subset $\mathcal{F}_{n,p}$ retains the Ramsey property w.h.p. if $p = \omega(\hat{p})$ and loses it w.h.p. if $p = o(\hat{p})$. Our general 1-statement and 0-statement for hypergraphs are formulated in a similar fashion.

Definition 7.6 (Probability threshold for hypergraphs). Given a sequence $(\mathcal{H}_n)_{n \in \mathbb{N}}$ of k -uniform ordered hypergraphs and $W \subseteq [k]$, let

$$(17) \quad f_{n,W} := \left(\frac{e(\mathcal{H}_{n,W})}{v(\mathcal{H}_n)} \right)^{-\frac{1}{|W|-1}} \quad \text{and} \quad \hat{p}(\mathcal{H}_n) := \max_{\substack{W \subseteq [k] \\ |W| \geq 2}} f_{n,W}.$$

Note that in Definition 7.6 we slightly abuse notation, writing $\mathcal{H}_{n,W}$ for the restriction $\mathcal{H}_n|_W$ of \mathcal{H}_n .

Our choice of $\hat{p}(\mathcal{H}_n)$ follows precisely the same heuristic as the probability threshold for the random Ramsey problem for graphs and the integers (see Section 4). Namely, consider the expected number of vertices and edges in the random hypergraph $\mathcal{H}_{n,p}^v$. Assuming that the edges in $\mathcal{H}_{n,p}^v$ are distributed uniformly in some way, if $\mathbb{E}(e(\mathcal{H}_{n,p}^v))$ is much less than $\mathbb{E}(v(\mathcal{H}_{n,p}^v))$ then intuitively the edges in $\mathcal{H}_{n,p}^v$ should be fairly spread out. Therefore, one should be able to 2-colour the vertices of $\mathcal{H}_{n,p}^v$ without creating monochromatic edges.

However, similarly to the random Ramsey problem for graphs and the integers, it may be the case that the above reasoning is applicable to a ‘lower’ level in the following sense. For any $W \subseteq [k]$, a 2-colouring that does not yield monochromatic edges in $\mathcal{H}_{n,W,p}^v$ also does not yield monochromatic edges in $\mathcal{H}_{n,p}^v$. By the previous reasoning, if $\mathbb{E}(e(\mathcal{H}_{n,W,p}^v))$ is much less than $\mathbb{E}(v(\mathcal{H}_{n,W,p}^v))$ then one should be able to 2-colour the vertices of $\mathcal{H}_{n,W,p}^v$ while avoiding any monochromatic edge in $\mathcal{H}_{n,W,p}^v$, and therefore in $\mathcal{H}_{n,p}^v$. For any $W \subseteq [k]$ with $|W| \geq 2$, we have

$$\mathbb{E}(v(\mathcal{H}_{n,W,p}^v)) = pv(\mathcal{H}_n) \quad \text{and} \quad \mathbb{E}(e(\mathcal{H}_{n,W,p}^v)) = p^{|W|}e(\mathcal{H}_{n,W}).$$

In particular, $\mathbb{E}(e(\mathcal{H}_{n,W,p}^v)) \leq \mathbb{E}(v(\mathcal{H}_{n,W,p}^v))$ if and only if

$$p \leq \left(\frac{e(\mathcal{H}_{n,W})}{v(\mathcal{H}_n)} \right)^{-\frac{1}{|W|-1}} = f_{n,W}.$$

It is then natural to let $\hat{p}(\mathcal{H}_n)$ be equal to the largest of the $f_{n,W}$'s.

7.3. General 1- and 0-statements for hypergraphs. We are now ready to state our general 1-statement and 0-statement for hypergraphs.

Theorem 7.7. *Let $r, k \geq 2$ be integers, and let $b > 0$. Let $(\mathcal{H}_n)_{n \in \mathbb{N}}$ be a sequence of k -uniform ordered hypergraphs and set $\hat{p}_n := \hat{p}(\mathcal{H}_n)$ as given by (17). Let (P1)–(P5) be the following properties.*

(P1) *Probabilities: $\hat{p}_n \rightarrow 0$ and $\hat{p}_n v(\mathcal{H}_n) \rightarrow \infty$ as $n \rightarrow \infty$.*

(P2) *Supersaturation: $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is r -supersaturated.*

(P3) *Bounded degree: for any $W \subseteq Y \subseteq [k]$, and any sufficiently large $n \in \mathbb{N}$ we have $\Delta_W(\mathcal{H}_{n,Y}) \leq b \frac{e(\mathcal{H}_{n,Y})}{e(\mathcal{H}_{n,W})}$.*

(P4) *Bounded 1-degree: for any $W \subset Y \subseteq [k]$ with $|W| = 1$ and any sufficiently large $n \in \mathbb{N}$, we have $\Delta_W(\mathcal{H}_{n,Y}) \leq b \frac{e(\mathcal{H}_{n,Y})}{v(\mathcal{H}_n)}$.*

(P5) *Bounded 2-degree: there exists a sequence $(X_n)_{n \in \mathbb{N}}$ of subsets $X_n \subseteq [k]$ with $|X_n| \geq 3$ for all $n \in \mathbb{N}$ such that*

$$(18) \quad f_{n,X_n} = \Omega(\hat{p}_n)$$

and

$$(19) \quad \Delta_{W_n}(\mathcal{H}_{n,X_n}) \cdot \Delta_{W'_n}(\mathcal{H}_{n,X_n}) \cdot e(\mathcal{H}_{n,X_n}) \cdot (\hat{p}_n)^{3|X_n|-2-|W'_n|} = o(1)$$

for any sequences $(W_n)_{n \in \mathbb{N}}$ and $(W'_n)_{n \in \mathbb{N}}$ with $W_n, W'_n \subset X_n$, $|W_n| = 2$ and $|W'_n| \geq 2$.

If (P1)–(P4) hold, then there exists a constant $C > 0$ such that if $q_n \geq C\hat{p}_n$ then

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}_{n,q_n}^v \text{ is } r\text{-Ramsey}] = 1.$$

If (P1) and (P3)–(P5) hold, then there exists a constant $c > 0$ such that if $q_n \leq c\hat{p}_n$ then

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}_{n,q_n}^v \text{ is } r\text{-Ramsey}] = 0.$$

All of our new random Ramsey-type results are essentially corollaries of Theorem 7.7. Indeed, we use Theorem 7.7 to prove the random Rado lemma (Lemma 6.8) which we have in turn used to prove the simplified random Rado lemma (Lemma 6.33). Further, the proofs of our new results for the primes (Theorem 2.7), integer lattices (Theorem 5.1), groups (Theorem 5.3) and finite vector spaces (Theorems 5.4 and 5.5) all apply one of these two random Rado lemmas (see Section 11). This in turn implies that the random Rado theorem (Theorems 4.2 and 4.3) is a consequence of Theorem 7.7. As we describe in Appendix A.3, the 1-statement of the random Ramsey theorem (Theorem 4.1) can also be deduced from Theorem 7.7.

The reader should therefore take away the following philosophy. Theorem 7.7 is the most general of our black box results, and its applications are not just restricted to systems of linear equations. On the other hand, the random Rado lemma is more readily applicable to the setting of Ramsey properties in arithmetic structures. The simplified random Rado lemma has the most restricted hypothesis of these three results, but has the advantage that it explicitly produces the corresponding probability threshold.

The proof of Theorem 7.7 is given in Section 8. The conditions required for the 1-statement to hold, namely (P1)–(P4), are natural. Condition (P1) is needed to ensure the expected value of $v(\mathcal{H}_{n,p}^v)$ tends to infinity. In Section 3.2, we already discussed the notion of supersaturation (condition (P2)) and how it is typically required in random Ramsey-type results. (P3) and (P4)

are recurring assumptions when applying the hypergraph container method, which we indeed use to prove Theorem 7.7. They are also natural conditions to make to ensure our heuristic described in Section 7.2 is valid. Indeed, (P3) and (P4) ensure that edges are fairly spread out in \mathcal{H}_n and its associated restrictions $\mathcal{H}_{n,Y}$.

On the other hand, condition (P5) is more artificial: it arises naturally in the setting of abelian groups but fails to hold in general. Indeed, while the 0-statement of the random Rado lemma (Lemma 6.8) follows from Theorem 7.7, the 0-statement for random Ramsey theorem (Theorem 4.1) does not. In Section 8 we explain how the 0-statement of Theorem 7.7 follows from the combination of a *deterministic lemma*, stating that certain substructures must arise in the hypergraph assuming it is r -Ramsey, and a *probabilistic lemma*, stating that such structures do not appear w.h.p. The proof of the latter consists of bounding the expected number of the forbidden structures so that the expectation tends to 0 as n goes to infinity. Condition (P5) is precisely designed to achieve these bounds.

Returning to Theorem 4.1, its 0-statement is proved by using a much stronger deterministic lemma than what we use here. In particular, this statement says that graphs with density at most $m_2(H)$ can be 2-coloured so to avoid a monochromatic copy of H . Using this strong result, one can prove a suitable corresponding probabilistic lemma without needing (P5) to be satisfied.

Note that the use of ordered hypergraphs in Theorem 7.7 is only really necessary for the 0-statement. Indeed, we will deduce the 1-statement of Theorem 7.7 very quickly from a theorem for unordered hypergraphs, Theorem 8.1. The proof of Theorem 8.1 is obtained via the method of containers, which was first formalised in the setting of hypergraphs by Balogh, Morris and Samotij [4] and independently by Saxton and Thomason [53]. The use of containers is now standard in proofs of 1-statements for various Ramsey-type problems, see, e.g., [1, 4, 5, 24, 25, 31, 32, 43, 44, 58]. We emphasise that the proof of Theorem 8.1 is not particularly difficult. Indeed, a researcher familiar with the container method would be able to come up with such a proof, and in particular, our proof here is an easy adaption of the short proof of the 1-statement of Theorem 4.1 given by Nenadov and Steger [44]. The proof of the 0-statement of Theorem 7.7 adapts a recent argument of the second and third authors [33].

There are several results in the literature of a similar flavour to Theorem 7.7. Indeed, Friedgut, Rödl and Schacht proved a similar result to our 1-statement; see [26, Theorem 2.5]. In their hypothesis though, they require the sequence $(\mathcal{H}_n)_{n \in \mathbb{N}}$ to be R -supersaturated for some R much larger than r . Furthermore, Aigner-Horev and Person [1, Theorem 2.8] proved an *asymmetric* version of the 1-statement of Theorem 7.7, where one considers a sequence of r -tuples of hypergraphs $(\mathcal{H}_{n,1}, \dots, \mathcal{H}_{n,r})_{n \in \mathbb{N}}$ on the same vertex set V_n and ask for the probability threshold such that whenever the vertices are r -coloured, there is some $i \in [r]$ such that there is an edge in $\mathcal{H}_{n,i,p}^v$ which is monochromatic in colour i . Note that their result requires some rather technical conditions on the $\mathcal{H}_{n,i}$ which are not necessary in the symmetric setting.

Finally, in their recent paper on sharp Ramsey thresholds, Friedgut, Kuperwasser, Samotij and Schacht [25, Theorem 2.1] gave general criteria for a sequence of hypergraphs $(\mathcal{H}_n)_{n \in \mathbb{N}}$ to have a sharp threshold for being r -Ramsey. The threshold function in their theorem is $\Theta(f_{n,[k]})$, which is the value of our \hat{p}_n in the case where $W = [k]$ is the maximiser of the $f_{n,W}$'s: note that this is indeed the case when considering the Schur property or sets avoiding monochromatic k -APs in \mathbb{Z}_n ; as mentioned at the end of Section 4, these are two of the specific settings for which sharp thresholds are obtained in [25].

8. PROOF OF THEOREM 7.7

8.1. Proof of the 1-statement of Theorem 7.7. In this subsection we state a general 1-statement, Theorem 8.1 for *hypergraphs*. We then show that the 1-statement for ordered hypergraphs (Theorem 7.7) follows immediately from it. We prove Theorem 8.1 in Section 8.1.3.

8.1.1. *The Ramsey property in supersaturated hypergraphs.*

Theorem 8.1. *Let $r, k \geq 2$ be integers and let $c > 0$. Suppose that $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is a sequence of k -uniform hypergraphs. Let $(p_n)_{n \in \mathbb{N}}$ with $p_n \in (0, 1)$ be such that $p_n \rightarrow 0$ and $p_n v(\mathcal{H}_n) \rightarrow \infty$ as $n \rightarrow \infty$, and for each sufficiently large $n \in \mathbb{N}$ and each $\ell \in [k]$, we have*

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

Further, suppose that $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is r -supersaturated. Then there exists $C \geq 1$ such that for $q_n \geq Cp_n$,

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}_{n, q_n}^v \text{ is } r\text{-Ramsey}] = 1.$$

The proof of Theorem 8.1 is in the same spirit as that of the proof of the 1-statement of Theorem 4.1 given by Nenadov and Steger in [44]. In particular, as mentioned in Section 7, it utilises the hypergraph container method.

For a set \mathcal{A} , let $\mathcal{P}(\mathcal{A})$ denote the powerset of \mathcal{A} . For a hypergraph \mathcal{H} , let $\mathcal{I}_r(\mathcal{H})$ be the collection of r -tuples (I_1, \dots, I_r) such that each $I_i \in \mathcal{P}(V(\mathcal{H}))$ is an independent set in \mathcal{H} and $I_i \cap I_j = \emptyset$ for every $i \neq j \in [r]$. For a sequence of hypergraphs $(\mathcal{H}_n)_{n \in \mathbb{N}}$, to show that w.h.p. $\mathcal{H}_{n, p}^v$ is r -Ramsey, it is equivalent to show that with probability tending to 0 the vertex set of $\mathcal{H}_{n, p}^v$ can be partitioned into r parts such that each part induces an independent set, i.e., it admits an r -partition $I \in \mathcal{I}_r(\mathcal{H}_n)$. Bounding this probability by using a union bound over all r -tuples in $\mathcal{I}_r(\mathcal{H}_n)$ does not work because $|\mathcal{I}_r(\mathcal{H}_n)|$ is too large. The power of a hypergraph container result is that one can instead use a union bound over all ‘containers’, since informally such a result says that every independent set lies in a container, and there are not too many containers overall.

8.1.2. *Deducing the 1-statement of Theorem 7.7 from Theorem 8.1.*

Proof of the 1-statement of Theorem 7.7. Let $(\mathcal{H}_n)_{n \in \mathbb{N}}$ be a sequence of k -uniform ordered hypergraphs satisfying (P1)–(P4). For all $W \subseteq [k]$ with $|W| \geq 2$ we have

$$\hat{p}_n = \max_{\substack{X \subseteq [n] \\ |X| \geq 2}} \left(\frac{e(\mathcal{H}_{n, X})}{v(\mathcal{H}_n)} \right)^{\frac{-1}{|X|-1}} \geq \left(\frac{e(\mathcal{H}_{n, W})}{v(\mathcal{H}_n)} \right)^{\frac{-1}{|W|-1}};$$

$$(20) \quad \implies e(\mathcal{H}_{n, W}) \geq (\hat{p}_n)^{-(|W|-1)} \cdot v(\mathcal{H}_n).$$

Therefore, for such $W \subseteq [k]$ with $|W| \geq 2$, and each $n \in \mathbb{N}$ sufficiently large, we have

$$(21) \quad \Delta_W(\mathcal{H}_n) \stackrel{(P3)}{\leq} b \frac{e(\mathcal{H}_n)}{e(\mathcal{H}_{n, W})} \stackrel{(20)}{\leq} b \cdot (\hat{p}_n)^{|W|-1} \cdot \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

For any $W \subseteq [k]$ with $|W| = 1$, and each $n \in \mathbb{N}$ sufficiently large, we have

$$(22) \quad \Delta_W(\mathcal{H}_n) \stackrel{(P4)}{\leq} b \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} = b \cdot (\hat{p}_n)^{|W|-1} \cdot \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

For each $n \in \mathbb{N}$, let \mathcal{H}'_n be the unordered restriction of \mathcal{H}_n . We now check the hypothesis of Theorem 8.1 with respect to $(\mathcal{H}'_n)_{n \in \mathbb{N}}$ and $(p_n)_{n \in \mathbb{N}}$ where $p_n := \hat{p}_n$. (P1) implies $p_n \rightarrow 0$ and $p_n v(\mathcal{H}'_n) \rightarrow \infty$ as $n \rightarrow \infty$. Remark 7.5 and (P2) imply $(\mathcal{H}'_n)_{n \in \mathbb{N}}$ is r -supersaturated. Finally, equations (21), (22) and (16) yield the required upper bound on $\Delta_\ell(\mathcal{H}'_n)$ where $c := b \cdot k!$.

Now, by Remark 7.5, the conclusion of Theorem 8.1 implies the 1-statement of Theorem 7.7. \square

8.1.3. *Proof of Theorem 8.1.* First, we introduce some notation. Let \mathcal{H} be a hypergraph and let $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$. We say \mathcal{F} is *increasing* if $A \in \mathcal{F}$ and $A \subseteq B \subseteq V(\mathcal{H})$ imply $B \in \mathcal{F}$. We write $\overline{\mathcal{F}}$ to denote the complement family of \mathcal{F} , i.e., the family of sets $A \in \mathcal{P}(V(\mathcal{H}))$ which are not in \mathcal{F} . Finally, for $\varepsilon \in (0, 1]$, we say that \mathcal{H} is $(\mathcal{F}, \varepsilon)$ -dense if $e(\mathcal{H}[A]) \geq \varepsilon e(\mathcal{H})$ for every $A \in \mathcal{F}$.

For this proof, when we consider r -tuples of sets, the r -tuples are *always* ordered. For two r -tuples of sets $I = (I_1, \dots, I_r)$ and $J = (J_1, \dots, J_r)$ we write $I \subseteq J$ if $I_i \subseteq J_i$ for all $i \in [r]$; we define $I \cup J := (I_1 \cup J_1, \dots, I_r \cup J_r)$. 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 $i \in [r]$.

We now state the following container result from [32], from which Theorem 8.1 will follow. Note that Proposition 8.2 itself follows from the general hypergraph container theorem of Balogh, Morris and Samotij [4, Theorem 2.2].

Proposition 8.2. [32, Proposition 3.2] *For every $k, r \in \mathbb{N}$ and all $c, \varepsilon > 0$, there exists $D > 0$ 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 $f : \mathcal{S}_r \rightarrow (\overline{\mathcal{F}})^r$ and $g : \mathcal{I}_r(\mathcal{H}) \rightarrow \mathcal{S}_r$ such that the following conditions hold:

- (1) *If $(S_1, \dots, S_r) \in \mathcal{S}_r$ then $\sum_{i=1}^r |S_i| \leq Dp \cdot v(\mathcal{H})$;*
- (2) *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 of Theorem 8.1. Note that it suffices to prove the theorem under the additional assumption that $c \geq k$. Pick $\varepsilon > 0$ so that, for all sufficiently large $n \in \mathbb{N}$, \mathcal{H}_n is (r, ε) -supersaturated. Define

$$(23) \quad \varepsilon' := \frac{\varepsilon}{3}, \quad \varepsilon'' := \frac{\varepsilon' k}{c} \quad \text{and} \quad \delta := \frac{\varepsilon}{3c}.$$

As $c \geq k$ we have that $\varepsilon'' \leq \varepsilon'$. Apply Proposition 8.2 with ε'' playing the role of ε to obtain $D > 0$. Since $p_n \rightarrow 0$ as $n \rightarrow \infty$, if n is sufficiently large then

$$(24) \quad cDp_n \leq \varepsilon'.$$

Next, we pick a sufficiently large constant $C > 0$ so that (\dagger) in (25) below holds for all sufficiently large $n \in \mathbb{N}$.

Let $n \in \mathbb{N}$ be sufficiently large and set $\mathcal{F}_n := \{F \subseteq V(\mathcal{H}_n) : e(\mathcal{H}_n[F]) \geq \varepsilon' e(\mathcal{H}_n)\}$. Clearly \mathcal{F}_n is increasing and \mathcal{H}_n is $(\mathcal{F}_n, \varepsilon')$ -dense, and thus also $(\mathcal{F}_n, \varepsilon'')$ -dense. Further, for all $F \in \mathcal{F}_n$, we have¹⁰

$$k\varepsilon' e(\mathcal{H}_n) \leq k e(\mathcal{H}_n[F]) = \sum_{y \in F} \deg_{\mathcal{H}_n[F]}(y) \leq |F| \Delta_1(\mathcal{H}_n) \leq \frac{|F| c \cdot e(\mathcal{H}_n)}{v(\mathcal{H}_n)},$$

where the last inequality uses the bound on $\Delta_1(\mathcal{H}_n)$ in the statement of Theorem 8.1. Rearranging, we obtain

$$|F| \geq \frac{\varepsilon' k}{c} v(\mathcal{H}_n) \stackrel{(23)}{=} \varepsilon'' v(\mathcal{H}_n),$$

¹⁰Here $\deg_{\mathcal{H}_n[F]}(y)$ denotes the number of edges y lies in within $\mathcal{H}_n[F]$.

for all $F \in \mathcal{F}_n$. Thus \mathcal{H}_n and \mathcal{F}_n satisfy the hypothesis of Proposition 8.2 with ε'' and p_n playing the roles of ε and p respectively. Apply Proposition 8.2 to obtain the family $\mathcal{S}_r \subseteq \mathcal{I}_r(\mathcal{H}_n)$ and functions $f : \mathcal{S}_r \rightarrow (\overline{\mathcal{F}}_n)^r$ and $g : \mathcal{I}_r(\mathcal{H}_n) \rightarrow \mathcal{S}_r$.

Now let $q_n \geq Cp_n$ and let B_{q_n} be the event that \mathcal{H}_{n,q_n}^v is not r -Ramsey. It suffices to show that $\mathbb{P}[B_{q_n}] \rightarrow 0$ as $n \rightarrow \infty$. Since $q_n \leq q'_n \leq 1$ implies $\mathbb{P}[B_{q'_n}] \leq \mathbb{P}[B_{q_n}]$, we may assume that $q_n = Cp_n$.

Suppose that \mathcal{H}_{n,q_n}^v is not r -Ramsey, i.e., there exists a partition $V(\mathcal{H}_{n,q_n}^v) = X_1 \cup \dots \cup X_r$ with $(X_1, \dots, X_r) \in \mathcal{I}_r(\mathcal{H}_n)$. By Proposition 8.2, there exists $S := (S_1, \dots, S_r) := g(X_1, \dots, X_r) \in \mathcal{S}_r$ such that $S \subseteq (X_1, \dots, X_r) \subseteq S \cup f(S)$ and $|\cup_{i \in [r]} S_i| \leq Dp_nv(\mathcal{H}_n)$. For brevity, we write $f(S) =: (f(S_1), \dots, f(S_r))$, $X(S) := \cup_{i \in [r]} S_i$, $Y(S) := \cup_{i \in [r]} f(S_i)$ and $Z(S) := X(S) \cup Y(S)$.¹¹

Claim 8.3. *We have $|Z(S)| \leq (1 - \delta)v(\mathcal{H}_n)$.*

Proof of the claim. Since \mathcal{H}_n is (r, ε) -supersaturated, in any partition $V(\mathcal{H}_n) = V_1 \cup \dots \cup V_r$ there exists $i \in [r]$ such that $e(\mathcal{H}_n[V_i]) \geq \varepsilon e(\mathcal{H}_n)$. Any vertex in \mathcal{H}_n is contained in at most $\Delta_1(\mathcal{H}_n)$ edges, and thus by deleting a set U of fewer than $\delta v(\mathcal{H}_n)$ vertices, we delete at most $|U|\Delta_1(\mathcal{H}_n) < \delta ce(\mathcal{H}_n)$ edges. Thus, given any induced subhypergraph of \mathcal{H}_n with more than $(1 - \delta)v(\mathcal{H}_n)$ vertices, for any partition $W_1 \cup \dots \cup W_r$ of it, there exists $i \in [r]$ such that $e(\mathcal{H}_n[W_i]) > (\varepsilon - \delta c)e(\mathcal{H}_n) = 2\varepsilon'e(\mathcal{H}_n)$ edges (where the last equality follows from (23)).

Now we bound $e(\mathcal{H}_n[S_i \cup f(S_i)])$ for each $i \in [r]$. Since $f(S_i) \in \overline{\mathcal{F}}_n$ we have $e(\mathcal{H}_n[f(S_i)]) < \varepsilon'e(\mathcal{H}_n)$. The number of edges in $\mathcal{H}_n[S_i \cup f(S_i)]$ containing at least one vertex from S_i is at most

$$|S_i| \cdot \Delta_1(\mathcal{H}_n) \leq Dp_nv(\mathcal{H}_n) \cdot c \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \stackrel{(24)}{\leq} \varepsilon'e(\mathcal{H}_n).$$

Thus, in total we have $e(\mathcal{H}_n[S_i \cup f(S_i)]) < 2\varepsilon'e(\mathcal{H}_n)$. Overall we see that the set $Z(S)$ can be partitioned into r sets so that each set induces fewer than $2\varepsilon'e(\mathcal{H}_n)$ edges in \mathcal{H}_n . Combining with the conclusion of the last paragraph, this implies that $|Z(S)| \leq (1 - \delta)v(\mathcal{H}_n)$ as required. \blacksquare

As $X_1 \cup \dots \cup X_r$ is a partition of $V(\mathcal{H}_{n,q_n}^v)$, we must have that $X(S) \subseteq V(\mathcal{H}_{n,q_n}^v) \subseteq Z(S)$. Crucially, this event depends uniquely on S and thus

$$\mathbb{P}[B_{q_n}] \leq \sum_{S \in \mathcal{S}_r} \mathbb{P}[X(S) \subseteq V(\mathcal{H}_{n,q_n}^v) \subseteq Z(S)] \leq \sum_{S \in \mathcal{S}_r} q_n^{|X(S)|} (1 - q_n)^{\delta v(\mathcal{H}_n)} \leq \sum_{S \in \mathcal{S}_r} q_n^{|X(S)|} e^{-\delta q_n v(\mathcal{H}_n)},$$

where the first inequality follows from the union bound, while the second inequality follows from Claim 8.3.

Recall that for every $S = (S_1, \dots, S_r) \in \mathcal{S}_r$ we have $|\cup_{i \in [r]} S_i| \leq Dp_nv(\mathcal{H}_n)$. We may bound $|\mathcal{S}_r|$ from above as follows: choose a set of $i \leq Dp_nv(\mathcal{H}_n)$ vertices from $V(\mathcal{H}_n)$ and then take an

¹¹Note the slight abuse of the f notation here.

(ordered) partition of these vertices into r classes. Therefore we have,

$$\begin{aligned}
\mathbb{P}[B_{q_n}] &\leq \sum_{S \in \mathcal{S}_r} q_n^{|X(S)|} e^{-\delta q_n v(\mathcal{H}_n)} \leq \sum_{i=0}^{D p_n v(\mathcal{H}_n)} \binom{v(\mathcal{H}_n)}{i} r^i q_n^i e^{-\delta q_n v(\mathcal{H}_n)} \\
&\leq (D p_n v(\mathcal{H}_n) + 1) (r q_n)^{D p_n v(\mathcal{H}_n)} \binom{v(\mathcal{H}_n)}{D p_n v(\mathcal{H}_n)} e^{-\delta q_n v(\mathcal{H}_n)} \\
&\leq (D p_n v(\mathcal{H}_n) + 1) \left(\frac{r q_n e v(\mathcal{H}_n)}{D p_n v(\mathcal{H}_n)} \right)^{D p_n v(\mathcal{H}_n)} e^{-\delta q_n v(\mathcal{H}_n)} \\
&= (D p_n v(\mathcal{H}_n) + 1) \left(\frac{r e C}{D} \right)^{\frac{D}{C} q_n v(\mathcal{H}_n)} e^{-\delta q_n v(\mathcal{H}_n)} \\
(25) \quad &\stackrel{(\dagger)}{\leq} e^{-\frac{\delta}{2} q_n v(\mathcal{H}_n)}.
\end{aligned}$$

Note that the equality follows since $q_n = C p_n$. Now using that $p_n v(\mathcal{H}_n)$ and hence $q_n v(\mathcal{H}_n)$ tends to infinity as $n \rightarrow \infty$, we have $\mathbb{P}[B_{q_n}] \rightarrow 0$ as $n \rightarrow \infty$, as required. \square

8.1.4. A remark on container theorems for solutions to systems of linear equations in abelian groups.

Note that we could have used Proposition 8.2 to obtain a container theorem for solutions to systems of linear equations in abelian groups, and then used this directly to prove the 1-statement of Lemma 6.8. Saxton and Thomason also gave a container theorem in this setting; see Theorem 4.2 in [54]. However, we do not use their result, as they give an alternative definition of $m_G(A)$ which we now state. (See Definitions 1.6 and 4.1 of their paper.) First define an $\ell \times k$ matrix A to have *full image* if given any $b \in G^\ell$, there exists $x \in G^k$ such that $Ax = b$. Now if G is a finite abelian group which is not a field, define $m'_G(A) := \frac{\ell+t-1}{t-1}$, where t is the maximum j for which $A_{\overline{W}}$ has full image for all $W \subseteq [k]$ with $|W| = j$.¹²

One can show that $m_G(A) \leq m'_G(A)$ for all finite abelian groups G and full image matrices A such that (A, G) is abundant. Further, there exist examples for which this inequality is strict. For example, consider the matrix

$$A := \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

together with any finite abelian group G which is not a field, for example \mathbb{Z}_6 . We have that (A, G) is abundant and in particular $t \geq 2$. By deleting the first three columns we obtain a submatrix which does not have full image, and so $t = 2$ and $m'_G(A) = 3$. However, under our definition, it is easy to compute that $m_G(A) = 2$. (Indeed, we have that $\text{rank}_G(A_{\overline{W}})$ is equal to the number of distinct columns in \overline{W} , except for when there are 3 distinct columns, in which case it equals 2.) Consequently, using Theorem 4.2 from [54] for such an example would lead to a suboptimal 1-statement.

8.2. Proof of the 0-statement of Theorem 7.7. The following proof adapts an argument of the second and third authors [33] (which in turn builds on work from [50]). Suppose $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is a sequence of k -uniform ordered hypergraphs that satisfies conditions (P1) and (P3)–(P5) of Theorem 7.7. In particular, (P5) yields a sequence $(X_n)_{n \in \mathbb{N}}$.

Let $c > 0$ be sufficiently small and pick $q_n \leq c \hat{p}(\mathcal{H}_n)$. Our aim is to show that w.h.p. there is a red-blue colouring of the vertices of $\mathcal{H}_{n, X_n, q_n}^v$ so that there are no monochromatic edges. By

¹²Note that Saxton and Thomason only consider G for which $t \geq 2$, and so $m'_G(A)$ is well-defined in this case.

definition of $\mathcal{H}_{n, X_n, q_n}^v$, such a vertex colouring also does not produce a monochromatic edge in \mathcal{H}_{n, q_n}^v , so this would prove the 0-statement of Theorem 7.7. We may further assume that $q_n = c\hat{p}(\mathcal{H}_n)$.

We say an (ordered) hypergraph is *Rado* if it has the property that however its vertices are red-blue coloured, there is always a monochromatic edge.¹³ Furthermore, a Rado hypergraph is *Rado minimal* if it is no longer Rado under the deletion of any edge. If $\mathcal{H}_{n, X_n, q_n}^v$ is Rado, fix a spanning subhypergraph H of $\mathcal{H}_{n, X_n, q_n}^v$ which is Rado minimal. Otherwise set $H := \emptyset$. So it suffices to prove that w.h.p. $H = \emptyset$. We will do this by combining two lemmas;

- A *deterministic lemma*, which states that if H is Rado then it must contain one of a finite list of subhypergraphs.
- A *probabilistic lemma*, which states that w.h.p. $\mathcal{H}_{n, X_n, q_n}^v$, and hence H , does not contain any of these subhypergraphs.

Before we state these two lemmas, we provide some motivation and state some required definitions.

The first step is to build some structure in H . The following claim is an analogue of Proposition 7.4 from [50] and Claim 1 from [33]. The proof of the claim follows in the same manner.

Claim 8.4. *Suppose H is non-empty, i.e., H is a Rado minimal ordered hypergraph. Then for every edge a of H and every vertex $v \in a$, there exists an edge b in H such that $a \cap b = v$.*

Proof. Let a be an edge, and let $v \in a$ be such that for all edges b such that $v \in b$, there exists another vertex $w \in a$ such that $w \in b$. Since H is Rado minimal, it is possible to red-blue colour the vertices of $H - a$ so that there are no red or blue edges. Thus once we add a back, it must be the case that a is (w.l.o.g.) red since H is Rado. But then change the colour of v to blue. If there is a monochromatic edge now, it must be a blue edge which contains v . However all edges containing v also contain another vertex from a which is red, thus we obtain a contradiction. \square

Using the above claim, one can build up more and more structure in H . Eventually, one can obtain enough structure so that H contains one of the subhypergraphs listed in Lemma 8.5 below. We now define some hypergraph notation in order to state it. In what follows, t is a positive integer. Further, in these definitions we just view edges as sets of vertices (i.e., we are not concerned about the order of an edge).

- A *simple path of length t* consists of edges e_1, \dots, e_t such that $|e_i \cap e_j| = 1$ if $j = i + 1$, and $|e_i \cap e_j| = 0$ if $j > i + 1$.
- A *fairly simple cycle of length t* ($t \geq 2$) consists of a simple path e_1, \dots, e_t and an edge e_0 such that $|e_0 \cap e_1| = 1$; $|e_0 \cap e_i| = 0$ for $2 \leq i \leq t - 1$; $|e_0 \cap e_t| = s \geq 1$.
- A *simple cycle* is a fairly simple cycle with $s = 1$.
- A simple path $P = e_1, \dots, e_t$ in H is called *spoiled* if there exists an edge $e^* \in E(H)$ such that $e^* \notin E(P)$, $e^* \subseteq V(P)$ and $|e^* \cap e_1| = 1$ where the vertex $e^* \cap e_1$ does not lie in e_2 .
- A subhypergraph H_0 of H is said to have a *handle* if there is an edge e^* in H such that $|e^*| > |e^* \cap V(H_0)| \geq 2$.
- A *faulty simple path of length t* ($t \geq 3$) is a simple path e_1, \dots, e_t together with two edges e_x and e_z such that e_1, e_2, e_x form a simple cycle with $|e_x \cap e_i| = 0$ for $i \geq 3$; e_{t-1}, e_t, e_z form a simple cycle with $|e_z \cap e_i| = 0$ for $i \leq t - 2$; each edge has size 3; the edges e_x and e_z may or may not be disjoint.
- A *bad triple* is a set of three edges e_1, e_x, e_y , where $e_1 \cap e_x = \{x\}$, $e_1 \cap e_y = \{y\}$, $x \neq y$, and $|e_x \cap e_y| \geq 2$.
- A *bad tight path* is a set of three edges e_1, e_2, e_3 each of size 3 such that $|e_1 \cap e_2| = 2$, $|e_1 \cap e_3| = 1$ and $|e_2 \cap e_3| = 2$.

We also require the following definition in the proof of Lemma 8.5.

¹³So being Rado is equivalent to being 2-Ramsey.

- A *Pasch configuration* is a set of four edges e_1, e_2, e_3, e_4 of size 3 such that $v_{ij} = e_i \cap e_j$ is a distinct vertex for each pair $i < j$.

Write $N := n \cdot q_n = \mathbb{E}(v(\mathcal{H}_{n, X_n, q_n}^v))$ for brevity.

Lemma 8.5 (Deterministic lemma). *If H is non-empty then it contains at least one of the following structures:*

- (i) *A simple path of length at least $\log N$.*
- (ii) *A spoiled simple path of length at most $\log N$.*
- (iii) *A fairly simple cycle of length at most $\log N$, with a handle.*
- (iv) *A faulty simple path of length at most $\log N$.*
- (v) *A bad triple.*
- (vi) *A bad tight path.*

Lemma 8.6 (Probabilistic lemma). *If $c > 0$ is sufficiently small then w.h.p. $\mathcal{H}_{n, X_n, q_n}^v$ does not contain any of the hypergraphs (i)–(vi) listed in Lemma 8.5.*

These two results immediately prove the 0-statement of Theorem 7.7.

8.2.1. *Proof of Lemma 8.5.* Suppose for a contradiction that H is non-empty but does not contain any of the structures defined in (i)–(vi). Note that H is an $|X_n|$ -uniform ordered hypergraph, and $|X_n| \geq 3$ by (P5).

Let $P = e_1, \dots, e_t$ be the longest simple path in H . As we have no simple paths of length at least $\log N$, we have $t \leq \log N$. By Claim 8.4, $t \geq 2$. Let x, y be two vertices which belong only to e_1 in P , and let e_x and e_y be two edges of H such that $e_z \cap e_1 = \{z\}$ for $z = x, y$ (such e_x, e_y exist by Claim 8.4). By the maximality of P , we have $h_z := |V(P) \cap e_z| \geq 2$ for $z = x, y$.

If $h_z = |X_n|$ for some $z = x, y$, then P together with e_z is a spoiled simple path, a contradiction. Otherwise, let $i_z := \min\{i \geq 2 : e_z \cap e_i \neq \emptyset\}$ for $z = x, y$, and assume without loss of generality that $i_y \leq i_x$. We know e_1, \dots, e_{i_x}, e_x must not form a fairly simple cycle for which e_y is a handle. This implies $e_y \subseteq e_1 \cup \dots \cup e_{i_x} \cup e_x$. In particular, this means e_x must contain all those vertices in e_y which do not lie on P . If $|e_y \cap e_x| \geq 2$ then e_1, e_x and e_y form a bad triple, a contradiction. Thus, $e_y \cap e_x$ consists of precisely one vertex v_{xy} (and v_{xy} lies outside of P). Now consider e_1, \dots, e_{i_y}, e_y . This is a fairly simple cycle that e_x intersects in at least two vertices (namely x and v_{xy}). Thus, we obtain a fairly simple cycle with a handle unless all the vertices in e_x lie in e_1, \dots, e_{i_y}, e_y . In particular, $e_x \subseteq (e_1 \cup e_{i_y} \cup e_y)$ as $i_y \leq i_x$. This in turn implies $e_{i_y} = e_{i_x}$. Indeed, otherwise e_x must contain one vertex from e_1 and $|X_n| - 1 \geq 2$ vertices from e_y , a contradiction as we already observed that e_x only intersects e_y in one vertex.

In summary, we have that $i_x = i_y$ and e_x and e_y intersect in a single vertex v_{xy} outside of P . As mentioned in the last paragraph, we must have $e_x \subseteq (e_1 \cup e_{i_y} \cup e_y)$. Similarly, we have that e_1, \dots, e_{i_x}, e_x form a fairly simple cycle for which e_y is a handle (a contradiction) unless $e_y \subseteq (e_1 \cup e_{i_x} \cup e_x)$.

The two previous inequalities imply $e_x \setminus \{x, v_{xy}\} \subseteq e_{i_y}$ and $e_y \setminus \{y, v_{xy}\} \subseteq e_{i_x}$. Recall here that $e_{i_x} = e_{i_y}$. We must have that $|e_{i_x} \setminus (e_x \cup e_y)| \geq 1$; indeed, otherwise either e_x or e_y contains the vertex $e_{i_x} \cap e_{i_x-1}$, which either contradicts the minimality of i_x or i_y . Note that $e_x \setminus \{x, v_{xy}\}$ and $e_y \setminus \{y, v_{xy}\}$ are disjoint, they have size $|X_n| - 2$ each, and their union is a strict subset of e_{i_x} . Thus $2(|X_n| - 2) \leq |e_{i_x}| - 1 = |X_n| - 1$. This implies $|X_n| \leq 3$, and so $|X_n| = 3$.

If $i_x \geq 3$ then e_1, e_x, e_y form a fairly simple cycle for which e_{i_x} is a handle (since $|e_{i_x} \setminus (e_x \cup e_y)| \geq 1$ and $e_1 \cap e_{i_x} = \emptyset$), a contradiction. Thus we have that $i_x = 2$, and so e_1, e_x, e_y, e_{i_x} form a Pasch configuration.

Now repeat the maximal path process which we did for e_1 to find e_x and e_y , except from the other end of the path. That is, there must exist edges e_z and e_w such that $e_z \cap e_t = \{z\}$, $e_w \cap e_t = \{w\}$, where z, w are vertices in e_t that are not in e_{t-1} . By repeating the previous case analysis, we arrive

at the conclusion that e_{t-1}, e_t, e_z, e_w must also form a Pasch configuration where $e_z \cap e_w$ is a vertex v_{zw} outside of P . If $t \geq 3$, then $e_1, \dots, e_t, e_x, e_z$ together form a faulty simple path. Hence we must have $t = 2$.

If the union of these two Pasch configurations contains 7 vertices (i.e., $v_{xy} \neq v_{zw}$), then e_1, e_2, e_x form a fairly simple cycle for which e_z is a handle. So we now suppose that the two Pasch configurations cover the same 6 vertices. If we do not have $\{e_x, e_y\} = \{e_z, e_w\}$ then e_x, e_z, e_y form a bad tight path. Hence we do have equality and the two Pasch configurations we found are identical.

Relabel the edges and vertices as in the definition of a Pasch configuration. We observe that H cannot be just these four edges: such a hypergraph is not Rado, e.g., colour v_{12}, v_{13}, v_{34} red, and the remaining vertices blue. Also, by definition of Rado minimal, this cannot be a component of H . That is, there is an edge e_5 in H , where $e_5 \neq e_i$ for $i \in [4]$, and e_5 contains s vertices from inside the Pasch configuration, where $s \geq 1$. If $s = 1$ then w.l.o.g. e_5 contains $v_{1,2}$; then e_5, e_1, e_3 is a simple path of length 3, a contradiction to the longest path in H of length 2 found earlier. If $s = 2$ then whichever two vertices of the Pasch configuration e_5 contains, taking any of the simple cycles of the Pasch configuration together with e_5 gives a (fairly) simple cycle with a handle. If $s = 3$, first suppose $e_5 = \{v_{1,2}, v_{1,3}, v_{2,3}\}$. Then e_1, e_5, e_2 is a bad tight path. If $e_5 = \{v_{1,2}, v_{1,3}, v_{2,4}\}$, then again e_1, e_5, e_2 is a bad tight path. For all other 3-sets of vertices e_5 could contain, a symmetrical argument shows that we find a bad tight path. Since all three values of s give a contradiction, this concludes the proof. \square

8.2.2. *Proof of Lemma 8.6.* Let $c > 0$ be sufficiently small and $n \in \mathbb{N}$ be sufficiently large. Recall that $q_n = c\hat{p}_n(\mathcal{H}_n)$. Given an $|X_n|$ -uniform ordered hypergraph G , we will obtain an upper bound on the expected number of copies of G in $\mathcal{H}_{n, X_n, q_n}^v$ by obtaining upper bounds on the expected number of possible assignments of vertices $v \in V(\mathcal{H}_{n, X_n, q_n}^v)$ to an edge in which some vertices are already fixed, and others are to be assigned. To make this precise, we need some notation. Given an edge order e_1, \dots, e_t of $E(G)$, we call a vertex v *new in e_i* if $v \in e_i$ but $v \notin e_j$ for all $j < i$. Otherwise we call $v \in e_i$ *old in e_i* . Clearly (provided G has no isolated vertices) each vertex in G is new in one edge, and old in any subsequent edge that it appears in. Now, for $0 \leq w \leq |X_n| - 1$, we will shortly provide an upper bound P_w on the expected number of edges in $\mathcal{H}_{n, X_n, q_n}^v$ that contain a fixed set of w (old) vertices; crucially, P_w is independent of the exact set of w vertices considered. Then one can obtain an upper bound on the expected number of copies of G in $\mathcal{H}_{n, X_n, q_n}^v$ directly as some function of the P_w 's (which will depend on G). We first obtain bounds for these P_w 's.

Given a fixed set of w vertices in $\mathcal{H}_{n, X_n, q_n}^v$, there are $\binom{|X_n|}{w} w! \leq |X_n|! \leq k!$ ways of selecting some $W \subseteq X_n$ with $|W| = w$ and picking an ordering v_1, \dots, v_w of these w vertices. Then, there are at most $\Delta_W(\mathcal{H}_{n, X_n})$ choices for an edge e in \mathcal{H}_{n, X_n} such that $e_W = (v_1, \dots, v_w)$. Each vertex in \mathcal{H}_{n, X_n} is included with probability q_n in $\mathcal{H}_{n, X_n, q_n}^v$, thus we obtain the upper bound

$$P_w = (k!) \cdot \left(\max_{W \subseteq X_n, |W|=w} \Delta_W(\mathcal{H}_{n, X_n}) \right) \cdot (q_n)^{|X_n|-w}.$$

For $w = 0$, we will use the bound

$$P_0 \leq e(\mathcal{H}_{n, X_n}) \cdot (q_n)^{|X_n|} \leq n^{|X_n|} \cdot (q_n)^{|X_n|} \leq N^k.$$

We now use conditions (P3), (P4) and (P5) to bound P_w further for values of w greater than 0. First, condition (P4) implies that,

$$P_1 \leq (k!) \cdot b \cdot \frac{e(\mathcal{H}_{n, X_n})}{v(\mathcal{H}_n)} \cdot (q_n)^{|X_n|-1}.$$

By (18) and since $n \in \mathbb{N}$ is sufficiently large, we have that $f_{n, X_n} \geq \eta \hat{p}_n(\mathcal{H}_n)$ for some constant $\eta > 0$. It follows from the above that

$$\begin{aligned} P_1 &\leq (k!) \cdot b \cdot \frac{e(\mathcal{H}_{n, X_n})}{v(\mathcal{H}_n)} \cdot (c \hat{p}_n(\mathcal{H}_n))^{|X_n|-1} \\ &\stackrel{(18)}{\leq} (k!) \cdot b \cdot \frac{e(\mathcal{H}_{n, X_n})}{v(\mathcal{H}_n)} \cdot (f_{n, X_n})^{|X_n|-1} \cdot (c/\eta)^{|X_n|-1} \\ &\stackrel{(17)}{=} (k!) \cdot b \cdot (c/\eta)^{|X_n|-1} \leq \sqrt{c}, \end{aligned}$$

where the last inequality follows from $c > 0$ being sufficiently small.

Similar calculations hold for higher values of w . Recall we have $q_n = c \hat{p}_n \geq c f_{n, W}$ for every $W \subseteq [k]$ with $|W| \geq 2$. Let $2 \leq w \leq |X_n| - 1$ and pick $W \subset X_n$ with $|W| = w$ maximising $\Delta_W(\mathcal{H}_{n, X_n})$. As $n \in \mathbb{N}$ is sufficiently large we have

$$\begin{aligned} P_w &\stackrel{(P3)}{\leq} (k!) \cdot b \cdot \frac{e(\mathcal{H}_{n, X_n})}{e(\mathcal{H}_{n, W})} \cdot (q_n)^{|X_n|-w} = (k!) \cdot b \cdot \frac{e(\mathcal{H}_{n, X_n})}{v(\mathcal{H}_n)} \cdot (q_n)^{|X_n|-1} \cdot \frac{v(\mathcal{H}_n)}{e(\mathcal{H}_{n, W})} \cdot (q_n)^{-(w-1)} \\ &\leq \left((k!) \cdot b \cdot (c/\eta)^{|X_n|-1} \right) \cdot \frac{v(\mathcal{H}_n)}{e(\mathcal{H}_{n, W})} \cdot (c f_{n, W})^{-(w-1)} \\ &\stackrel{(17)}{=} \left((k!) \cdot b \cdot (c/\eta)^{|X_n|-1} \right) \cdot c^{-(w-1)} \leq (k!) \cdot b \cdot c \cdot \eta^{-(|X_n|-1)} \leq \sqrt{c}, \end{aligned}$$

where the last two inequalities follow from $|X_n| > w$ and $c > 0$ being sufficiently small.

Finally, for any $2 \leq w \leq |X_n| - 1$, we have that $P_0 \cdot P_2 \cdot P_w$ is at most

$$\begin{aligned} &(k!)^2 \cdot e(\mathcal{H}_{n, X_n}) \cdot \left(\max_{W \subseteq X_n, |W|=2} \Delta_W(\mathcal{H}_{n, X_n}) \right) \cdot \left(\max_{W \subseteq X_n, |W|=w} \Delta_W(\mathcal{H}_{n, X_n}) \right) \cdot (q_n)^{3|X_n|-2-w} \\ &\stackrel{(19)}{=} o(1). \end{aligned}$$

For brevity, let $P_{\geq w} := \max_{i \geq w} P_i$. To recap, we have shown that

$$(26) \quad P_0 \leq N^k, \quad P_{\geq 1} \leq \sqrt{c} \quad \text{and} \quad P_0 \cdot P_2 \cdot P_{\geq 2} = o(1).$$

To prove the lemma it suffices to show that, for each set of hypergraphs stated in Lemma 8.5, the expected number of such structures in $\mathcal{H}_{n, X_n, q_n}^v$ is $o(1)$. Indeed, then by Markov's inequality w.h.p. $\mathcal{H}_{n, X_n, q_n}^v$ does not contain any of these structures. In each case, we will find an edge order and then use (26) to bound the expected number of these structures. Let G be an ordered hypergraph. Given an edge order e_1, \dots, e_t of $E(G)$, we call an edge e_i

- *initial* if all its vertices are new;
- *normal* if it has precisely one old vertex;
- *2-bad* if it has precisely two old vertices;
- *bad* if it has at least two and at most $|X_n| - 1$ old vertices.

Now observe the bounds P_0, P_1, P_2 and $P_{\geq 2}$ correspond to upper bounds on the expected number of ways one can place an initial, normal, 2-bad and bad edge in $\mathcal{H}_{n, X_n, q_n}^v$ respectively.

Case (i). Note that for this case it suffices to show that w.h.p. $\mathcal{H}_{n, X_n, q_n}^v$ contains no simple path of length exactly $s := \lceil \log N \rceil$; thus we only need to consider the expected number of simple paths of length s . Let e_1, \dots, e_s be the simple path. Here, e_1 is initial, and all other edges are normal. Thus, we can bound the expected number of simple paths of length s in $\mathcal{H}_{n, X_n, q_n}^v$ by

$$P_0 \cdot P_1^{s-1} \stackrel{(26)}{\leq} N^k \cdot (\sqrt{c})^{s-1} \leq (\sqrt{c})^{(\log N)-1} \cdot N^k \stackrel{(P1)}{=} o(1),$$

which follows since $c > 0$ is sufficiently small.

Case (ii). Let e_1, \dots, e_t be the simple path and let e^* be the spoiling edge. Choose i to be the smallest $i \geq 2$ such that $|e_i \cap e^*| \geq 1$. Consider the edge order $e^*, e_i, e_{i-1}, \dots, e_1, e_{i+1}, \dots, e_t$. Here e^* is initial. If $|e_i \cap e^*| \geq 2$ then e_i is bad, the e_j for $2 \leq j \leq i-1$ are normal, e_1 is 2-bad, and each e_j for $i+1 \leq j \leq t$ is normal or bad. If $|e_i \cap e^*| = 1$ then e_j for $2 \leq j \leq i$ are normal, e_1 is 2-bad, and each e_j for $i+1 \leq j \leq t$ is normal or bad. In particular, for at least one such j , e_j is bad, since e^* must intersect at least one of these edges (as $|X_n| \geq 3$). In either case, we have found an edge order with one initial edge, one 2-bad edge, at least one further bad edge, and all other edges normal or bad edges. We have $t \leq \log N$ and there are at most tk choices for the location on the simple path of each vertex of e^* intersecting it. Overall, the expected number of spoiled paths of length at most $\log N$ is at most

$$\sum_{t=2}^{\log N} (tk)^k \cdot P_0 \cdot P_{\geq 1}^{t-2} \cdot P_2 \cdot P_{\geq 2} \stackrel{(26)}{\leq} \sum_{t=2}^{\log N} (tk)^k \cdot (\sqrt{c})^{t-2} \cdot o(1) = o(1),$$

where the last inequality holds as $c > 0$ is sufficiently small.

Case (iii). Let e_1, \dots, e_t, e_0 be the fairly simple cycle and let e^* be the handle. Consider the edge order $e_0, e_t, \dots, e_1, e^*$. Here e_0 is initial, e_t is normal or bad, the e_j for $2 \leq j \leq t-1$ are normal, e_1 is 2-bad and e^* is bad. We have $t \leq \log N$ and there are at most tk choices for the location on the fairly simple cycle for the choice of where each of the u vertices of the handle e^* intersect it, where $2 \leq u \leq k-1$. Overall, the expected number of fairly simple cycles of length at most $\log N$ with a handle is at most

$$\sum_{t=2}^{\log N} \sum_{u=2}^{k-1} (tk)^u \cdot P_0 \cdot P_{\geq 1}^{t-1} \cdot P_2 \cdot P_{\geq 2} \stackrel{(26)}{\leq} \sum_{t=2}^{\log N} \sum_{u=2}^{k-1} (tk)^u \cdot (\sqrt{c})^{t-1} \cdot o(1) = o(1).$$

Case (iv). Let $e_1, \dots, e_t, e_x, e_z$ be the faulty simple path. Consider the edge order $e_1, e_x, e_2, \dots, e_{t-1}, e_z, e_t$. Here e_1 is initial, e_2 and e_t are 2-bad, e_z is normal or 2-bad (depending whether e_z intersects e_x), and all other edges are normal. We have $t \leq \log N$, and a choice of whether e_x and e_z intersect or not. Overall, we can bound the expected number of faulty simple paths of length at most $\log N$ by

$$\sum_{t=3}^{\log N} P_0 \cdot P_1^{t-1} \cdot P_2^2 + \sum_{t=3}^{\log N} P_0 \cdot P_1^{t-2} \cdot P_2^3 \stackrel{(26)}{\leq} 2 \sum_{t=3}^{\log N} (\sqrt{c})^{t-1} \cdot o(1) = o(1).$$

Case (v). Let e_1, e_x, e_y be the bad triple. Consider the edge order e_x, e_y, e_1 . Here, e_x is initial, e_y is bad and e_1 is 2-bad. We have a choice of the size of $t := |e_x \cap e_y|$. Overall, the expected number of bad triples is at most

$$\sum_{t=2}^{k-1} P_0 \cdot P_2 \cdot P_{\geq 2} \stackrel{(26)}{=} o(1).$$

Case (vi). Let e_1, e_2, e_3 be a bad tight path. In this edge order, e_1 is initial, and e_2 and e_3 are 2-bad. Overall, we can bound the expected number of bad tight paths by

$$P_0 \cdot P_2^2 \stackrel{(26)}{=} o(1).$$

□

9. PROOF OF THE RANDOM RADO LEMMA

Before we deduce the random Rado lemma (Lemma 6.8) from Theorem 7.7, we next prove Lemma 6.9.

Proof of Lemma 6.9. We first prove the second statement of the lemma.

Claim 9.1.

$$\lim_{n \rightarrow \infty} \frac{|k\text{-Sol}_{S_n}^A([k])|}{|\text{Sol}_{S_n}^A([k])|} = 1.$$

Proof of the claim. We bound from above the number of non k -distinct solutions in $\text{Sol}_{S_n}^A([k])$. For each such solution, we have $|S_n|$ choices for the repeated entry. Using conditions (A3) and (A6), for n sufficiently large, we obtain

$$\begin{aligned} \sum_{z \in S_n} \sum_{\substack{W \subseteq [k] \\ |W|=2}} |\text{Sol}_{S_n}^A((z, z), W, [k])| &\stackrel{(A3)}{\leq} B \sum_{z \in S_n} \sum_{\substack{W \subseteq [k] \\ |W|=2}} \frac{|\text{Sol}_{S_n}^A([k])|}{|\text{Sol}_{S_n}^A(W)|} \\ &= B|S_n| \sum_{\substack{W \subseteq [k] \\ |W|=2}} \frac{|\text{Sol}_{S_n}^A([k])|}{|\text{Sol}_{S_n}^A(W)|} \stackrel{(A6)}{=} o(|\text{Sol}_{S_n}^A([k])|). \end{aligned}$$

In other words, $|\text{Sol}_{S_n}^A([k])| - |k\text{-Sol}_{S_n}^A([k])| = o(|\text{Sol}_{S_n}^A([k])|)$ which immediately implies the claim. \blacksquare

Next, we prove the inequality stated in the lemma. Let $Y \subseteq [k]$. Note that $k\text{-Sol}_{S_n}^A(Y) \subseteq \text{Sol}_{S_n}^A(Y)$, so the upper bound follows trivially. By definition, if $y \in k\text{-Sol}_{S_n}^A(Y)$ then there exists some $x \in k\text{-Sol}_{S_n}^A([k])$ such that $x_Y = y$. Thus,

$$|k\text{-Sol}_{S_n}^A(Y)| \cdot \max_{y_0 \in S^{|Y|}} |\text{Sol}_{S_n}^A(y_0, Y, [k])| \geq |k\text{-Sol}_{S_n}^A([k])|.$$

As (A, S_n) is B -extendable, for n large the above implies

$$|k\text{-Sol}_{S_n}^A(Y)| \cdot \frac{B|\text{Sol}_{S_n}^A([k])|}{|\text{Sol}_{S_n}^A(Y)|} \geq |k\text{-Sol}_{S_n}^A([k])|.$$

Rearranging gives

$$\frac{|k\text{-Sol}_{S_n}^A(Y)|}{|\text{Sol}_{S_n}^A(Y)|} \geq \frac{|k\text{-Sol}_{S_n}^A([k])|}{B|\text{Sol}_{S_n}^A([k])|} \geq \frac{1}{2B},$$

where the last inequality holds for n sufficiently large by Claim 9.1. \square

Lemma 6.8 follows easily from Theorem 7.7. We use Lemma 6.9 to restrict our attention to k -distinct solutions.

Proof of Lemma 6.8. Let $r \geq 2$, $k \geq 3$ and $\ell \geq 1$ be integers. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of abelian groups and A be an $\ell \times k$ integer matrix. Suppose that conditions (A1)–(A6) from Lemma 6.8 hold. Let \mathcal{H}_n be the k -uniform ordered hypergraph with vertex set S_n , and edge set consisting of all k -distinct solutions $x = (x_1, \dots, x_k)$ to $Ax = 0$ in S_n .

Let $Y \subseteq [k]$. Observe that $e(\mathcal{H}_{n,Y}) = |k\text{-Sol}_{S_n}^A(Y)|$. As (A3) and (A6) hold, Lemma 6.9 implies that

$$(27) \quad e(\mathcal{H}_{n,Y}) = \Theta(|\text{Sol}_{S_n}^A(Y)|) \quad \text{and} \quad e(\mathcal{H}_n) \sim |\text{Sol}_{S_n}^A([k])|.$$

It follows that, for $|Y| \geq 2$,

$$(28) \quad f_{n,Y} = \left(\frac{e(\mathcal{H}_{n,Y})}{v(\mathcal{H}_n)} \right)^{-\frac{1}{|Y|-1}} = \Theta \left(\frac{|\text{Sol}_{S_n}^A(Y)|}{|S_n|} \right)^{-\frac{1}{|Y|-1}} = \Theta(p_Y(A, S_n))$$

and thus $\hat{p}(\mathcal{H}_n) = \Theta(\hat{p}(A, S_n))$. We now prove that conditions (P1)–(P5) hold.

(P1): Since $\hat{p}_n(\mathcal{H}_n) = \Theta(\hat{p}_n(A, S_n))$ and $|S_n| = v(\mathcal{H}_n)$, (A1) implies that (P1) holds.

(P2): Since $e(\mathcal{H}_n) = |k\text{-Sol}_{S_n}^A([k])| \sim |\text{Sol}_{S_n}^A([k])|$ by (27), (A2) implies that (P2) holds.

(P3): For any $W \subseteq Y \subseteq [k]$ we have

$$\Delta_W(\mathcal{H}_{n,Y}) \leq \max_{w_0 \in S_n^{|W|}} \{|\text{Sol}_{S_n}^A(w_0, W, Y)|\} \stackrel{(A3)}{\leq} \Theta \left(\frac{|\text{Sol}_{S_n}^A(Y)|}{|\text{Sol}_{S_n}^A(W)|} \right) \stackrel{(27)}{\leq} \Theta \left(\frac{|e(\mathcal{H}_{n,Y})|}{|e(\mathcal{H}_{n,W})|} \right),$$

and so (P3) holds.

(P4): Pick $W \subset Y \subseteq [k]$ with $|W| = 1$. We have

$$\Delta_W(\mathcal{H}_{n,Y}) \leq \max_{w_0 \in S_n} |\text{Sol}_{S_n}^A(w_0, W, Y)| \stackrel{(A4)}{=} O \left(\frac{|\text{Sol}_{S_n}^A(Y)|}{|S_n|} \right) \stackrel{(27)}{=} O \left(\frac{e(\mathcal{H}_{n,Y})}{v(\mathcal{H}_n)} \right).$$

(P5): Pick a sequence $(X_n)_{n \in \mathbb{N}}$ of subsets $X_n \subseteq [k]$ with $|X_n| \geq 3$ for all $n \in \mathbb{N}$ as in Definition 6.6.

Note that

$$f_{n,X_n} = \left(\frac{e(\mathcal{H}_{n,X_n})}{v(\mathcal{H}_n)} \right)^{-\frac{1}{|X_n|-1}} \stackrel{(28)}{=} \Theta \left(\frac{|\text{Sol}_{S_n}^A(X_n)|}{|S_n|} \right)^{-\frac{1}{|X_n|-1}} = \Theta(\hat{p}(A, S_n)) = \Theta(\hat{p}(\mathcal{H}_n)),$$

and so $f_{n,X_n} = \Omega(\hat{p}(\mathcal{H}_n))$. It remains to verify equation (19). Pick sequences $(W_n)_{n \in \mathbb{N}}$ and $(W'_n)_{n \in \mathbb{N}}$ with $W_n, W'_n \subset X_n$, $|W_n| = 2$ and $|W'_n| \geq 2$. Then

$$\begin{aligned} & \Delta_{W_n}(\mathcal{H}_{n,X_n}) \cdot \Delta_{W'_n}(\mathcal{H}_{n,X_n}) \cdot e(\mathcal{H}_{n,X_n}) \cdot \hat{p}(\mathcal{H}_n)^{3|X_n|-2-|W'_n|} \\ & \stackrel{(P3),(27)}{\leq} \Theta \left(\frac{|\text{Sol}_{S_n}^A(X_n)|}{|\text{Sol}_{S_n}^A(W_n)|} \cdot \frac{|\text{Sol}_{S_n}^A(X_n)|}{|\text{Sol}_{S_n}^A(W'_n)|} \cdot |\text{Sol}_{S_n}^A(X_n)| \cdot \left(\frac{|\text{Sol}_{S_n}^A(X_n)|}{|S_n|} \right)^{-\frac{3|X_n|-2-|W'_n|}{|X_n|-1}} \right) \\ & = \Theta \left(\frac{|S_n|^2}{|\text{Sol}_{S_n}^A(W_n)|} \cdot \frac{|S_n|}{|\text{Sol}_{S_n}^A(W'_n)|} \cdot \left(\frac{|\text{Sol}_{S_n}^A(X_n)|}{|S_n|} \right)^{\frac{|W'_n|-1}{|X_n|-1}} \right) \\ & = \Theta \left(\frac{|S_n|^2}{|\text{Sol}_{S_n}^A(W_n)|} \cdot \left(\frac{p_{W'_n}(A, S_n)}{p_{X_n}(A, S_n)} \right)^{|W'_n|-1} \right) \stackrel{(A5)}{=} o(1). \end{aligned}$$

As conditions (P1)–(P5) are satisfied, by Theorem 7.7 there exist constants $c, C > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}_{n,q_n}^v \text{ is } r\text{-Ramsey}] = \begin{cases} 0 & \text{if } q_n \leq c\hat{p}(\mathcal{H}_n); \\ 1 & \text{if } q_n \geq C\hat{p}(\mathcal{H}_n). \end{cases}$$

Note that \mathcal{H}_{n,q_n}^v is r -Ramsey if and only if S_{n,q_n} is (A, r) -Rado. Furthermore, we have $\hat{p}(\mathcal{H}_n) = \Theta(\hat{p}(A, S_n))$ and so there exist constants c_0, C_0 such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,q_n} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } q_n \leq c_0 \cdot \hat{p}(A, S_n); \\ 1 & \text{if } q_n \geq C_0 \cdot \hat{p}(A, S_n). \end{cases}$$

□

10. PROOF OF TWO AUXILIARY RESULTS

In this section, we prove two auxiliary results. Lemma 2.3 gives an upper bound on the number of k -APs in \mathbf{P}_n containing a fixed prime number and it is used in the proofs of Theorem 2.2 and Theorem 2.7. Theorem 3.9 is a supersaturation result for integer lattices which is used to prove Theorem 5.1.

10.1. Proof of Lemma 2.3. We make use of the following sieve result.

Lemma 10.1 ([60]). *Let $k \in \mathbb{N}$ with $k \geq 2$ and let q be a prime number. For each prime $p \geq k$ with $p \neq q$, let E_p be the union of $k - 1$ residue classes modulo p . Then $|E(n)| = O(n/\log^{k-1} n)$ where*

$$E(n) := [n] \setminus \bigcup_{\substack{p \neq q \text{ prime} \\ k \leq p \leq \sqrt{n}}} E_p.$$

Lemma 10.1 is a weakened version of Theorem 32 from [60]. A precise explanation of how to deduce Lemma 10.1 from the original statement in [60] can be found in Appendix A.1.

Proof of the Lemma 2.3. Let k, ℓ, n and q be as in the statement of Lemma 2.3. Let $\mathcal{A}(n)$ be the set of k -APs in \mathbb{Z} such that the ℓ th term of the progression is q and the common difference is an integer between 1 and n inclusive. In particular, we have $|\mathcal{A}(n)| = n$. Furthermore, the collection of k -APs in \mathbf{P}_n whose ℓ th element is q is a subset of $\mathcal{A}(n)$. Thus, it suffices to show that the number of k -APs in $\mathcal{A}(n)$ consisting entirely of primes is $O(n/\log^{k-1} n)$.

For every prime number $p \neq q$, let \mathcal{A}_p be the set of k -APs in $\mathcal{A}(n)$ with at least one term divisible by p . Consider the set

$$\mathcal{E}(n) := \mathcal{A}(n) \setminus \bigcup_{\substack{p \neq q \text{ prime} \\ k \leq p \leq \sqrt{n}}} \mathcal{A}_p.$$

In other words, $\mathcal{E}(n)$ is the set of all k -APs in $\mathcal{A}(n)$ such that no term of the progression is divisible by a prime between k and \sqrt{n} (except q). In particular, $\mathcal{E}(n)$ contains all k -APs in $\mathcal{A}(n)$ whose terms (other than perhaps q) are all primes greater than \sqrt{n} . Conversely, the number of k -APs in $\mathcal{A}(n)$ containing at least one prime $p' \leq \sqrt{n}$ where $p' \neq q$ is $o(n/\log^{k-1} n)$. This follows from (i) the fact that there are at most k arithmetic progressions of length k containing both q (in the ℓ th position) and a fixed prime $p \neq q$ and as (ii) $\mathbf{P}_{\sqrt{n}} \sim \frac{2\sqrt{n}}{\log n}$ by the prime number theorem.

By the observations above, the number of k -APs in $\mathcal{A}(n)$ consisting entirely of primes is at most $o(n/\log^{k-1} n) + |\mathcal{E}(n)|$; so it suffices to prove that $|\mathcal{E}(n)| = O(n/\log^{k-1} n)$.

Now we reformulate the problem in terms of common differences. Since the ℓ th term of all k -APs in $\mathcal{A}(n)$ is equal to q , each k -AP in $\mathcal{A}(n)$ is uniquely determined by its common difference. For every prime $p \neq q$, let E_p be the set of all common differences of k -APs in \mathcal{A}_p . It follows immediately that $|E_p| = |\mathcal{A}_p|$. Furthermore, the set of common differences of k -APs in $\mathcal{A}(n)$ is exactly $[n]$ by definition. Thus we have $|\mathcal{E}(n)| = |E(n)|$ where

$$E(n) := [n] \setminus \bigcup_{\substack{p \neq q \text{ prime} \\ k \leq p \leq \sqrt{n}}} E_p.$$

Applying Theorem 10.1 yields $|\mathcal{E}(n)| = |E(n)| = O(n/\log^{k-1} n)$, as required. The only assumption of Theorem 10.1 that needs to be checked is that, for every prime $p \geq k$ with $p \neq q$, E_p is the union of $k - 1$ residue classes modulo p .

Let $a \in [n]$ and $b_i := q + a(i - \ell)$ for every $i \in [k]$. Then, $a \in E_p$ if and only if (b_1, \dots, b_k) is a k -AP in \mathcal{A}_p , i.e., p divides some b_i . Now, p divides $b_i = q + a(i - \ell)$ if and only if $i \neq \ell$ and

$a \equiv -q(i - \ell)^{-1} \pmod{p}$. Note that $i - \ell \neq 0$ is always invertible modulo p since $|i - \ell| < k \leq p$. It follows that $E_p = \{a \in [n] : a \equiv -q(i - \ell)^{-1} \pmod{p}, i \in [k], i \neq \ell\}$. Hence, E_p is the set of precisely $k - 1$ residue classes modulo p , as required. \square

10.2. Supersaturation in the integer lattice. To prove Theorem 3.9, we first need a few intermediate results. Throughout this subsection, we will repeatedly use the following fact.

Fact 10.2. *Let $d, \ell, k \in \mathbb{N}$ and A be an $\ell \times k$ integer matrix of rank ℓ . Let $b \in (\mathbb{Z}^d)^\ell$ and $S \subseteq \mathbb{Z}^d$ be a finite subset. The number of solutions to $Ax = b$ in S is at most $|S|^{k-\ell}$.*

Proof. As A has rank ℓ , there are ℓ linearly independent columns in A , w.l.o.g. the first ℓ columns. We bound the number of solutions $x = (x_1, \dots, x_k)$ to $Ax = 0$ in S as follows. There are at most $|S|^{k-\ell}$ choices for $x_{\ell+1}, \dots, x_k$. For each such choice, $x' := (x_1, \dots, x_\ell)$ is a solution to $A'x' = b'$ in $S \subseteq \mathbb{Z}^d$ where A' is the $\ell \times \ell$ matrix consisting of the first ℓ columns of A and b' is some vector in $(\mathbb{Z}^d)^\ell$. Since the columns of A' are linearly independent, there is at most one solution to $A'x' = b'$ in \mathbb{Z}^d . Thus, there are at most $|S|^{k-\ell}$ solutions to $Ax = 0$ in S , as required. \square

The following lemma essentially states that if a supersaturation result holds for $S \subseteq \mathbb{Z}^d$ then it also holds for $S' \subseteq S$, provided that S' contains almost all elements of S .

Lemma 10.3. *Let $r, d, \ell, k \in \mathbb{N}$ and $\gamma > 0$. Let A be a partition regular $\ell \times k$ integer matrix of rank ℓ and let S and S' be finite sets such that $S' \subseteq S \subseteq \mathbb{Z}^d$ and $|S'| \geq (1 - \gamma/(2k))|S|$. Fix an r -colouring of S (and thus S'). If there are at least $\gamma|S|^{k-\ell}$ monochromatic solutions to $Ax = 0$ in S , then there are at least $(\gamma/2)|S|^{k-\ell}$ monochromatic solutions to $Ax = 0$ in S' .*

Proof. We determine an upper bound on the number of monochromatic solutions $(\mathbf{z}_1, \dots, \mathbf{z}_k)$ which lie in S but not in S' . For one such solution, there must be some \mathbf{z}_i which lies in S but not in S' . We have k choices for i and $|S| - |S'| \leq \gamma|S|/(2k)$ choices for \mathbf{z}_i . The remaining elements $(\mathbf{z}_1, \dots, \mathbf{z}_{i-1}, \mathbf{z}_{i+1}, \dots, \mathbf{z}_k)$ form a solution to $A'x = \mathbf{b}$ in S where A' is obtained from A by removing the i th column and \mathbf{b} is some vector. Note that A' has rank ℓ . If not, then the i th column of A cannot be expressed as a linear combination of the columns of A' . In particular, the i th entry of any solution to $Ax = 0$ in \mathbb{Z} must be 0, contradicting the assumption that A is partition regular. Since A' has rank ℓ , there are at most $|S|^{k-1-\ell}$ solutions to $A'x = \mathbf{b}$ in S by Fact 10.2. Overall, there are at most $k \cdot \frac{\gamma|S|}{2k} \cdot |S|^{k-\ell-1} = \frac{\gamma}{2}|S|^{k-\ell}$ possible choices for $(\mathbf{z}_1, \dots, \mathbf{z}_k)$.

Therefore, the number of monochromatic solutions to $Ax = 0$ in S' is at least

$$\gamma|S|^{k-\ell} - \frac{\gamma}{2}|S|^{k-\ell} = \frac{\gamma}{2}|S|^{k-\ell}.$$

\square

The following proposition asserts that if the statement of Theorem 3.9 holds for $d \in \mathbb{N}$ then it also holds for $d - 1$.

Proposition 10.4. *Let $r, d, n \in \mathbb{N}$ with $d \geq 2$ and $\delta > 0$. Let A be an $\ell \times k$ integer matrix of rank ℓ . If every r -colouring of $[n]^d$ yields at least $\delta n^{d(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^d$, then every r -colouring of $[n]^{d-1}$ yields at least $\delta n^{(d-1)(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^{d-1}$.*

Proof. Let A, r, d, n and δ be as in the statement of the proposition. Let $f : [n]^d \rightarrow [n]^{d-1}$ be such that for $\mathbf{x} = (x_1, \dots, x_d) \in [n]^d$ we have $f(\mathbf{x}) := (x_1, \dots, x_{d-1})$, i.e., $f(\mathbf{x})$ is the vector consisting of the first $d - 1$ entries of \mathbf{x} . We first prove the following property of f .

Claim 10.5. *Given a solution $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ to $Ax = 0$ in $[n]^{d-1}$, there are at most $n^{k-\ell}$ solutions $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ to $Ax = 0$ in $[n]^d$ such that $f(\mathbf{x}_i) = \mathbf{y}_i$ for every $i \in [k]$.*

Proof of the claim. Let S be the set of solutions $(\underline{\mathbf{x}}_1, \dots, \underline{\mathbf{x}}_k)$ in $[n]^d$ such that $f(\underline{\mathbf{x}}_i) = \underline{\mathbf{y}}_i$ for every $i \in [k]$. Let $q : S \rightarrow [n]^k$ with $q((\underline{\mathbf{x}}_1, \dots, \underline{\mathbf{x}}_k)) := (x_1, \dots, x_k)$ where x_i is the d th entry of $\underline{\mathbf{x}}_i$. Since S is a set of solutions in $[n]^d$, it follows that the image $q(S)$ of q is a set of solutions to $Ax = 0$ in $[n]$. In particular, $|q(S)| \leq n^{k-\ell}$ by Fact 10.2. Note that the assumption $f(\underline{\mathbf{x}}_i) = \underline{\mathbf{y}}_i$ for every $i \in [k]$ implies q is injective, hence $|S| = |q(S)| \leq n^{k-\ell}$, as required. \blacksquare

Pick an arbitrary r -colouring $\mathcal{C} : [n]^{d-1} \rightarrow [r]$ of $[n]^{d-1}$. Let $\mathcal{C}^* : [n]^d \rightarrow [r]$ be the r -colouring of $[n]^d$ where $\mathcal{C}^*(\underline{\mathbf{x}}) := \mathcal{C}(f(\underline{\mathbf{x}}))$. If $(\underline{\mathbf{x}}_1, \dots, \underline{\mathbf{x}}_k)$ is a monochromatic solution to $Ax = 0$ in $[n]^d$ (with respect to \mathcal{C}^*) then $(f(\underline{\mathbf{x}}_1), \dots, f(\underline{\mathbf{x}}_k))$ is a monochromatic solution to $Ax = 0$ in $[n]^{d-1}$ (with respect to \mathcal{C}). By assumption, there exist at least $\delta n^{d(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^{d-1}$ with respect to \mathcal{C} . Hence, by Claim 10.5, the number of monochromatic solutions in $[n]^d$ with respect to \mathcal{C}^* is at least

$$\frac{\delta n^{d(k-\ell)}}{n^{k-\ell}} = \delta n^{(d-1)(k-\ell)},$$

as required. \square

By Proposition 10.4, it is clear that if the statement of Theorem 3.9 holds for an infinite sequence $d_1 < d_2 < \dots$ of positive integers then it holds for any $d \in \mathbb{N}$. Therefore, it suffices to prove the following relaxation of Theorem 3.9.

Theorem 10.6. *Let $r \in \mathbb{N}$, $c \in \mathbb{N} \cup \{0\}$ and let A be a partition regular $\ell \times k$ integer matrix of rank ℓ . There exist $\delta, n_0 > 0$ such that for all $n \geq n_0$ the following holds: every r -colouring of $[n]^{2^c}$ yields at least $\delta n^{2^c(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^{2^c}$.*

Proof of Theorem 3.9. Let $r, d \in \mathbb{N}$ and let A be a partition regular $\ell \times k$ integer matrix of rank ℓ . Choose $c \in \mathbb{N}$ such that $d \leq 2^c$. Let δ, n_0 be the constants obtained by applying Theorem 10.6 with parameters r, c and A . By Proposition 10.4, the statement of Theorem 3.9 holds for A, r, δ, n_0 and every $d' \in \mathbb{N}$ with $d' \leq 2^c$. In particular, it holds for A, r, d, δ, n_0 . \square

We conclude this subsection with the proof of Theorem 10.6. Before proceeding with the proof, we give a rough intuition of the main idea by considering an easier case. For integers $a, b \in \mathbb{Z}$, we write $[a, b] := \{m \in \mathbb{Z} : a \leq m \leq b\}$. Let $A = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$. Given an arbitrary r -colouring \mathcal{C} of $[-n, n]^2$, our aim is to find $\Theta(n^4)$ monochromatic solutions to $Ax = 0$.

Define a colouring \mathcal{C}^* of $[(\varepsilon n)^2]$, where $\varepsilon > 0$ is a small constant, as follows. If $z \in [(\varepsilon n)^2]$, then $z = x + (\varepsilon n)y$ for a unique pair $(x, y) \in [\varepsilon n] \times [0, \varepsilon n - 1]$. Then colour z with $(c_{-2}, c_{-1}, c_0, c_1, c_2)$ where c_λ is the colour of $(x - \lambda(\varepsilon n), y + \lambda) \in [-n, n]^2$ with respect to \mathcal{C} . Note that if $z_1 + z_2 - z_3 = 0$ is a solution in $[(\varepsilon n)^2]$, then $(x_1 + x_2 - x_3) + (\varepsilon n)(y_1 + y_2 - y_3) = 0$. Since the x_i 's lie in $[\varepsilon n]$, $|y_1 + y_2 - y_3| \leq 2$. Thus, there exists $\lambda \in \mathbb{Z}$ with $|\lambda| \leq 2$ such that $(y_1 + \lambda) + (y_2 + \lambda) - (y_3 + \lambda) = 0$, which implies $(x_1 - (\varepsilon n)\lambda) + (x_2 - (\varepsilon n)\lambda) - (x_3 - (\varepsilon n)\lambda) = 0$. In particular, $\{(x_j - \lambda(\varepsilon n), y_j + \lambda)\}_{j=1,2,3}$ is a solution to $Ax = 0$ in $[-n, n]^2$. Furthermore, suppose $\{z_j\}_{j=1,2,3}$ is a monochromatic solution to $Ax = 0$ in $[(\varepsilon n)^2]$ with respect to \mathcal{C}^* ; say the colour is $(c_{-2}, c_{-1}, c_0, c_1, c_2)$. Then there exists $\lambda \in \mathbb{Z}$ with $|\lambda| \leq 2$ such that $\{(x_j - \lambda(\varepsilon n), y_j + \lambda)\}_{j=1,2,3}$ is a monochromatic solution to $Ax = 0$ in $[-n, n]^2$ with respect to \mathcal{C} , the colour being c_λ .

It is not difficult to see that different monochromatic solutions in $[(\varepsilon n)^2]$ with respect to \mathcal{C}^* give rise to different monochromatic solutions in $[-n, n]^2$ with respect to \mathcal{C} . That is, we have defined an injection from monochromatic solutions in $[(\varepsilon n)^2]$ with respect to \mathcal{C}^* to monochromatic solutions in $[-n, n]^2$ with respect to \mathcal{C} . By Theorem 3.5 there are $\Theta(n^4)$ many of the former, yielding the required result.

The proof of Theorem 10.6 proceeds by induction on c and follows a similar idea, where instead of Theorem 3.5 we invoke the induction hypothesis. Essentially, we project a solution to $Ax = 0$ in $[(\varepsilon n)^2]^{2^c}$ to an ‘almost’ solution in $[\varepsilon n]^{2^{c+1}}$, and then apply a shifting procedure to obtain an exact solution in $[n]^{2^{c+1}}$. The colouring \mathcal{C}^* we require is more involved than the one above due to two technicalities: (i) ensuring that the parameter λ (which is the analogue of λ in the general case) is a vector with integer entries, thereby implying that the shifting procedure is well-defined; (ii) ensuring that solutions generated by the shifting procedure lie in $[n]^{2^{c+1}}$ rather than $[-n, n]^{2^{c+1}}$.

We will use the following notation. Given two vectors $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$, we denote their concatenation as $\underline{\mathbf{x}} * \underline{\mathbf{y}}$. Furthermore, we let $\text{abs}[\underline{\mathbf{x}}]$ be the vector obtained by taking the absolute value of all entries of $\underline{\mathbf{x}}$. For example, if $\underline{\mathbf{x}} = (1, -2)$ and $\underline{\mathbf{y}} = (3, 4)$ then $\underline{\mathbf{x}} * \underline{\mathbf{y}} = (1, -2, 3, 4)$ and $\text{abs}[\underline{\mathbf{x}}] = (1, 2)$. We write $\underline{\mathbf{x}} \cdot \underline{\mathbf{y}}$ for the scalar product of $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$. Let $\underline{\mathbf{1}}$ denote the all one vector.

Proof of Theorem 10.6. Let $r \in \mathbb{N}$ and let A be a partition regular $\ell \times k$ integer matrix of rank ℓ . We proceed by induction on c . The case $c = 0$ follows immediately from Theorem 3.5. For the induction hypothesis, suppose that the statement of Theorem 10.6 holds for some $c \in \mathbb{N} \cup \{0\}$.

Denote the i th row of A as $\underline{\mathbf{a}}_i$. Let $C := \max_{i \in [\ell]} \{|\underline{\mathbf{a}}_i \cdot \underline{\mathbf{1}}|, 1\}$. By relabelling the rows of A and changing their sign (these operations do not affect the set of solutions), we may assume that $C = \max\{\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{1}}, 1\} \geq 1$. Let $E \geq 0$ be the largest entry of A in absolute value. Fix constants ε, P such that

$$Ek \ll 1/\varepsilon \ll P,$$

where P is a prime number that does not divide C . Let $d := 2^c$ and $R := r^{(2Ek+1)^d} \cdot (CP)^d$. By assumption, we can apply Theorem 10.6 with input A, R and c to obtain $\gamma, n_0 > 0$. Finally, we take $n \in \mathbb{N}$ with $n \geq \max\{n_0/\varepsilon, 2^{d+2}Ek^2d/(\varepsilon\gamma)\}$. For simplicity, we will assume that εn is an integer.

Let \mathcal{C} be an arbitrary r -colouring of $[0, n]^{2d}$. We define an R -colouring \mathcal{C}^* of $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ as follows. For any $\underline{\mathbf{z}} \in [Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$, there exists a unique pair of vectors $\underline{\mathbf{x}} \in [\varepsilon n]^d$ and $\underline{\mathbf{y}} \in [Ek, \varepsilon n]^d$ such that $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}(\varepsilon n)$. Let $I := [-Ek, Ek]^d$. We colour $\underline{\mathbf{z}}$ with $(\{c_{\underline{\mathbf{s}}}\}_{\underline{\mathbf{s}} \in I}, \underline{\mathbf{t}})$ where for each $\underline{\mathbf{s}} \in I$, $c_{\underline{\mathbf{s}}}$ is the colour of $\text{abs}[(\underline{\mathbf{y}} + \underline{\mathbf{s}}) * (\underline{\mathbf{x}} - \underline{\mathbf{s}}(\varepsilon n))] \in [0, n]^{2d}$ with respect to \mathcal{C} and $\underline{\mathbf{t}} \in [CP]^d$ is the unique vector such that $\underline{\mathbf{y}} \equiv \underline{\mathbf{t}} \pmod{CP}$. Note that all entries of $(\underline{\mathbf{y}} + \underline{\mathbf{s}}) * (\underline{\mathbf{x}} - \underline{\mathbf{s}}(\varepsilon n))$ have absolute value at most $(\varepsilon n)(Ek + 1) \ll n$, so $\text{abs}[(\underline{\mathbf{y}} + \underline{\mathbf{s}}) * (\underline{\mathbf{x}} - \underline{\mathbf{s}}(\varepsilon n))]$ indeed lies in $[0, n]^{2d}$.

Observe that \mathcal{C}^* is a well-defined R -colouring of $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$: $\underline{\mathbf{x}}, \underline{\mathbf{y}}$ and $\underline{\mathbf{t}}$ are unique, so the colour of $\underline{\mathbf{z}} \in [Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ in \mathcal{C}^* is uniquely determined. There are r possible colours for each $c_{\underline{\mathbf{s}}}$ and $(CP)^d$ possible values for $\underline{\mathbf{t}}$; since $|I| = (2Ek + 1)^d$, at most R colours have been used.

Next, we describe a shifting procedure which generates a monochromatic solution in $[0, n]^{2d}$ with respect to \mathcal{C} given a monochromatic solution in $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ with respect to \mathcal{C}^* .

Suppose $(\underline{\mathbf{z}}_1, \dots, \underline{\mathbf{z}}_k)$ is a monochromatic solution to $Ax = 0$ in $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ with respect to \mathcal{C}^* . For every $1 \leq i \leq k$ we have $\underline{\mathbf{z}}_i = \underline{\mathbf{x}}_i + \underline{\mathbf{y}}_i(\varepsilon n)$ for some unique $\underline{\mathbf{x}}_i \in [\varepsilon n]^d$ and $\underline{\mathbf{y}}_i \in [Ek, \varepsilon n]^d$. Let $\underline{\mathbf{z}}_i := (z_i^1, \dots, z_i^d)$, $\underline{\mathbf{x}}_i := (x_i^1, \dots, x_i^d)$ and $\underline{\mathbf{y}}_i := (y_i^1, \dots, y_i^d)$. Pick an arbitrary $j \in [d]$ and let $\underline{\mathbf{x}} := (x_1^j, \dots, x_k^j)$, $\underline{\mathbf{y}} := (y_1^j, \dots, y_k^j)$ and $\underline{\mathbf{z}} := (z_1^j, \dots, z_k^j)$. Then we have $A\underline{\mathbf{z}} = 0$ and $\underline{\mathbf{z}} = \underline{\mathbf{x}} + (\varepsilon n)\underline{\mathbf{y}}$. In particular, for every $i \in [\ell]$ we have

$$(29) \quad \underline{\mathbf{a}}_i \cdot \underline{\mathbf{z}} = \underline{\mathbf{a}}_i \cdot \underline{\mathbf{x}} + (\varepsilon n)\underline{\mathbf{a}}_i \cdot \underline{\mathbf{y}} = 0.$$

Note that the absolute value of every entry of $\underline{\mathbf{a}}_i$ is at most E and every entry of $\underline{\mathbf{x}}$ lies in $[\varepsilon n]$, thus we obtain

$$(30) \quad |\underline{\mathbf{a}}_i \cdot \underline{\mathbf{x}}| \leq \varepsilon n Ek.$$

Now (29) and (30) immediately imply

$$(31) \quad |\underline{\mathbf{a}}_i \cdot \underline{\mathbf{y}}| \leq Ek.$$

Next, we distinguish two cases.

Case 1. Suppose that $C = \underline{\mathbf{a}}_1 \cdot \underline{\mathbf{1}} \geq 1$. Since $(\underline{\mathbf{z}}_1, \dots, \underline{\mathbf{z}}_k)$ is monochromatic with respect to \mathcal{C}^* , there exists some $t \in [CP]$ such that $\underline{\mathbf{y}} \equiv t\underline{\mathbf{1}} \pmod{CP}$. Hence, we have

$$\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{y}} \equiv \underline{\mathbf{a}}_1 \cdot (t\underline{\mathbf{1}}) \equiv t(\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{1}}) \equiv Ct \equiv 0 \pmod{C}.$$

Set $\lambda := -(\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{y}})/C$. The previous equality implies that λ is an integer, while by (31) we have $|\lambda| \leq Ek/C$. Observe that

$$(32) \quad \underline{\mathbf{a}}_1 \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) = \underline{\mathbf{a}}_1 \cdot \underline{\mathbf{y}} + \lambda C = 0.$$

As $\underline{\mathbf{y}} \equiv t\underline{\mathbf{1}} \pmod{CP}$, we have

$$0 \stackrel{(32)}{=} \underline{\mathbf{a}}_1 \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) \equiv (t + \lambda)\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{1}} \equiv C(t + \lambda) \pmod{P},$$

and so we must have that $t + \lambda \equiv 0 \pmod{P}$ since P is a prime number which does not divide C . It follows that

$$(33) \quad \underline{\mathbf{a}}_i \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) \equiv (t + \lambda)(\underline{\mathbf{a}}_i \cdot \underline{\mathbf{1}}) \equiv 0 \pmod{P}$$

for every $i \in [\ell]$. Note that

$$|\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}})| \leq |\underline{\mathbf{a}}_i \cdot \underline{\mathbf{y}}| + |\lambda(\underline{\mathbf{a}}_i \cdot \underline{\mathbf{1}})| \stackrel{(31)}{\leq} Ek + |\lambda C| \leq 2Ek.$$

As we picked P so that $P \gg Ek$, the previous inequality and (33) imply that

$$\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) = 0$$

for every $i \in [\ell]$. Furthermore, this together with (29) implies that, for every $i \in [\ell]$ we have

$$\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{x}} - \lambda(\varepsilon n)\underline{\mathbf{1}}) = 0.$$

Case 2. Suppose $\underline{\mathbf{a}}_1 \cdot \underline{\mathbf{1}} = 0$ (and so $\underline{\mathbf{a}}_i \cdot \underline{\mathbf{1}} = 0$ for every $i \in [\ell]$). Then set $\lambda := 0$ and observe that

$$\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) = \underline{\mathbf{a}}_i \cdot \underline{\mathbf{y}} \equiv t(\underline{\mathbf{a}}_i \cdot \underline{\mathbf{1}}) \equiv 0 \pmod{P}.$$

By (31) and as $P \gg Ek$, we deduce that $\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{y}} + \lambda\underline{\mathbf{1}}) = 0$ and so $\underline{\mathbf{a}}_i \cdot (\underline{\mathbf{x}} - \lambda(\varepsilon n)\underline{\mathbf{1}}) = 0$, for every $i \in [\ell]$.

In both cases, we obtain two solutions $(y_i^j + \lambda)_{i \in [k]}$ and $(x_i^j - \lambda(\varepsilon n))_{i \in [k]}$ to $Ax = 0$ in \mathbb{Z} , where $|\lambda| \leq Ek$. In general, since j was arbitrarily chosen, for every $j \in [d]$ there exists $\lambda_j \in [-Ek, Ek]$ such that $(y_i^j + \lambda_j)_{i \in [k]}$ and $(x_i^j - \lambda_j(\varepsilon n))_{i \in [k]}$ are two solutions to $Ax = 0$ in \mathbb{Z} . By setting $\underline{\lambda} := (\lambda_1, \dots, \lambda_d)$, we obtain that $(\underline{\mathbf{y}}_i + \underline{\lambda})_{i \in [k]}$ and $(\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))_{i \in [k]}$ are two solutions to $Ax = 0$ in \mathbb{Z}^d .

Note that $\underline{\mathbf{y}}_i + \underline{\lambda} \in [0, 2\varepsilon n]^d$ for every $i \in [k]$ since $\underline{\mathbf{y}}_i \in [Ek, \varepsilon n]^d$ and $\underline{\lambda} \in [-Ek, Ek]^d$. Furthermore, for every $m \in [d]$, the m th entries of the vectors $(\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))_{i \in [k]}$ must have the same sign since $\underline{\mathbf{x}}_i \in [\varepsilon n]^d$ and $\underline{\lambda}$ has integer entries. Hence $(\text{abs}[(\underline{\mathbf{y}}_i + \underline{\lambda}) * (\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))])_{i \in [k]}$ is a solution to $Ax = 0$ in $[0, n]^{2d}$.

Crucially, all vectors $(\text{abs}[(\underline{\mathbf{y}}_i + \underline{\lambda}) * (\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))])_{i \in [k]}$ have the same colour in \mathcal{C} by the construction of \mathcal{C}^* and the fact that $\underline{\lambda} \in I = [-Ek, Ek]^d$. Namely, if all $(\underline{\mathbf{z}}_1, \dots, \underline{\mathbf{z}}_k)$ are coloured with, say, $(\{c_{\mathbf{s}}\}_{\mathbf{s} \in I}, \mathbf{t})$ in \mathcal{C}^* then all $(\text{abs}[(\underline{\mathbf{y}}_i + \underline{\lambda}) * (\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))])_{i \in [k]}$ are coloured with $c_{\underline{\lambda}}$ in \mathcal{C} . Hence, $(\text{abs}[(\underline{\mathbf{y}}_i + \underline{\lambda}) * (\underline{\mathbf{x}}_i - \underline{\lambda}(\varepsilon n))])_{i \in [k]}$ is a monochromatic solution to $Ax = 0$ in $[0, n]^{2d}$.

Arbitrarily extend the R -colouring \mathcal{C}^* of $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ to an R -colouring \mathcal{C}^{**} of $[(\varepsilon n)^2]^d$. As $(\varepsilon n)^2 \geq n_0$, the induction hypothesis implies that there are at least $\gamma(\varepsilon n)^{2d(k-\ell)}$ monochromatic solutions in $[(\varepsilon n)^2]^d$ with respect to \mathcal{C}^{**} . Note that $[Ek(\varepsilon n) + 1, (\varepsilon n)^2]^d$ contains all but at most $dEk(\varepsilon n)^{2d-1} \leq \gamma(\varepsilon n)^{2d}/(2k)$ of the elements of $[(\varepsilon n)^2]^d$. Thus, by Lemma 10.3, there are at least

$(\gamma/2)(\varepsilon n)^{2d(k-\ell)}$ monochromatic solutions in $[Ek(\varepsilon n)+1, (\varepsilon n)^2]^d$ with respect to \mathcal{C}^* ; moreover, using the procedure described above, we can produce a monochromatic solution in $[n]^{2d}$ with respect to \mathcal{C} from each one of them. It remains to check how many times we could have counted each solution in $[n]^{2d}$.

Observe that a different monochromatic solution $(\mathbf{z}_1, \dots, \mathbf{z}_k)$ in $[(Ek(\varepsilon n)+1, (\varepsilon n)^2]^d$ would have generated different k -tuples $(y_i^j + \lambda)_{i \in [k]}$ and $(x_i^j - \lambda(\varepsilon n))_{i \in [k]}$, since \mathbf{x}_i and \mathbf{y}_i are unique and by (29). On the other hand, the preimage of a singleton in $\text{abs} : [-n, n]^d \rightarrow [0, n]^d$ has size at most 2^d . Hence, a monochromatic solution in $[n]^{2d}$ with respect to \mathcal{C} could be counted at most 2^d times.

In conclusion, there are at least $(\gamma/2^{d+1})(\varepsilon n)^{2d(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[0, n]^{2d}$ with respect to \mathcal{C} . Using Lemma 10.3, there are at least $(\gamma/2^{d+2})(\varepsilon n)^{2d(k-\ell)}$ monochromatic solutions to $Ax = 0$ in $[n]^{2d}$. Taking $\delta := (\gamma/2^{d+2})\varepsilon^{2d(k-\ell)}$, the statement of Theorem 10.6 holds for $c+1$ since $2d = 2^{c+1}$. This concludes the inductive step and the proof. \square

11. PROOF OF THE APPLICATIONS OF THE RANDOM RADO LEMMA

In this section, we prove the applications of the random Rado lemma presented in Sections 2 and 5, that is Theorems 2.7, 5.1, 5.2, 5.3, 5.4 and 5.5. The first uses the full version of the random Rado lemma, Lemma 6.8. The others all use the simplified version, Lemma 6.33.

Proof of Theorem 2.7. Let $r, k \in \mathbb{N}$ with $r \geq 2$ and $k \geq 3$. Let $A = (a_{ij})$ be the $(k-2) \times k$ matrix where

$$a_{ij} := \begin{cases} 1 & \text{if } i = j; \\ -2 & \text{if } i + 1 = j; \\ 1 & \text{if } i + 2 = j; \\ 0 & \text{otherwise.} \end{cases}$$

Hence $k\text{-Sol}_{\mathbf{P}_n}^A([k])$ corresponds to the set of all k -APs in \mathbf{P}_n .

The prime number theorem yields

$$(34) \quad |\mathbf{P}_n| = \Theta\left(\frac{n}{\log n}\right).$$

Note that any solution to $Ax = 0$ that does not correspond to a k -AP must be a trivial solution (i.e., all entries in x are the same). Moreover, given fixed integers a, b and $1 \leq i < j \leq k$, there is at most one k -AP that has a in its i th position and b in its j th position. Thus, for any $Y \subseteq [k]$ with $2 \leq |Y| \leq k$, each element of $\text{Sol}_{\mathbf{P}_n}^A(Y)$ corresponds to either a trivial solution to $Ax = 0$ or a single k -AP. Hence, by Theorem 2.1 we have that

$$(35) \quad |\text{Sol}_{\mathbf{P}_n}^A(Y)| = \Theta\left(\frac{n^2}{\log^k n}\right) \quad \text{for each } Y \subseteq [k] \text{ with } 2 \leq |Y| \leq k.$$

Therefore, (34) and (35) imply that

$$(36) \quad p_Y(A, \mathbf{P}_n) = \Theta\left(\left(\frac{n}{\log^{k-1} n}\right)^{-\frac{1}{|Y|-1}}\right) \quad \text{for each } Y \subseteq [k] \text{ with } 2 \leq |Y| \leq k.$$

Note that $[k]$ maximises $p_Y(A, \mathbf{P}_n)$ over all $Y \subseteq [k]$ with $|Y| \geq 2$, and so

$$(37) \quad \hat{p}_n := \hat{p}(A, \mathbf{P}_n) = \Theta\left(n^{-\frac{1}{k-1}} \log n\right).$$

We now verify the conditions of Lemma 6.8.

(A1): By (34) and (37), we have $|\mathbf{P}_n| \hat{p}_n = \Theta(n^{\frac{k-2}{k-1}}) \rightarrow \infty$ and $\hat{p}_n \rightarrow 0$ as $n \rightarrow \infty$, since $k \geq 3$.

(A2): Theorem 3.6 immediately implies that \mathbf{P}_n is (A, r) -supersaturated.

(A3): Let $W \subseteq Y \subseteq [k]$. First, consider the case $|W| = 1$. Fix a prime $q \in \mathbf{P}_n$. If $|Y| = 1$ then $|\text{Sol}_{\mathbf{P}_n}^A(q, W, Y)| \leq 1$. If $|Y| \geq 2$, we have

$$(38) \quad |\text{Sol}_{\mathbf{P}_n}^A(q, W, Y)| = O\left(\frac{n}{\log^{k-1} n}\right) \stackrel{(34),(35)}{=} O\left(\frac{|\text{Sol}_{\mathbf{P}_n}^A(Y)|}{|\mathbf{P}_n|}\right) = O\left(\frac{|\text{Sol}_{\mathbf{P}_n}^A(Y)|}{|\text{Sol}_{\mathbf{P}_n}^A(W)|}\right),$$

where we used Lemma 2.3 for the first equality and $|\text{Sol}_{\mathbf{P}_n}^A(W)| \leq |\mathbf{P}_n|$ for the third.

Now suppose $|W| \geq 2$. Since specifying two elements (and their position) of a k -AP uniquely determines the latter, for any $z_W \in S^{|W|}$, we have

$$|\text{Sol}_{\mathbf{P}_n}^A(z_W, W, Y)| \leq 1 \stackrel{(35)}{=} O\left(\frac{|\text{Sol}_{\mathbf{P}_n}^A(Y)|}{|\text{Sol}_{\mathbf{P}_n}^A(W)|}\right).$$

Therefore, there exists $B \geq 1$ such that (A, \mathbf{P}_n) is B -extendable for every n sufficiently large, as required.

(A4): Let $W \subset Y \subseteq [k]$ with $|W| = 1$. Since $|Y| \geq 2$, equation (38) holds. The first two equalities of (38) immediately imply condition (A4).

(A5): We take $X_n = [k]$ for all $n \in \mathbb{N}$. We have $|X_n| = k \geq 3$ and $p_{X_n}(A, \mathbf{P}_n) = \hat{p}_n$. Let $W, W' \subset [k]$ with $|W| = 2$ and $|W'| \geq 2$. As $|W'| \leq k - 1$ we have

$$\frac{|\mathbf{P}_n|^2}{|\text{Sol}_{\mathbf{P}_n}^A(W)|} \left(\frac{p_{W'}(A, \mathbf{P}_n)}{p_{[k]}(A, \mathbf{P}_n)}\right)^{|W'|-1} \stackrel{(34),(35),(36)}{=} O\left(\log^{k-2} n \left(\frac{n}{\log^{k-1} n}\right)^{\frac{|W'|-k}{k-1}}\right) \rightarrow 0,$$

as $n \rightarrow \infty$, as required.

(A6): Let $W \subseteq [k]$ with $|W| = 2$. Then (34) and (35) imply

$$\frac{|\mathbf{P}_n|}{|\text{Sol}_{\mathbf{P}_n}^A(W)|} = O\left(\frac{\log^{k-1} n}{n}\right) \rightarrow 0,$$

as $n \rightarrow \infty$, as required.

Since conditions (A1)–(A6) hold, we can apply Lemma 6.8. Note that $\mathbf{P}_{n,p}$ is (A, r) -Rado if and only if $\mathbf{P}_{n,p}$ is (r, k) -van der Waerden. Thus Lemma 6.8 yields Theorem 2.7. \square

Proof of Theorem 5.1. Let $r, d \in \mathbb{N}$ with $r \geq 2$. Let A satisfy the assumptions of Theorem 5.1. We will verify that the conditions of Lemma 6.33 hold, where we take $S_n := [n]^d$. Note that $[n]$ is a finite subset of the field \mathbb{Q} and $m_{S_n}(A) = m(A)$.

(C2): Theorem 3.9 implies that $([n]^d)_{n \in \mathbb{N}}$ is (A, r) -supersaturated.

(C3): An immediate consequence of Theorem 3.9 is that there exists some $\delta > 0$ such that the number of solutions to $Ax = 0$ in $[n]^d$ is at least $\delta n^{d(k-\ell)}$, for all n sufficiently large. Since there are at most $n^{d(k-\ell)}$ solutions to $Ax = 0$ in $[n]^d$ (by e.g., Lemma 6.16), $(A, [n]^d)$ is δ -rich for n sufficiently large.

(C4): Since A is irredundant, A is irredundant with respect to \mathbb{Q}^d . Furthermore, by Theorem 3.9, A is 3-partition regular in $[n]^d$ for n sufficiently large. Thus, we can apply Lemma 6.23 with $\mathbb{F} := \mathbb{Q}$ and $S := [n]^d$ to obtain that $(A, [n]^d)$ is abundant for n sufficiently large.

(C1): By (C4), $m(A)$ is well-defined (see Remark 6.28). As A is full rank, considering $W = [k]$ yields $m(A) \geq \frac{k-1}{k-1-\ell} > 1$. Since (C3) holds, Lemma 6.32 yields $\hat{p}_n = \Theta(n^{-d/m(A)})$. Therefore, we have $|[n]^d| \hat{p}_n \rightarrow \infty$ and $\hat{p}_n \rightarrow 0$ as $n \rightarrow \infty$.

(C5): Note that $\text{rank}_{S_n}(A) > 0$ for every n since $S_n \neq \{0\}$. As (C3) and (C4) are satisfied, we can apply Lemma 6.26. Lemma 6.26 and (C1) imply that (C5) holds.

As conditions (C1)–(C5) hold, we can apply Lemma 6.33 to obtain Theorem 5.1. \square

Proof of Theorem 5.3. Let $r \in \mathbb{N}$ with $r \geq 2$ and let A and G be as in the statement of Theorem 5.3. Note that G^n is a finite abelian group. We will verify that the hypothesis of Lemma 6.33 holds, where we take $S_n := G^n$. We will repeatedly use the fact that $\text{rank}_{G^n}(A_W) = \text{rank}_G(A_W)$ for any $W \subseteq [k]$; this follows easily from Definition 6.11. Thus, $m_{G^n}(A) = m_G(A)$.

(C2): Since G is a finite abelian group with exponent s and A satisfies the s -columns condition, Theorem 3.8 implies $(G^n)_{n \in \mathbb{N}}$ is (A, r) -supersaturated.

(C3): Equation (4) implies (A, G^n) is 1-rich.

(C4): Let $W \subseteq [k]$ with $|W| = 2$. By assumption, (A, G) is abundant, and so $\text{rank}_G(A_{\overline{W}}) = \text{rank}_G(A)$. This implies $\text{rank}_{G^n}(A_{\overline{W}}) = \text{rank}_{G^n}(A)$. As W is arbitrary, (A, G^n) is abundant.

(C1): By (C4), $m_{G^n}(A)$ is well-defined and positive (see Remark 6.28). Since (C3) holds, Lemma 6.32 yields $\hat{p}_n := \hat{p}(A, G^n) = \Theta((|G|^n)^{-1/m_{G^n}(A)})$. As $\text{rank}_{G^n}(A) = \text{rank}_G(A) > 0$ and by using $W = [k]$, we have $m_{G^n}(A) \geq \frac{k-1}{k-1-\text{rank}_{G^n}(A)} > 1$. Therefore we have $|G|^n \hat{p}_n \rightarrow \infty$ and $\hat{p}_n \rightarrow 0$ as $n \rightarrow \infty$.

(C5): Let $Z \subseteq [k]$ and

$$D(Z) := \frac{|Z| - 1 - \text{rank}_G(A) + \text{rank}_G(A_{\overline{Z}})}{|Z| - 1}.$$

As (A, G^n) is 1-rich, (7) implies that $p_Z(A, G^n) = |G|^{-n \cdot D(Z)}$ for every $n \in \mathbb{N}$. Pick $X \subseteq [k]$, $|X| \geq 2$ so that the function $D(X)$ is minimised; additionally choose X so that $|X|$ is as small as possible under this assumption. It follows that $p_X(A, G^n) = \hat{p}(A, G^n)$. For every $|Z| = 2$ we have $D(Z) = 1$, by (C4), while $D([k]) < 1$ since $\text{rank}_G(A) > 0$. Hence, $|X| \geq 3$.

Let $W' \subset X$ with $|W'| \geq 2$. We have

$$\frac{p_{W'}(A, G^n)}{p_X(A, G^n)} = (|G|^n)^{-D(W') + D(X)}.$$

Since X is minimal and there are finitely many subsets $Z \subset X$, there exists $\varepsilon > 0$ such that we have $D(Z) - D(X) \geq \varepsilon$ for every $Z \subset X$. It follows that

$$\frac{p_{W'}(A, G^n)}{p_X(A, G^n)} \leq |G|^{-\varepsilon n} \rightarrow 0$$

as $n \rightarrow \infty$.

As conditions (C1)–(C5) hold, we can apply Lemma 6.33 to obtain Theorem 5.3. □

Proof of Theorem 5.2. Since $m_{\mathbb{Z}_4}(A) = 4/3$ when $A = (2 \ 2 \ -2)$ (see Example 6.29), to prove the theorem it suffices to verify that the hypothesis of Theorem 5.3 holds in our setting.

Firstly, \mathbb{Z}_4 is an abelian group with exponent 4. The matrix A satisfies the 4-columns condition: the last two entries of A sum to zero and they span the first entry.

Next, we check (A, \mathbb{Z}_4) is abundant and $\text{rank}_{\mathbb{Z}_4}(A) > 0$. The total number of solutions to $2x_1 + 2x_2 - 2x_3 = 0$ in \mathbb{Z}_4 is $4 \cdot 4 \cdot 2 = 4^{3-1/2}$; so $\text{rank}_{\mathbb{Z}_4}(A) = 1/2 > 0$. For any $W \subseteq [3]$ with $|W| = 2$, the equation $A_{\overline{W}}x = 0$ in \mathbb{Z}_4 is the same as $2x = 0$ in \mathbb{Z}_4 . We have $2 = 4^{1-1/2}$ solutions to $2x = 0$ in \mathbb{Z}_4 , thus $\text{rank}_{\mathbb{Z}_4}(A_{\overline{W}}) = 1/2 = \text{rank}_{\mathbb{Z}_4}(A)$. In particular, (A, \mathbb{Z}_4) is abundant. □

Proof of Theorem 5.4. Consider any full rank $\ell \times k$ matrix A with entries from \mathbb{F} that satisfies the hypothesis of the theorem. By Definition 6.12, $\text{rank}_{\mathbb{F}}(A) = \text{rank}_{\mathbb{F}^n}(A)$. Moreover, as \mathbb{F}^n is an abelian group, one can also define $\text{rank}_{\mathbb{F}^n}(A)$ analogously to Definition 6.11. In particular, both these notions of $\text{rank}_{\mathbb{F}^n}(A)$ are equivalent. Thus, (4) holds in our setting (with $S = \mathbb{F}^n$).

Similarly, one can define $m_{\mathbb{F}^n}(A)$ analogously to Definition 6.27. Further, as $\text{rank}_{\mathbb{F}}$ is the same as $\text{rank}_{\mathbb{F}^n}$ we have that $m_{\mathbb{F}^n}(A) = m_{\mathbb{F}}(A)$, where the latter parameter was defined in (2).

As (4) holds, one can check that all the results used to prove Lemma 6.33 extend to this setting of matrices with entries from \mathbb{F} . Thus, the statement of Lemma 6.33 holds for $S_n := \mathbb{F}^n$ and full rank matrices A with entries from \mathbb{F} . In particular, the notions of (A, r) -saturated, abundant and ε -rich can be analogously defined in this setting. (For a formal justification of all of this, see the PhD thesis of the first author [23].)

Since $m_{\mathbb{F}^n}(A) = m_{\mathbb{F}}(A)$, to prove the theorem it suffices to show that (C1)–(C5) hold with respect to $S_n = \mathbb{F}^n$.

(C2): Theorem 3.7 implies that $(\mathbb{F}^n)_{n \in \mathbb{N}}$ is (A, r) -supersaturated.

(C3): Equation (4) implies that (A, \mathbb{F}^n) is 1-rich.

(C4): Theorem 3.3 implies that A is 3-partition regular in $\mathbb{F}^n \setminus \{0\}^n$ (for n sufficiently large). By assumption A is irredundant with respect to \mathbb{F}^n . Arguing precisely as in the proof of Lemma 6.23, one can conclude that (A, \mathbb{F}^n) is abundant for all sufficiently large n .

(C1): As (A, \mathbb{F}^n) is abundant for sufficiently large n , (A, \mathbb{F}) is abundant. Remark 6.28 implies that $m_{\mathbb{F}}(A)$ is well-defined and positive. As A is full rank, considering $W = [k]$ yields $m_{\mathbb{F}}(A) \geq \frac{k-1}{k-1-\ell} > 1$. Since (C3) holds, (the analogue of) Lemma 6.32 yields $\hat{p}_n = \Theta(|\mathbb{F}^n|^{-1/m_{\mathbb{F}}(A)})$. Therefore we have $|\mathbb{F}^n| \hat{p}_n \rightarrow \infty$ and $\hat{p}_n \rightarrow 0$ as $n \rightarrow \infty$.

(C5): Since (C3) and (C4) are satisfied, we can apply (the analogue of) Lemma 6.26. This together with (C1) implies that (C5) holds. □

Proof of Theorem 5.5. Let A be as in the theorem. Recall that as A is an integer matrix, we can view it as a matrix with entries from \mathbb{F} (i.e., an entry a_{ij} corresponds to $a_{ij} \cdot 1_{\mathbb{F}} \in \mathbb{F}$). Thus, we show that the theorem follows directly from Theorem 5.4.

Note that the finite field \mathbb{F} of order q^k has exponent q (in particular, $\langle 1_{\mathbb{F}} \rangle \cong \mathbb{Z}_q$) so the fact that A satisfies the q -columns condition immediately implies it also satisfies the columns condition over \mathbb{F} . We can therefore apply Theorem 5.4 to obtain the theorem. □

12. FURTHER APPLICATIONS AND CONCLUDING REMARKS

We have seen a number of applications of the (simplified) random Rado lemma. Of course, there are many other potential applications; we give a couple more in the next subsection. In Section 12.2 we consider a resilience result. In the final subsection, we consider some possible future research directions. All proofs omitted in this section appear in the PhD thesis of the first author [23].

12.1. Further applications of the random Rado lemma. The next result is an analogue of the random Rado theorem for \mathbb{Z}_n .

Theorem 12.1. *For all irredundant partition regular full rank integer matrices A , and all $r \geq 2$, there exist constants $C, c > 0$ such that the following holds.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{Z}_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m(A)}; \\ 1 & \text{if } p \geq Cn^{-1/m(A)}. \end{cases}$$

□

Theorem 12.1 follows from Lemma 6.8,¹⁴ though actually the 1-statement of Theorem 12.1 is also an immediate consequence of Theorem 4.3. The reader may wonder why the parameter $m(A)$ rather than $m_{\mathbb{Z}_n}(A)$ appears in the statement of Theorem 12.1. In particular, $m_{\mathbb{Z}_n}(A)$ may not equal $m(A)$ and it can be the case that $m_{\mathbb{Z}_n}(A)$ may not be the same for all values of n . For

¹⁴In particular, one cannot apply Lemma 6.33 here as not all choices of A ensure (A, \mathbb{Z}_n) is abundant; see [23].

example, for $A = \begin{pmatrix} 2 & 2 & -2 \end{pmatrix}$ and even n we have $m_{\mathbb{Z}_n}(A) = 2/(1 + \log_n 2)$. However, in this case we have $n^{-1/m_{\mathbb{Z}_n}(A)} = \frac{1}{\sqrt{2}}n^{-1/m(A)}$. More generally, we have that $n^{-1/m(A)} = \Theta(n^{-1/m_{\mathbb{Z}_n}(A)})$ and so we can use $m(A)$ in the statement of Theorem 12.1 instead of $m_{\mathbb{Z}_n}(A)$.

In general, our machinery ensures that, given any ‘well behaved’ sequence $(S_n)_{n \in \mathbb{N}}$ of finite abelian groups, one can obtain a corresponding random Ramsey-type theorem if one can prove a supersaturation result. It would therefore be interesting to prove other supersaturation results for abelian groups à la Theorem 3.8. Furthermore, it is likely that ‘projection’ type arguments such as in the proof of Theorem 3.9 can be employed to obtain new supersaturation results from existing results (such as Theorems 3.8 and 3.9).

As noted in Section 3.2, sometimes supersaturation results are easy to obtain, for example, for matrices A that are translation-invariant. In such cases there are fewer conditions to check in order to prove a random Ramsey-type result. By applying Lemma 6.33 we obtain the following.

Theorem 12.2. *Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of finite abelian groups and A be an integer matrix that is translation-invariant with respect to S_n for all sufficiently large $n \in \mathbb{N}$. If (C1), (C4) and (C5) are satisfied then there exist constants $C, c > 0$ such that the following holds.*

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p} \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_{S_n}(A)}; \\ 1 & \text{if } p \geq Cn^{-1/m_{S_n}(A)}. \end{cases}$$

□

12.2. Resilience in the primes. The following is an immediate consequence of Theorem 2.1.

Lemma 12.3. *Let $r \geq 1$ and $k \geq 3$. Given any $\gamma > 0$ there exist $n_0, \varepsilon > 0$ such that for all $n \geq n_0$, the following holds. Suppose $X \subseteq \mathbf{P}_n$ with $|X| \geq \gamma|\mathbf{P}_n|$. Then for any r -colouring of X there are at least $\varepsilon n^2 / \log^k n$ monochromatic k -APs in X .* □

Using Lemma 12.3 one can obtain the following resilience-type result.

Theorem 12.4. *Let $r \geq 1$, $k \geq 3$ and $\delta > 0$. There exists a constant $C > 0$ such that if $p \geq Cn^{-1/(k-1)} \log n$, then w.h.p. $\mathbf{P}_{n,p}$ has the property that for any subset $X \subseteq \mathbf{P}_{n,p}$ with $|X| \geq \delta|\mathbf{P}_{n,p}|$, whenever X is r -coloured, there is a monochromatic k -AP.* □

By setting $r = 1$ in Theorem 12.4, we obtain a sharpening of Theorem 2.4, since the property considered is that of being (δ, k) -Szemerédi. For $r \geq 2$, Theorem 12.4 strengthens the 1-statement of Theorem 2.7 as it states that w.h.p. not only can we find a monochromatic k -AP in $\mathbf{P}_{n,p}$ whenever r -coloured, but we can afford to delete a $(1 - \delta)$ -proportion of $\mathbf{P}_{n,p}$ and still guarantee such a monochromatic k -AP.

The proof of Theorem 12.4 is an easy adaption of the proof of the 1-statement of Theorem 7.7, using ideas from the proof of a resilience-type result in [32].

12.3. Further directions and examples. Another direction in which one can generalise the random Rado lemma is by removing the requirement of matrices having integer entries. Given a matrix A with entries in a ring R one can ask whether a finite colouring of an R -module yields a monochromatic solution to $Ax = 0$ or not. If one can in fact prove a suitable supersaturation result, then our machinery could yield a random Ramsey-type theorem for this setting. Theorem 5.4 already provides one such result in this direction.

Condition (P5) in the statement of Theorem 7.7 arises as an artifact of our proof. It would be interesting to investigate possible relaxations of this condition; this could in turn lead to a relaxation of condition (A5) from Lemma 6.8.

Condition (C4) in Theorem 6.33 states that (A, S_n) is abundant for every sufficiently large $n \in \mathbb{N}$. This property is used to verify, e.g., (A5) when deducing Lemma 6.33 from Lemma 6.8. However,

(as briefly touched upon when discussing Theorem 12.1) just because a sequence $(S_n)_{n \in \mathbb{N}}$ of finite abelian groups does not satisfy (C4) does not necessarily mean one cannot apply Lemma 6.8 instead.

The following is another example of this. Let $A := \begin{pmatrix} 2 & -2 & 63 & 65 \end{pmatrix}$ and $S_n := \mathbb{Z}_{128}^n$ for every $n \in \mathbb{N}$. The pair (A, S_n) is not abundant since $\text{rank}_{S_n}(A) = 1$, whereas $\text{rank}_{S_n}(A_{\{1,2\}}) = 6/7$, so we cannot apply Lemma 6.33. However, conditions (A1)–(A6) do hold for $(S_n)_{n \in \mathbb{N}}$ and A , and so we can apply Lemma 6.8 to deduce the following theorem.

Theorem 12.5. *Let $A = \begin{pmatrix} 2 & -2 & 63 & 65 \end{pmatrix}$ and $r \geq 2$. There exist constants $c, C > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[(\mathbb{Z}_{128}^n)_p \text{ is } (A, r)\text{-Rado}] = \begin{cases} 0 & \text{if } p \leq c 128^{-2n/3}; \\ 1 & \text{if } p \geq C 128^{-2n/3}. \end{cases}$$

□

Conversely, consider $B := \begin{pmatrix} 4 & -4 & 63 & 65 \end{pmatrix}$. The sequence $(S_n)_{n \in \mathbb{N}}$ is (B, r) -supersaturated for every $r \in \mathbb{N}$: this follows from Theorem 3.8, as \mathbb{Z}_{128} is an abelian group with exponent 128 and B satisfies the 128-columns condition as $4 - 4 + 63 + 65 \equiv 0 \pmod{128}$. However, we cannot apply Lemma 6.8, as condition (A5) does not hold. Indeed, it is easy to check that, for $W \subseteq [4]$,

$$|\text{Sol}_{S_n}^B(W)| = \begin{cases} |S_n|^3 & \text{if } |W| = [4]; \\ |S_n|^{|W|-2/7} & \text{if } \{3, 4\} \subseteq W \neq [4]; \\ |S_n|^{|W|} & \text{otherwise.} \end{cases}$$

It immediately follows that, for $|W| \geq 2$,

$$p_W(B, S_n) = \left(\frac{|\text{Sol}_{S_n}^B(W)|}{|S_n|} \right)^{-\frac{1}{|W|-1}} = \begin{cases} |S_n|^{-2/3} & \text{if } |W| = [4]; \\ |S_n|^{-(|W|-9/7)/(|W|-1)} & \text{if } \{3, 4\} \subseteq W \neq [4]; \\ |S_n|^{-1} & \text{otherwise.} \end{cases}$$

Note that if $(X_n)_{n \in \mathbb{N}}$ is a sequence such that $X_n \subseteq [4]$ and $p_{X_n}(B, S_n) = \Omega(\hat{p}(B, S_n))$, then $X_n = [4]$ for all sufficiently large $n \in \mathbb{N}$. In this case, for $W_n := \{3, 4\} \subset [4]$ we have

$$\frac{|S_n|^2}{|\text{Sol}_{S_n}^B(W_n)|} \left(\frac{p_{W_n}(B, S_n)}{p_{[4]}(B, S_n)} \right) = |S_n|^{5/21} = (128)^{5n/21} \not\rightarrow 0$$

as $n \rightarrow \infty$. So $(S_n)_{n \in \mathbb{N}}$ is not compatible with respect to B .

Open access statement. This research was funded in part by EPSRC grant EP/V048287/1. For the purpose of open access, a CC BY public copyright licence is applied to any Author Accepted Manuscript arising from this submission.

Data availability statement. The proofs omitted in this paper can be found in the PhD thesis of the first author [23].

REFERENCES

- [1] E. Aigner-Horev and Y. Person, An asymmetric random Rado theorem: 1-statement, *J. Combin. Theory Ser. A* **193** (2023), 105687.
- [2] D. Altman and A. Liebenau, On the uncommonness of minimal rank-2 systems of linear equations, arXiv:2404.18908.
- [3] J. Balogh, H. Liu and M. Sharifzadeh, The number of subsets of integers with no k -term arithmetic progression, *Int. Math. Res. Not.* **2017(20)** (2017), 6168–6186.
- [4] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* **28** (2015), 669–709.
- [5] G.F. Barros, B.P. Cavalari, Y. Kohayakawa, G.O. Mota and T. Naia, Directed graphs with lower orientation Ramsey thresholds, *Rairo-OP*, to appear.
- [6] V. Bergelson, W.A. Deuber and N. Hindman, Rado’s theorem for finite fields, *Proceedings of the Conference on Sets, Graphs, and Numbers*, Budapest, 1991, *Colloq. Math. Soc. János Bolyai* **60** (1992), 77–88.

- [7] V. Bergelson, W. Deuber, N. Hindman and H. Lefmann, Rado’s theorem for commutative rings, *J. Combin. Theory Ser. A* **66(1)** (1994), 68–92.
- [8] B. Bollobás and A.G. Thomason, Threshold functions, *Combinatorica* **7(1)** (1987), 35–38.
- [9] C. Bowtell, R. Hancock and J. Hyde, Proof of the Kohayakawa–Kreuter conjecture for the majority of cases, arXiv:2307.16760.
- [10] S.A. Burr and V. Rosta, On the Ramsey multiplicity of graphs - problems and recent results, *J. Graph Theory* **4(4)** (1980), 347–361.
- [11] J. Byszewski and E. Krawczyk, Rado’s theorem for rings and modules, *J. Combin. Theory Ser. A* **180** (2021), 105402.
- [12] P.J. Cameron, J. Cilleruelo and O. Serra, On monochromatic solutions of equations in groups, *Rev. Mat. Iberoam.* **23(1)** (2007), 385–395.
- [13] M. Christoph, A. Martinsson, R. Steiner and Y. Wigderson, Resolution of the Kohayakawa-Kreuter conjecture, arXiv:2402.03045.
- [14] D. Conlon, Combinatorial theorems relative to a random set, Proceedings of the International Congress of Mathematicians 2014, Vol. 4, 303–328.
- [15] D. Conlon and W.T. Gowers, Combinatorial theorems in sparse random sets, *Ann. Math.* **84** (2016), 367–454.
- [16] B. Cook, Á. Magyar and T. Titichetrakun, A Multidimensional Szemerédi theorem in the primes via combinatorics, *Ann. Combin.* **22** (2018), 711–768.
- [17] K.P. Costello and G. Elvin, Avoiding monochromatic solutions to 3-term equations, *J. Combin.* **14** (2023), 281–304.
- [18] B.A. Datskovsky, On the number of monochromatic Schur triples, *Adv. Appl. Math.* **31(1)** (2003), 193–198.
- [19] W. Deuber, Partition theorems for abelian groups, *J. Combin. Theory Ser. A* **19(1)** (1975), 95–108.
- [20] D. Dong, A. Li and Y. Zhao, Uncommon linear systems of two equations, arXiv:2404.17005.
- [21] J. Fox, H.T. Pham and Y. Zhao, Common and Sidorenko linear equations, *Q. J. Math.* **72(4)** (2021), 1223–1234.
- [22] P. Frankl, R.L. Graham and V. Rödl, Quantitative theorems for regular systems of equations, *J. Combin. Theory Ser. A* **47** (1988), 246–261.
- [23] A. Freschi, PhD thesis, in preparation.
- [24] E. Friedgut, H. Hàn, Y. Person and M. Schacht, A sharp threshold for van der Waerden’s theorem in random subsets, *Discrete Analysis* 2016:7, 19 pp.
- [25] E. Friedgut, E. Kuperwasser, W. Samotij and M. Schacht, Sharp thresholds for Ramsey properties, arXiv:2207.13982.
- [26] E. Friedgut, V. Rödl and M. Schacht, Ramsey properties of random discrete structures, *Random Structures & Algorithms* **37(4)** (2010), 407–436.
- [27] A.W. Goodman, On sets of acquaintances and strangers at any party, *Amer. Math. Monthly* **66(9)** (1959), 778–783.
- [28] R. Graham, V. Rödl and A. Ruciński, On Schur properties of random subsets of integers, *J. Number Theory* **61(2)** (1996), 388–408.
- [29] B. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Anal.* **15** (2005), 340–376.
- [30] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.* **167(2)** (2008), 481–547.
- [31] L. Gugelmann, R. Nenadov, Y. Person, N. Škorić, A. Steger and H. Thomas, Symmetric and asymmetric Ramsey properties in random hypergraphs, *Forum Math. Sigma* **5** (2017), e28.
- [32] R. Hancock, K. Staden and A. Treglown, Independent sets in hypergraphs and Ramsey properties of graphs and the integers, *SIAM J. Discrete Math.* **33(1)** (2019), 153–188.
- [33] R. Hancock and A. Treglown, An asymmetric random Rado theorem for single equations: the 0-statement, *Random Structures & Algorithms* **60(4)** (2022), 529–550.
- [34] D. Hilbert, Über die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten, *J. für die Reine und Angew. Math.* **110** (1892), 104–129.
- [35] J. Hyde, Towards the 0-statement of the Kohayakawa–Kreuter conjecture, *Combin. Probab. Comput.* **32(2)** (2022), 225–268.
- [36] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, 2000.
- [37] N. Kamčev, A. Liebenau and N. Morrison, On uncommon systems of equations, *Israel J. Math.*, to appear.
- [38] Y. Kohayakawa and B. Kreuter, Threshold functions for asymmetric Ramsey properties involving cycles, *Random Structures & Algorithms* **11(3)** (1997), 245–276.
- [39] D. Král’, O. Serra and L. Vena, On the removal lemma for linear systems over abelian groups, *Eur. J. Comb.* **34(2)** (2013), 248–259.
- [40] E. Kuperwasser, W. Samotij and Y. Wigderson, On the Kohayakawa–Kreuter conjecture, arXiv:2307.16611.

- [41] A. Liebenau, L. Mattos, W. Mendonça and J. Skokan, Asymmetric Ramsey properties of random graphs involving cliques and cycles, *Random Structures & Algorithms* **62(4)** (2023), 1035–1055.
- [42] M. Marciniszyn, J. Skokan, R. Spöhel and A. Steger, Asymmetric Ramsey properties of random graphs involving cliques, *Random Structures & Algorithms* **34(4)** (2009), 419–453.
- [43] F. Mousset, R. Nenadov and W. Samotij, Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties, *Combin. Probab. Comput.* **29(6)** (2020), 943–955.
- [44] R. Nenadov and A. Steger, A short proof of the random Ramsey theorem, *Combin. Probab. Comput.* **25(1)** (2016), 130–144.
- [45] R. Rado, Studien zur kombinatorik, *Mathematische Zeitschrift* **36** (1933), 424–470.
- [46] A. Robertson and D. Zeilberger, A 2-coloring of $[1, n]$ can have $(1/22)n^2 + O(n)$ monochromatic Schur triples, but not less!, *Electron. J. Combin.* **5** (1998), #R19.
- [47] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, in *Combinatorics, Paul Erdős is Eighty*, Vol. 1, 317–346, Bolyai Soc. Math. Studies, János Bolyai Math. Soc., Budapest, 1993.
- [48] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures & Algorithms* **5(2)** (1994), 253–270.
- [49] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.
- [50] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers, *Proc. London Math. Soc.* **74(3)** (1997), 481–502.
- [51] J. Rué and C. Spiegel, The Rado multiplicity problem in vector spaces over finite fields, arXiv:2304.00400.
- [52] A. Saad and J. Wolf, Ramsey multiplicity of linear patterns in certain finite abelian groups, *Q. J. Math.* **68(1)** (2017), 125–140.
- [53] D. Saxton and A. Thomason, Hypergraph containers, *Invent. Math.* **201** (2015), 925–992.
- [54] D. Saxton and A. Thomason, Online containers for hypergraphs, with applications to linear equations, *J. Combin. Theory Ser. B* **121** (2016), 248–283.
- [55] T. Schoen, The number of monochromatic Schur triples, *Eur. J. Comb.* **20(8)** (1999), 855–866.
- [56] I. Schur, Über die kongruenz $x^m + y^m \equiv z^m \pmod{p}$, *Jber. Deutsch. Math. Verein.* **25** (1916), 114–117.
- [57] O. Serra and L. Vena, On the number of monochromatic solutions of integer linear systems on Abelian groups, *European J. Combin.* **35** (2014), 459–473.
- [58] C. Spiegel, A note on sparse supersaturation and extremal results for linear homogeneous systems, *Electron. J. Combin.* **24(3)** (2017), #P3.38.
- [59] A.P. Street and E.G. Whitehead, Jr., Group Ramsey Theory, *J. Combin. Theory Ser. A* **17(2)** (1974), 219–226.
- [60] T. Tao, 254A, Notes 4: Some sieve theory, <https://terrytao.wordpress.com>.
- [61] B.L. van der Waerden, Beweis einer baudeutschen vermutung, *Nieuw Arch. Wisk.* **15** (1927), 212–216.
- [62] L. Vena, On the removal lemma for linear configurations in finite abelian groups, arXiv:1505.02280.
- [63] L. Versteegen, Common and Sidorenko equations in Abelian groups, *J. Comb.* **14(1)** (2023), 53–67.

APPENDIX A

A.1. Proof of Lemma 10.1. Recall Lemma 10.1 is the sieve result we used in the proof of Lemma 2.3 in Section 10. Lemma 10.1 is a corollary of the following result from [60].

Lemma 12.6 ([60, Theorem 32]). *Let t and C be fixed natural numbers and let $n \in \mathbb{N}$ with $n \geq 2$. For each prime number $p \leq \sqrt{n}$, let E_p be the union of $\omega(p)$ residue classes modulo p , where $\omega(p) = t$ for all $p \geq C$, and $\omega(p) < p$ for all p . Then for any $\varepsilon > 0$, one has*

$$|[n] \setminus \bigcup_{\substack{p \text{ prime} \\ p \leq \sqrt{n}}} E_p| \leq (2^t t! + \varepsilon) \mathfrak{S} \frac{n}{\log^t n}$$

whenever n is sufficiently large depending on k, C, ε , and where \mathfrak{S} is the singular series

$$\mathfrak{S} := \prod_{\text{prime } p} \left(1 - \frac{1}{p}\right)^{-t} \left(1 - \frac{\omega(p)}{p}\right).$$

We now explain how to deduce Lemma 10.1 from Lemma 12.6.

Proof of Lemma 10.1. Let $k \in \mathbb{N}$ with $k \geq 2$ and let q be a prime number. For each prime $p \geq k$ with $p \neq q$, let E_p be the union of $k - 1$ residue classes modulo p . We want to prove that

$|E(n)| = O(n/\log^{k-1} n)$ where

$$E(n) := [n] \setminus \bigcup_{\substack{p \neq q \text{ prime} \\ k \leq p \leq \sqrt{n}}} E_p.$$

Fix $C := k$, $t := k - 1$, $\varepsilon := 1$ and pick $n \in \mathbb{N}$ sufficiently large so that Lemma 12.6 holds. For every prime $p < k$, let $\omega(p) := 0$ and $E_p := \emptyset$. We have

$$E(n) = [n] \setminus \bigcup_{\substack{p \neq q \text{ prime} \\ p \leq \sqrt{n}}} E_p.$$

Note that if $q < k$ then $E_q = \emptyset$ by definition. In this case set $E'(n) := E(n) \setminus E_q = E(n)$.

If $q \geq k$, let E_q be the union of $k - 1$ residue classes modulo q that maximises $|E'(n)|$ where

$$E'(n) := E(n) \setminus E_q = [n] \setminus \bigcup_{\substack{p \text{ prime} \\ p \leq \sqrt{n}}} E_p.$$

Since the residue classes modulo q induce a partition of $E(n)$, by the maximality of E_q we obtain

$$|E'(n)| = |E(n) \setminus E_q| \geq \left(1 - \frac{k-1}{q}\right) |E(n)| \geq \frac{1}{k} |E(n)|.$$

Applying Lemma 12.6 with respect to parameters t, C, ε and the E_p 's, we obtain $|E'(n)| \leq Tn/\log^{k-1} n$ where T is a constant depending uniquely on k . Thus, $|E(n)| \leq kTn/\log^{k-1} n$ and so indeed $|E(n)| = O(n/\log^{k-1} n)$, as required. \square

A.2. Proof of Theorem 2.2. The proof of Theorem 2.2 is an easy application of the second moment method and it is nearly identical to the analogous proof for arithmetic progressions in the integers (see [36, Example 3.2]). We include it here for completeness.

Proof of Theorem 2.2. Let X be the number of k -APs in $\mathbf{P}_{n,p}$. By Theorem 2.1 there are $\Theta(n^2/\log^k n)$ k -APs in \mathbf{P}_n and so we have $\mathbb{E}(X) = \Theta(p^k n^2/\log^k n)$. If $p = o(n^{-2/k} \log n)$ then $\mathbb{E}(X) = o(1)$. By Markov's inequality, $\mathbb{P}(X \geq 1) \leq \mathbb{E}(X) = o(1)$ and in particular $\mathbb{P}(X = 0) = 1 - o(1)$, as required.

For the 1-statement, pick an arbitrary ordering of the k -APs in \mathbf{P}_n and let X_i be the indicator variable for the i th k -AP to appear in $\mathbf{P}_{n,p}$. Formally,

$$X_i = \begin{cases} 1 & \text{if the } i\text{th } k\text{-AP lies in } \mathbf{P}_{n,p} \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Now we bound the number of pairs of k -APs in \mathbf{P}_n that share exactly one element: first fix a prime $q \in \mathbf{P}_n$ ($O(n/\log n)$ choices) and then pick two k -APs in \mathbf{P}_n that contain q ($O(n^2/\log^{2k-2} n)$ choices by Lemma 2.3). Thus, there are $O(n^3/\log^{2k-1} n)$ such pairs. Furthermore, there are $O(n^2/\log^k n)$ pairs of k -APs in \mathbf{P}_n that share at least two elements (since a k -AP is uniquely determined by two elements and their positions within the k -AP).

Note that $X = \sum_i X_i$. It follows that

$$\text{Var}(X) = \sum_{i,j} \text{Cov}(X_i, X_j) \leq \sum_{i,j} \mathbb{E}(X_i X_j) = O\left(\frac{n^3}{\log^{2k-1} n} p^{2k-1} + \sum_{t=2}^k \frac{n^2}{\log^k n} p^{2k-t}\right),$$

where here the sums are only for i, j (including $i = j$) such that the i th and j th k -APs in \mathbf{P}_n intersect. Chebyshev's inequality yields

$$\mathbb{P}(X = 0) \leq \frac{\text{Var}(X)}{\mathbb{E}(X)^2} = O\left(\frac{\log n}{np} + \sum_{t=2}^k \frac{\log^k n}{n^2 p^t}\right).$$

If $p = \omega(n^{-2/k} \log n)$ the RHS of the equality above is $o(1)$ and so $\mathbb{P}(X = 0) = o(1)$, as required. \square

A.3. Deducing the 1-statement of the random Ramsey theorem from Theorem 7.7. In this section we show that Theorem 7.7 implies the 1-statement of Theorem 4.1.

Let H be as in the statement of Theorem 4.1. Write $E(H) = \{h_1, \dots, h_{e(H)}\}$ and let $r \geq 2$. Without loss of generality we may assume that H has no isolated vertices. For $n \in \mathbb{N}$, let \mathcal{H}_n be the $e(H)$ -uniform ordered hypergraph with vertex set $V(\mathcal{H}_n) := E(K_n)$ where an ordered $e(H)$ -tuple $(e_1, \dots, e_{e(H)}) \in (E(K_n))^{e(H)}$ is an edge of \mathcal{H}_n if and only if there is an injective graph homomorphism $g : H \rightarrow K_n$ such that $g(h_i) = e_i$ for every $i \in [e(H)]$; in other words, there is a copy of H in K_n with edges $e_1, \dots, e_{e(H)}$ where e_i plays the role of h_i for every $i \in [e(H)]$.

For brevity, we write $\mathcal{H} := \mathcal{H}_n$. For $W \subseteq [e(H)]$, let H_W be the subgraph of H spanned by the set of edges $\{h_i : i \in W\} \subseteq E(H)$; so H_W contains no isolated vertices. Recall that the restriction \mathcal{H}_W of \mathcal{H} was defined just after Remark 7.5.

Claim 12.7. *We have $e(\mathcal{H}_W) = \Theta(n^{v(H_W)})$ for every $W \subseteq [e(H)]$.*

Proof of the claim. Say $W = \{t_1, \dots, t_{|W|}\}$ where $t_1 < \dots < t_{|W|}$. Then, provided $n \geq v(H)$, an ordered $|W|$ -tuple $(e_1, \dots, e_{|W|}) \in (E(K_n))^{|W|}$ is an edge of \mathcal{H}_W if and only if there is an injective graph homomorphism $g : H_W \rightarrow K_n$ such that $g(h_{t_i}) = e_i$ for every $i \in [|W|]$. There are $\Theta(n^{v(H_W)})$ such homomorphisms. Conversely, at most $2^{|W|}$ injective homomorphisms $g : H_W \rightarrow K_n$ correspond to the same edge in \mathcal{H}_W .¹⁵ Hence $e(\mathcal{H}_W) = \Theta(n^{v(H_W)})$, as claimed. \blacksquare

We can now compute $\hat{p}(\mathcal{H})$. We have $v(\mathcal{H}) = e(K_n) = \binom{n}{2}$. This and Claim 12.7 yield

$$(39) \quad f_{n,W} = \left(\frac{e(\mathcal{H}_W)}{v(\mathcal{H})}\right)^{-\frac{1}{|W|-1}} = \Theta\left(\left(\frac{n^{v(H_W)}}{n^2}\right)^{-\frac{1}{e(H_W)-1}}\right) = \Theta\left(n^{-\frac{v(H_W)-2}{e(H_W)-1}}\right).$$

Note that $\frac{e(H_W)-1}{v(H_W)-2} = d_2(H_W)$ when $|W| \geq 2$. Thus, we have

$$(40) \quad \max_{\substack{W \subseteq [k] \\ |W| \geq 2}} \frac{e(H_W) - 1}{v(H_W) - 2} \leq \max_{H' \subseteq H} d_2(H') = m_2(H).$$

By assumption, H contains a connected component K with at least two edges. We have $d_2(K) = \frac{e(K)-1}{v(K)-2} \geq 1$ since K is connected. Suppose the value of $m_2(H)$ is achieved by $H' \subseteq H$. Then H' contains at least two edges, since $d_2(H') \geq d_2(K) \geq 1$. Furthermore, H' has no isolated vertices. Indeed, deleting an isolated vertex from H' would yield a subgraph H'' of H with $d_2(H'') > d_2(H')$, a contradiction. Set $W := \{i : h_i \in E(H')\}$. Since H' has no isolated vertices, it follows that $H_W =$

¹⁵More precisely, there are exactly 2^a injective homomorphisms $g : H_W \rightarrow K_n$ that corresponds to the same edge in \mathcal{H}_W , where a is the number of isolated edges in H which lie in H_W .

H' . Furthermore, $|W| \geq 2$ since H' has at least two edges. We have $d_2(H_W) = d_2(H') = m_2(H)$, and so equality holds in (40). That is,

$$\max_{\substack{W \subseteq [k] \\ |W| \geq 2}} \frac{e(H_W) - 1}{v(H_W) - 2} = m_2(H).$$

Using the equality above, taking the maximum over all $W \subseteq [e(H)]$ of size at least 2 in (39) yields

$$\hat{p}(\mathcal{H}) = \Theta\left(n^{-\frac{1}{m_2(H)}}\right).$$

Next, we verify conditions (P1)–(P4).

(P1): Recall we proved that $m_2(H) \geq 1$. This inequality implies $n^{-\frac{1}{m_2(H)}} \rightarrow 0$ and $n^{2-\frac{1}{m_2(H)}} \rightarrow \infty$ as $n \rightarrow \infty$. In particular, $\hat{p}(\mathcal{H}) \rightarrow 0$ and $\hat{p}(\mathcal{H})v(\mathcal{H}) \rightarrow \infty$ as $n \rightarrow \infty$.

(P2): By Ramsey's theorem, there exists an $m \in \mathbb{N}$ such that any r -edge-colouring of K_m yields a monochromatic copy of H . For $n \geq m$, pick an r -edge-colouring of K_n . Every subset of $V(K_n)$ of size m contains a monochromatic copy of H , and each copy of H belongs to $\binom{n-v(H)}{m-v(H)}$ subsets of $V(K_n)$ of size m . Thus, the number of monochromatic copies of H in K_n is at least

$$\frac{\binom{n}{m}}{\binom{n-v(H)}{m-v(H)}} \geq \left(\frac{n}{m}\right)^{v(H)}.$$

Equivalently, given any r -colouring of the vertices of \mathcal{H} , there are at least $(n/m)^{v(H)}$ monochromatic edges in \mathcal{H} . By Claim 12.7 we have $e(\mathcal{H}) = e(\mathcal{H}_{[e(H)]}) = \Theta(n^{v(H)})$. Together, this implies that the sequence of ordered hypergraphs $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is r -supersaturated.

(P3): Let $W \subseteq Y \subseteq [e(H)]$. Given $(e_1, \dots, e_{|W|}) \in \mathcal{H}_W$, there are at most $n^{v(H_Y)-v(H_W)}$ edges in \mathcal{H}_Y that restrict to $(e_1, \dots, e_{|W|})$. This is because each such edge corresponds to an injective graph homomorphism $g : H_Y \rightarrow K_n$ with $g(h_i) = e_i$ for every $i \in W$. For all such homomorphisms, the images of the edges in $E(H_W)$ are fixed, and we have at most $n^{v(H_Y)-v(H_W)}$ choices for the image of $V(H_Y) \setminus V(H_W)$. Thus,

$$\Delta_W(\mathcal{H}_Y) \leq n^{v(H_Y)-v(H_W)}.$$

By Claim 12.7, we have $e(\mathcal{H}_W) = \Theta(n^{v(H_W)})$ and $e(\mathcal{H}_Y) = \Theta(n^{v(H_Y)})$. Combining this with the previous inequality yields

$$\Delta_W(\mathcal{H}_Y) \leq \Theta\left(\frac{e(\mathcal{H}_Y)}{e(\mathcal{H}_W)}\right),$$

and so condition (P3) holds.

(P4): Let $W \subset Y \subseteq [e(H)]$ with $|W| = 1$. Note that $E(\mathcal{H}_W) = V(\mathcal{H})$. This combined with condition (P3) gives

$$\Delta_W(\mathcal{H}_Y) \leq \Theta\left(\frac{e(\mathcal{H}_Y)}{e(\mathcal{H}_W)}\right) = \Theta\left(\frac{e(\mathcal{H}_Y)}{v(\mathcal{H}_W)}\right),$$

and so condition (P4) holds.

Since conditions (P1)–(P4) hold, by Theorem 7.7 we obtain the following 1-statement. There exists some $C > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}_{n,q}^v \text{ is } r\text{-Ramsey}] = 1 \quad \text{if } q \geq Cn^{-\frac{1}{m_2(H)}}.$$

This is equivalent to

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H, r)\text{-Ramsey}] = 1 \quad \text{if } p \geq Cn^{-\frac{1}{m_2(H)}},$$

which is precisely the 1-statement of Theorem 4.1.