INSTITUTO DE MATEMÁTICA PURA E APLICADA

Counting in additive combinatorics

by

Marcelo Campos

January 2020

INSTITUTO DE MATEMÁTICA PURA E APLICADA

Abstract

by Marcelo Campos

The aim of this thesis is to present advances in problems from additive combinatorics that can be approached by counting the number of objects with some specific structure.

In the Chapter 2, we study the number of k-element subsets J of a given abelian group G, such that $|J + J| \leq \lambda|J|$. Proving a conjecture of Alon, Balogh, Morris and Samotij, and improving a result of Green and Morris, who proved the conjecture for λ fixed, we provide an upper bound on the number of such sets which is tight up to a factor of $2^{o(k)}$, when $G = \mathbb{Z}$ and $\lambda = o(k/(\log n)^3)$. We also provide a generalization of this result to arbitrary abelian groups which is tight up to a factor of $2^{o(k)}$ in many cases. The main tool used in the proof is the asymmetric container lemma, introduced recently by Morris, Samotij and Saxton.

In the Chapter 3, joint with Collares, Morris, Morrison and Seixas, we determine the number and typical structure of sets of integers with bounded doubling. In particular, improving recent results of Green and Morris, and of Mazur, we show that the following holds for every fixed $\lambda > 2$ and every $k \geq (\log n)^4$: if $\omega \to \infty$ as $n \to \infty$ (arbitrarily slowly), then almost all sets $A \subset [n]$ with $|A| = k$ and $|A + A| \leq \lambda k$ are contained in an arithmetic progression of length $\lambda k/2 + \omega$.

Chapter 4, joint with Mattos, Morris and Morrison, presents progress in a well-known conjecture that states that a random symmetric $n \times n$ matrix with entries in $\{-1, 1\}$ is singular with probability $\Theta\big(n^2 2^{-n}\big)$. More precisely we prove that the probability of this event is at most $\exp\big(-\Omega(\sqrt{n})\big)$, improving the best known bound of $\exp\big(-\Omega(n^{1/4}\sqrt{\log n})\big)$, which was obtained recently by Ferber and Jain. The main new ingredient is an inverse Littlewood–Offord theorem in $\mathbb{Z}_p^n$ that applies under very mild conditions, whose statement is inspired by the method of hypergraph containers.

Contents

Chapter 1

Introduction

In this thesis we present new results in counting and probabilistic problems in additive combinatorics using combinatorial techniques. Chapters 2 and 3 are concerned with counting and typical structure problems for sets of a given size and doubling constant. While Chapter 4 is concerned with singularity of random symmetric matrices.

1.1 Counting and Typical Structure for sets with given doubling

One of the central objects of interest in additive combinatorics is the sumset

$$A + B := \{a + b : a \in A,\, b \in B\}$$

of two sets $A, B \subset \mathbb{Z}$. If $|A + A| = \lambda |A|$ we say A has *doubling constant* (or sometimes simply *doubling*) λ. A cornerstone of the theory is the celebrated theorem of Freĭman [17, 18] (later reproved by Ruzsa [43]), which states that if $|A + A| \leqslant \lambda |A|$, then A is contained in a generalised arithmetic progression[1] of dimension $O_\lambda(1)$ and size $O_\lambda(|A|)$, where the implicit constants depend only on λ. For an overview of the area, see the book of Tao and Vu [48], or the surveys by Green [19] and Sanders [45].

In Chapters 2 and 3 we will be interested in the number and typical structure of sets with sumset of a given size. The study of this problem was initiated in 2005 by Green [21], who was motivated by applications to random Cayley graphs, and in recent years there has been significant interest in related questions [1, 3, 4, 10, 22]. In particular Alon, Balogh, Morris and Samotij [1] proved a refinement of the Cameron-Erdős conjecture about the number of sum-free subsets of $[n]$, which was solved independently by Green [20] and Sapozhenko [46]. In [1] the

[1] That is, a set of the form $P = \{a + i_1 d_1 + \cdots + i_s d_s : i_j \in \{0, \ldots, k_j\}\}$ for some $a, d_1, \ldots, d_s, k_1, \ldots, k_s \in \mathbb{N}$.

author used an early form of the method of hypergraph containers and also needed to prove a bound on the number of k-sets $A \subset [n]$ with doubling constant λ. They moreover conjectured that the following stronger (and, if true, best possible) bound holds.

Conjecture 1.1.1 (Alon, Balogh, Morris and Samotij)**.** *For every $\delta > 0$, there exists $C > 0$ such that the following holds. If $k \geqslant C \log n$ and if $\lambda \leqslant k/C$, then there are at most*

$$2^{\delta k}\binom{\frac{1}{2}\lambda k}{k}$$

sets $J \subset [n]$ with $|J| = k$ and $|J + J| \leqslant \lambda |J|$.

The conjecture was later confirmed for λ constant by Green and Morris [22]; in fact they proved a slightly more general result: for each fixed λ and as $k \to \infty$, the number of sets $J \subset [n]$ with $|J| = k$ and $|J + J| \leqslant \lambda |J|$ is at most

$$2^{o(k)}\binom{\frac{1}{2}\lambda k}{k} n^{\lfloor \lambda + o(1) \rfloor}.$$

We improve this result in 2 directions, in Chapter 2 we prove Conjecture 1.1.1 for all $\lambda = o\left(k/(\log n)^3\right)$ and in Chapter 3 we obtain bounds tight up to a constant factor for fixed λ.

In order to understand such why results should be true, recall first that, by Freĭman's theorem, a set has bounded doubling if and only if it is a subset of positive density of a generalised arithmetic progression P of bounded dimension. Now, if A were a random subset of P of positive density, then $A+A$ would be unlikely to 'miss' many elements of $P+P$, and this suggests that most sets of bounded doubling should in fact be contained in an arithmetic progression of size roughly $|A + A|/2$. If this intuition was true we should expect to have about

$$\binom{\lambda k/2}{k}$$

choices for A, which is roughly what Conjecture 1.1.1 states. This intuition about the typical structure of A will be confirmed in Chapters 2 and 3, building upon the works of Green, Morris [22] and Mazur []. In Chapter 2 we will show that there typically exists an arithmetic progression P of length $(1/2+o(1))|A+A|$ such that $|A \setminus P| = o(|A|)$, as long as $|A+A| = o\left(|A|^2(\log n)^{-3}\right)$. We prove a more refined result for small doubling in Chapter 3, more precisely we show that typically there exists an arithmetic progression P of lenght $\frac{|A+A|}{2} + \omega$, where $\omega \to \infty$ arbitrarily slow, such that $A \subset P$, as long as $|A + A| = O(|A|)$.

1.2 Random Symmetric Matrices

Let A_n denote a (uniformly-chosen) random $n \times n$ matrix with entries in the set $\{-1, 1\}$. An old and notorious conjecture (see, for example, the discussion in [26]) states that the probability that $\det(A_n) = 0$ is asymptotically equal to the probability that two of the rows or columns of A_n are equal (up to a factor of ± 1), and hence is equal to $\big(1+o(1)\big)n^2 2^{-n+1}$. The first progress on this conjecture was made in 1967, by Komlós [27], who used Erdős' celebrated solution [12] of the Littlewood–Offord problem (see below) to deduce that A_n is singular with probability at most $O(n^{-1/2})$. However, the first exponential bound on the probability was only obtained in 1995, by Kahn, Komlós and Szemerédi [26]. Subsequent improvements were made by Tao and Vu [49] and by Bourgain, Vu and Wood [8], culminating in the recent work of Tikhomirov [52], who proved that

$$\mathbb{P}\big(\det(A_n) = 0\big) = \left(\frac{1}{2} + o(1)\right)^n.$$

In Chapter 4 we will consider the analogous problem for random *symmetric* ± 1 matrices, for which significantly less is known. As in the case of A_n, it is natural to conjecture that such a matrix is singular with probability $\Theta\big(n^2 2^{-n}\big)$; however, it turns out to be extremely difficult even to prove that this probability tends to zero as $n \to \infty$. This problem was apparently first posed by Weiss in the early 1990s (see [9]), but only resolved in 2005, by Costello, Tao and Vu [9], who proved that

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant n^{-1/8+o(1)}, \tag{1.1}$$

where we write M_n for a (uniformly-chosen) random $n \times n$ symmetric matrix with entries in the set $\{-1, 1\}$. The first super-polynomial bound on the probability that M_n is singular, and the first exponential-type bound (i.e., of the form $\exp(-n^c)$ for some $c > 0$), were obtained almost simultaneously, by Nguyen [31] and Vershynin [53], respectively. We remark that the proof in [31] was based on earlier work of Nguyen and Vu [32], which relied on deep results from additive combinatorics, while the proof in [53] built on the earlier breakthroughs of Rudelson and Vershynin [38, 40].

Recently, a new 'combinatorial' approach was introduced by Ferber, Jain, Luh, and Samotij [13], and applied by Ferber and Jain [14] to prove that

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant \exp\big(-cn^{1/4}\sqrt{\log n}\big)$$

for some $c > 0$. In Chapter 4 we use a different combinatorial approach (inspired by the method of [13, 14]) to obtain that

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant \exp\big(-c\sqrt{n}\big).$$

Chapter 2

On the number of sets with a given doubling constant

2.1 Introduction

In this chapter we study the number and typical structure of k-sets with doubling λ, where λ is allowed to be large. Our main theorem confirms Conjecture 1.1.1 for all $\lambda = o(k/(\log n)^3)$.

Theorem 2.1.1. *Let k, n be integers and $2 \leqslant \lambda \leqslant o\big(\frac{k}{(\log n)^3}\big)$. The number of sets $J \subset [n]$ with $|J| = k$ such that $|J + J| \leqslant \lambda |J|$ is at most*

$$2^{o(k)} \binom{\frac{1}{2}\lambda k}{k}.$$

We will in fact prove stronger bounds on the error term than those stated above, see Theorem 2.4.1. Nevertheless, we are unable to prove the conjecture in the range $\lambda = \Omega(k/(\log n)^3)$, and actually the conjecture is false for a certain range of values of k and $\lambda \gg k/\log n$. More precisely in Proposition 2.6.1 we pro that for any integers n, k, and any positive numbers λ, ϵ with $\min\{k, n^{1/2-\epsilon}\} \geqslant \lambda \geqslant \frac{4\log(24C)k}{\epsilon \log n}$, there are at least

$$\binom{\frac{n}{2}}{\frac{\lambda}{4}} \binom{\frac{\lambda k}{8}}{k - \frac{\lambda}{4}} \geqslant \binom{C\lambda k}{k}$$

sets $J \subset [n]$ with $|J| = k$ and $|J + J| \leqslant \lambda k$. The construction[1] is very simple: let P be an arithmetic progression of size $\lambda k/8$ and set $J = J_0 \cup J_1$, where J_0 is any subset of P of size $k - \lambda/4$, and J_1 is any subset of $[n] \setminus P$ of size $\lambda/4$.

Our methods also allow us to characterize the typical structure of an k-set with doubling constant λ, and obtain the following result.

[1] We would like to thank Rob Morris for pointing out this construction.

Theorem 2.1.2. *Let k, n be integers and $2 \leqslant \lambda \leqslant o\big(\frac{k}{(\log n)^3}\big)$. For almost all sets $J \subset [n]$ with $|J| = k$ such that $|J + J| \leqslant \lambda |J|$, there is a set $T \subset J$ such that $J \setminus T$ is contained in an arithmetic progression of size $\frac{1+o(1)}{2}\lambda k$ and $|T| = o(k)$.*

In the case $k = \Omega(n)$ (and hence $\lambda = O(1)$), this result was proved by Mazur [29]. We will provide better bounds for the error terms in Theorem 2.5.1, below.

2.1.1 Abelian Groups

Notice that the doubling constant is defined for finite subsets of any abelian group. So, given a finite subset Y of an abelian group, one might ask: how many subsets of Y of size k with doubling constant λ there are? We are also able to provide an answer to this more general question. From now on, fix an arbitrary abelian group G throughout this chapter. To state our main result formally in the context of general abelian groups we define, for each positive real number t, the quantity $\beta(t)$ to be the size of the biggest subgroup of G of size at most t, that is,

$$\beta(t) = \max\big\{|H| : H \leqslant G,\ |H| \leqslant t\big\}. \tag{2.1}$$

Theorem 2.1.3. *Let k, n be integers, $2 \leqslant \lambda \leqslant o(\frac{k}{(\log n)^3})$, and $Y \subset G$ with $|Y| = n$. The number of sets $J \subset Y$ with $|J| = k$ such that $|J + J| \leqslant \lambda |J|$ is at most*

$$2^{o(k)} \binom{\frac{1}{2}(\lambda k + \beta)}{k},$$

where $\beta := \beta((1 + o(1))\lambda k)$.

Again we will actually prove somewhat stronger (although slightly more convoluted) bounds for Theorem 2.1.3, see Theorem 2.4.1. We remark that Theorem 2.1.3 implies Theorem 2.1.1, since the only finite subgroup of $\mathbb{Z}$ is the trivial one, so in this case $\beta(t) = 1$ for all t. Finally let us remark that Theorem 2.1.3 is best possible in many cases. Indeed suppose for some integers l, m, that the largest subgroup $H \leqslant G$ with $|H| \leqslant m \leqslant |G|$ is of size $\beta = \frac{m}{2l-1}$, then there are at least

$$\binom{\frac{m+\beta}{2}}{k}$$

sets $J \subset G$ of size k such that $|J + J| \leqslant m$. To see this, take an arithmetic progression $P \subset G/H$ of size l (there exists one because of the choice of H) and consider $B = P + H$. Since $|B + B| \leqslant |P + P||H| = m$, for every set $J \subset B$ of size k we have $|J + J| \leqslant |B + B| \leqslant m$. Therefore, there are at least

$$\binom{\frac{lm}{2l-1}}{k} = \binom{\frac{m+\beta}{2}}{k}$$

sets $J \subset B$ of size k with $|J + J| \leqslant m$.

2.1.2 The method of hypergraph containers

Before diving into the proof of the main results, let us briefly mention the main tool used in the proof of Theorem 2.1.1. The method of hypergraph containers, introduced by Balogh, Morris and Samotij [5] and independently by Saxton and Thomason [47], has proven to be a very useful tool in counting problems that involve forbidden structures, for a general overview of the method and its applications see [6]. More recently, Morris, Samotij and Saxton [30] introduced asymmetric containers, a generalization of hypergraph containers for forbidden structures with some sort of asymmetry, and applied the method to give a structural characterization of almost all graphs with a given number of edges free of an induced C_4. A variant of the asymmetric container lemma, which follows essentially from a minor modification of the proof in [30], will be our main tool in this article, we give more details in the next section.

2.2 The Asymmetric Container Lemma

In this section we will state our main tool and give a brief explanation of how we will apply it to our problem. Let $Y \subset G$, with $|Y| = n$, and observe that when trying to count sets $J \subset Y$ with $|J| = k$ and $|J + J| \leqslant \lambda k$, one may instead count sets $J \subset Y$ such that there is a set $I \subset Y$ with $J + J \subset I$ and $|I| \leqslant \lambda k$. Keeping this in mind, the following definition will be useful.

Definition 2.2.1. *Given disjoint copies of $Y+Y$ and Y, namely Y_0, Y_1 respectively, and $A \subset Y_0$ and $B \subset Y_1$, we define $\mathcal{H}(A, B)$ to be the hypergraph with vertex set $V(\mathcal{H}(A, B)) := (Y_0 \setminus A) \cup B$ and edge set*

$$E(\mathcal{H}(A, B)) := \big\{(\{c\}, \{a, b\}) : c \in Y_0 \setminus A,\ a, b \in B,\ a + b = c\big\}.$$

Sometimes when A and B are clear from the context we will denote $\mathcal{H}(A, B)$ simply by $\mathcal{H}$. Notice that $\mathcal{H}(A, B)$ is not uniform since there are edges $(\{c\}, \{a\})$ corresponding to $a + a = c$, but these will not be a problem. The usefulness of Definition 2.2.1 is that now for every pair of sets (I, J) with $J + J \subset I$ we know that $(Y_0 \setminus I) \cup J$ doesn't contain any edges of $\mathcal{H}(A, B)$, so $(Y_0 \setminus I) \cup J$ would usually be called an independent set, but instead we will call the pair (I, J) independent for convenience. Since we have a method for counting what are usually called independent sets in hypergraphs, and each of those is in correspondence to what we call an independent pair, we can obtain a theorem for counting independent pairs.

To state the main tool in this article we will need to go into some more slightly technical definitions. We first define a useful generalization of uniform hypergraphs, that includes the hypergraph presented in Definition 2.2.1. Given disjoint finite sets V_0, V_1 we define an (r_0, r_1)-bounded hypergraph $\mathcal{H}$ on the vertex set $V = V_0 \cup V_1$ to be a set of edges $E(\mathcal{H}) \subset \binom{V_0}{\leqslant r_0} \times \binom{V_1}{\leqslant r_1}$.

Note that the hypergraph in Definition 2.2.1 is $(1,2)$-bounded. Given a pair $(W_0, W_1) \in 2^{V_0} \times 2^{V_1}$, we say (W_0, W_1) *violates* $(e_0, e_1) \in E(\mathcal{H})$ if $e_0 \subset V_0 \setminus W_0$ and $e_1 \subset W_1$. If a set (W_0, W_1) doesn't violate any $(e_0, e_1) \in E(\mathcal{H})$ then we call (W_0, W_1) *independent* with respect to $\mathcal{H}$. Let $\mathcal{F}_{\leqslant m}(\mathcal{H}) \subset 2^{V(\mathcal{H})}$ be the family of independent pairs (W_0, W_1) such that $|W_0| \leqslant m$, and observe that for any pair of sets (I, J), with $|I| \leqslant m$ and $J + J \subset I$, we have $(I, J) \in \mathcal{F}_{\leqslant m}(\mathcal{H}(\emptyset, Y))$. We define the codegree $d_{(L_0,L_1)}(\mathcal{H})$ of $L_0 \subset V_0,\ L_1 \subset V_1$ to be the size of the set

$$\{(e_0, e_1) \in E(\mathcal{H}) :\ L_0 \subset e_0,\ L_1 \subset e_1\}$$

and we define the maximum (ℓ_0, ℓ_1)-codegree of $\mathcal{H}$ to be

$$\Delta_{(\ell_0,\ell_1)} := \max\{d_{(L_0,L_1)}(\mathcal{H}) :\ L_0 \subset V_0,\ L_1 \subset V_1,\ |L_0| = \ell_0,\ |L_1| = \ell_1\}.$$

With all of this in mind we introduce a variant of the asymmetric container lemma of Morris, Samotij and Saxton [30] that we can, once we have suitable supersaturation theorem to check the codegree condition, apply iteratively and prove Theorem 2.1.1.

Theorem 2.2.2. *For all non-negative integers r_0, r_1, not both zero, and each $R > 0$, the following holds. Suppose that $\mathcal{H}$ is a non-empty (r_0, r_1)-bounded hypergraph with $V(\mathcal{H}) = V_0 \cup V_1$, and b, m, and q are integers with $b \leqslant \min\{m, |V_1|\}$, satisfying*

$$\Delta_{(\ell_0,\ell_1)}(\mathcal{H}) \leqslant R\frac{b^{\ell_0+\ell_1-1}}{m^{\ell_0}|V_1|^{\ell_1}}e(\mathcal{H})\left(\frac{m}{q}\right)^{1[\ell_0>0]} \tag{2.2}$$

for every pair $(\ell_0, \ell_1) \in \{0, 1, \dots, r_0\} \times \{0, 1, \dots, r_1\} \setminus \{(0, 0)\}$. Then there exists a family $\mathcal{S} \subset \binom{V_0}{\leqslant r_0 b} \times \binom{V_1}{\leqslant r_1 b}$ and functions $f\colon \mathcal{S} \to 2^{V_0} \times 2^{V_1}$ and $g\colon \mathcal{F}_{\leqslant m}(\mathcal{H}) \to \mathcal{S}$, such that, letting $\delta = 2^{-(r_0+r_1+1)(r_0+r_1)}R^{-1}$:

(i) If $f(g(I, J)) = (A, B)$ with $A \subset V_0$ and $B \subset V_1$, then $A \subset I$ and $J \subset B$.

(ii) For every $(A, B) \in f(\mathcal{S})$ either $|A| \geqslant \delta q$ or $|B| \leqslant (1 - \delta)|V_1|$.

(iii) If $g(I, J) = (S_0, S_1)$ and $f(g(I, J)) = (A, B)$ then $S_0 \subset V_0 \setminus I$ and $S_1 \subset J$, and $|S_0| > 0$ only if $|A| \geqslant \delta q$.

Let us remark that the main difference between this statement of the asymmetric container lemma and the one in [30] is that we partition the vertex set in two parts and treat them differently, which is essential in our application. More specifically, we will apply the container lemma iteratively in such a way that V_1 will shrink much more than V_0, and to account for this imbalance we must differentiate between the two sets of the partition. Another small difference is that the hypergraph $\mathcal{H}$ doesn't need to be uniform. Finally we observe that if S_0 is non-empty, where $g(I, J) = (S_0, S_1)$, then we must have $|A| \geqslant \delta q$, where $f(g(I, J)) = (A, B)$. The proof given below is essentially identical to that presented in [30] with some adaptations to the

notation. We would like to thank the authors of [30] for allowing us to reproduce their proof in this appendix.

2.2.1 Setup

Let r_0 and r_1 be non-negative integers and let R be a positive real. Let b, m, and r be positive integers and suppose that $\mathcal{H}$ is a (r_0, r_1)-bounded hypergraph[2] with vertex set $V = (V_0, V_1)$ satisfying (2.2) for each pair (ℓ_0, ℓ_1) and $b \leqslant \min\{m, |V_1|\}$ as in the statement of Theorem 2.2.2. We claim that, without loss of generality, denoting from now on $v_0(\mathcal{H}) = |V_0|$ and $v_1(\mathcal{H}) = |V_1|$, we may assume that $m \leqslant v_0(\mathcal{H})$. Indeed, if $m > v_0(\mathcal{H})$, then we may replace m with $v_0(\mathcal{H})$ as $\mathcal{F}_{\leqslant m} \subseteq \mathcal{F}(\mathcal{H}) = \mathcal{F}_{\leqslant v_0(\mathcal{H})}(\mathcal{H})$ and the right-hand side of (2.2) is a non-increasing function of m. We shall be working only with hypergraphs whose uniformities come from the set

$$\mathcal{U} := \big\{(1,0), (2,0), \ldots, (r_0,0), (r_0,1), \ldots, (r_0,r_1)\big\}.$$

The maximum codegrees we must check for each uniformity will come from the set

$$\mathcal{V}(i_0,i_1) := \big\{0,1,\ldots,i_0\big\} \times \big\{0,1,\ldots,i_1\big\} \setminus \{(0,0)\}.$$

We now define a collection of numbers that will be upper bounds on the maximum degrees of the hypergraphs constructed by our algorithm. To be more precise, for each $(i_0,i_1) \in \mathcal{U}$ and all $(\ell_0,\ell_1) \in \mathcal{V}(i_0,i_1)$, we shall force the maximum (ℓ_0,ℓ_1)-degree of the (i_0,i_1)-uniform hypergraph not to exceed the quantity $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$, defined as follows.

Definition 2.2.3. *For every $(i_0,i_1) \in \mathcal{U}$ and every $(\ell_0,\ell_1) \in \mathcal{V}(i_0,i_1)$, we define the number $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$ using the following recursion:*

(1) Set $\Delta^{(r_0,r_1)}_{(\ell_0,\ell_1)} := \Delta_{(\ell_0,\ell_1)}(\mathcal{H})$ for all $(\ell_0,\ell_1) \in \mathcal{V}(r_0,r_1)$.

(2) If $i_0 = r_0$ and $0 \leqslant i_1 < r_1$, then

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} := \max\left\{2 \cdot \Delta^{(i_0,i_1+1)}_{(\ell_0,\ell_1+1)},\ \frac{b}{v_1(\mathcal{H})} \cdot \Delta^{(i_0,i_1+1)}_{(\ell_0,\ell_1)}\right\}.$$

(3) If $0 < i_0 < r_0$ and $i_1 = 0$, then

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} := \max\left\{2 \cdot \Delta^{(i_0+1,i_1)}_{(\ell_0+1,\ell_1)},\ \frac{b}{m} \cdot \Delta^{(i_0+1,i_1)}_{(\ell_0,\ell_1)}\right\}.$$

The above recursive definition will be convenient in some parts of the analysis. In other parts, we shall require the following explicit formula for $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$, which one easily derives from Definition 2.2.3 using a straightforward induction on $r_0 + r_1 - i_0 - i_1$.

[2] We remark that from now on all hypergraphs are allowed to have multi-edges, and the edges are counted with multiplicity.

Observation 2.2.4. *For all i_0, i_1, ℓ_0, and ℓ_1 as in Definition 2.2.3,*

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} = \max\left\{2^{d_0+d_1}\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-i_1-d_1}\left(\frac{b}{m}\right)^{r_0-i_0-d_0}\Delta_{(\ell_0+d_0,\ell_1+d_1)}(\mathcal{H}) : 0 \leqslant d_j \leqslant r_j - i_j\right\}.$$

For future reference, we note the following two simple corollaries of Observation 2.2.4 and our assumptions on the maximum degrees of $\mathcal{H}$, see (2.2). Suppose that $(i_0,i_1) \in \mathcal{U}$. If $i_1 > 0$, then necessarily $i_0 = r_0$ and hence,

$$\begin{aligned}\Delta^{(i_0,i_1)}_{(0,1)} &\leqslant \max\left\{2^{d_1}\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-i_1-d_1} R\cdot\frac{b^{d_1}}{v_1(\mathcal{H})^{d_1+1}}\cdot e(\mathcal{H}) : 0 \leqslant d_1 \leqslant r_1 - i_1\right\}\\ &\leqslant 2^{r_1}R\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-i_1}\frac{e(\mathcal{H})}{v_1(\mathcal{H})} = 2^{r_1}R\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-i_1}\left(\frac{b}{m}\right)^{r_0-i_0}\frac{e(\mathcal{H})}{v_1(\mathcal{H})}.\end{aligned} \tag{2.3}$$

Moreover, if $i_0 > 0$ and $i_1 = 0$, then

$$\begin{aligned}\Delta^{(i_0,i_1)}_{(1,0)} &\leqslant \max\left\{2^{d_0+d_1}\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-d_1}\left(\frac{b}{m}\right)^{r_0-i_0-d_0} R\cdot\frac{b^{d_0+d_1}}{m^{d_0}\cdot v_1(\mathcal{H})^{d_1}}\cdot\frac{e(\mathcal{H})}{q}\right\}\\ &\leqslant 2^{r_0+r_1}R\left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1}\left(\frac{b}{m}\right)^{r_0-i_0}\frac{e(\mathcal{H})}{q},\end{aligned} \tag{2.4}$$

where the maximum is over all pairs (d_0,d_1) of integers satisfying $0 \leqslant d_j \leqslant r_j - i_j$.

We will build a sequence of hypergraphs with decreasing uniformity, starting with $\mathcal{H}$, and making sure that, for each hypergraph $\mathcal{G}$ in the sequence, we have an appropriate bound on its maximum codegrees. To this end we define the following set of pairs with large codegree.

Definition 2.2.5. *Given $(i_0,i_1) \in \mathcal{U}$, $(\ell_0,\ell_1) \in \mathcal{V}(i_0,i_1)$, and an (i_0,i_1)-uniform hypergraph $\mathcal{G}$, we define*

$$M^{(i_0,i_1)}_{(\ell_0,\ell_1)}(\mathcal{G}) = \left\{(T_0,T_1) \in \binom{V(\mathcal{G})}{\ell_0}\times\binom{V(\mathcal{G})}{\ell_1} : \deg_{\mathcal{G}}(T_0,T_1) \geqslant \frac{1}{2}\cdot\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}\right\}.$$

Finally, let us say that $c \in \{0,1\}$ is *compatible* with $(i_0,i_1) \in \mathcal{U}$ if the unique pair $(i'_0,i'_1) \in \mathcal{U}\cup\{(0,0)\}$ with $i'_0+i'_1 = i_0+i_1-1$ satisfies $i'_c = i_c - 1$ (and $i'_{1-c} = i_{1-c}$). By the definition of $\mathcal{U}$, it follows that $c = 1$ for $(i_0,i_1) \in \mathcal{U}$ if and only if $i_1 > 0$.

2.2.2 The algorithm

We shall now define precisely a single round of the algorithm we use to prove the container lemma. To this end, fix some $(i_0,i_1) \in \mathcal{U}$ and a compatible $c \in \{0,1\}$ and (as in the definition of a compatible c) set

$$i'_c = i_c - 1 \qquad \text{and} \qquad i'_{1-c} = i_{1-c}. \tag{2.5}$$

Suppose that $\mathcal{G}$ is an (i_0, i_1)-bounded hypergraph with $V(\mathcal{G}) = V(\mathcal{H})$. A single round of the algorithm takes as input an arbitrary $(I, J) \in \mathcal{F}(\mathcal{G})$ and outputs an (i_0', i_1')-bounded hypergraph $\mathcal{G}^*$ satisfying $V(\mathcal{G}^*) = V(\mathcal{G})$ and $(I, J) \in \mathcal{F}(\mathcal{G}^*)$ as well as a set of vertices S of $\mathcal{G}$ such that $|S| \leqslant b$ and either $S \subset J$ or $S \subset V_0 \setminus I$. Crucially, the number of possible outputs of the algorithm (over all possible inputs $(I, J) \in \mathcal{F}(\mathcal{G})$) is at most $\binom{v_c(\mathcal{H})}{\leqslant b}$.

Assume that there is an implicit linear order $\preccurlyeq$ on $V(\mathcal{G})$. The *c-maximum vertex* of a hypergraph $\mathcal{A}$ with $V(\mathcal{A}) = V(\mathcal{G})$ is the $\preccurlyeq$-smallest vertex among those v that maximise $|\{(A_0, A_1) \in \mathcal{A} : v \in A_c\}|$.

The algorithm. Set $\mathcal{A}^{(0)} := \mathcal{G}$, let S be the empty set, and let $\mathcal{G}_*^{(0)}$ be the empty (i_0', i_1')-bounded hypergraph on $V(\mathcal{G})$. Do the following for each integer $j \geqslant 0$ in turn:

(S1) If $|S| = b$ or $\mathcal{A}^{(j)}$ is empty, then set $J := j$ and `STOP`.

(S2) Let $v_j \in V_c$ be the c-maximum vertex of $\mathcal{A}^{(j)}$.

(S3) If $c = 0$ and $v_j \notin I$ or $c = 1$ and $v_j \in J$, then add j to the set S and let

$$\mathcal{G}_*^{(j+1)} := \mathcal{G}_*^{(j)} \cup \Big\{ \big(A_0 \setminus \{v_j\}, A_1 \setminus \{v_j\}\big) : (A_0, A_1) \in \mathcal{A}^{(j)} \text{ and } v_j \in A_c \Big\}.$$

(S4) Let $\mathcal{A}^{(j+1)}$ be the hypergraph obtained from $\mathcal{A}^{(j)}$ by removing from it all pairs (A_0, A_1) such that either of the following hold:

(a) $v_j \in A_c$;

(b) there exist $T_0 \subseteq A_0$ and $T_1 \subseteq A_1$, not both empty, such that

$$(T_0, T_1) \in M_{(\ell_0,\ell_1)}^{(i_0',i_1')}\big(\mathcal{G}_*^{(j+1)}\big)$$

for some $(\ell_0, \ell_1) \in \mathcal{V}(i_0', i_1')$.

Finally, set $\mathcal{A} := \mathcal{A}^{(L)}$ and $\mathcal{G}_* := \mathcal{G}_*^{(L)}$. Moreover, set

$$W := \big\{0, \dots, L-1\big\} \setminus S = \Big\{ j \in \big\{0, \dots, L-1\big\} : v_j \notin V_0 \setminus I \text{ and } v_j \notin J \Big\}.$$

Observe that the algorithm always stops after at most $v(\mathcal{G})$ iterations of the main loop. Indeed, since all constraints (A_0, A_1) with $v_j \in A_c$ are removed from $\mathcal{A}^{(j+1)}$ in part (a) of step (S4), the vertex v_j cannot be the c-maximum vertex of any $\mathcal{A}^{(j')}$ with $j' > j$ and hence the map $\{0, \dots, L-1\} \ni j \mapsto v_j \in V(\mathcal{G})$ is injective.

2.2.3 The analysis

We shall now establish some basic properties of the algorithm described in the previous subsection. To this end, let us fix some $(i_0, i_1) \in \mathcal{U}$ and a compatible $c \in \{0,1\}$ and let i_0' and i_1' be the numbers defined in (2.5). Moreover, suppose that $\mathcal{G}$ is an (i_0, i_1)-bounded hypergraph and that we have run the algorithm with input $(I, J) \in \mathcal{F}(\mathcal{G})$ and obtained the (i_0', i_1')-bounded hypergraph $\mathcal{G}_*$, the integer L, the injective map $\{0, \dots, L-1\} \ni j \mapsto v_j \in V(\mathcal{G})$, and the partition of $\{0, \dots, L-1\}$ into S and W such that $v_j \in J$ or $v_j \in V_0 \setminus I$ if and only if $j \in S$. We first state two straightforward, but fundamental, properties of the algorithm.

Observation 2.2.6. *If $(I, J) \in \mathcal{F}(\mathcal{G})$, then $(I, J) \in \mathcal{F}(\mathcal{G}_*)$.*

Proof. Observe that $\mathcal{G}_*$ contains only constraints of the form:

(*i*) $(A_0 \setminus \{v\}, A_1)$, where $v \in A_0$ and $v \in V_0 \setminus I$, or

(*ii*) $(A_0, A_1 \setminus \{v\})$, where $v \in A_1$ and $v \in J$,

where $(A_0, A_1) \in \mathcal{G}$, see (S3). Hence, if (I, J) violated a constraint of type (*i*) (resp. (*ii*)) then (I, J) would also violate the constraint (A_0, A_1), as $v \in V_0 \setminus I$ (resp. $v \in J$). □

The next observation says that if the algorithm applied to two pairs (I, J) and (I', J') outputs the same set $\{v_j : j \in S\}$, then the rest of the output is also the same.

Observation 2.2.7. *Fix the hypergraph $\mathcal{G}$ we input in the algorithm, suppose that the algorithm applied to $(I', J') \in \mathcal{F}(\mathcal{G})$ outputs a hypergraph $\mathcal{G}_*'$, an integer L', a map $j \mapsto v_j'$, and a partition of $\{0, \dots, L'-1\}$ into S' and W'. If $\{v_j : j \in S\} = \{v_j' : j \in S'\}$, then $\mathcal{G}_* = \mathcal{G}_*'$, $L = L'$, $v_j = v_j'$ for all j, and $W = W'$.*

Proof. The only step of the algorithm that depends on the input pair (I, J) is (S3). There, an index j is added to the set S if and only if $v_j \in V_0 \setminus I$ or $v_j \in J$. Therefore, the execution of the algorithm depends only on the set $\{v_j : j \in S\}$ and the hypergraph $\mathcal{G}$. □

The next two lemmas will allow us to maintain suitable upper and lower bounds on the degrees and densities of the hypergraphs obtained by applying the algorithm iteratively. The first lemma, which is the easier of the two, states that if all the maximum degrees of $\mathcal{G}$ are appropriately bounded, then all the maximum degrees of $\mathcal{G}_*$ are also appropriately bounded.

Lemma 2.2.8. *Given $(\ell_0, \ell_1) \in \mathcal{V}(i_0, i_1)$ and $\ell_c > 0$, set $\ell_c' = \ell_c - 1$ and $\ell_{1-c}' = \ell_{1-c}$. If $\Delta_{(\ell_0,\ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0,\ell_1)}^{(i_0,i_1)}$, then $\Delta_{(\ell_0',\ell_1')}(\mathcal{G}_*) \leqslant \Delta_{(\ell_0',\ell_1')}^{(i_0',i_1')}$.*

Proof. Suppose (for a contradiction) that there exist sets T_0' and T_1', with $|T_0'| = \ell_0'$ and $|T_1'| = \ell_1'$, such that $\deg_{\mathcal{G}_*}(T_0', T_1') > \Delta^{(i_0',i_1')}_{(\ell_0',\ell_1')}$. Let j be the smallest integer satisfying

$$\deg_{\mathcal{G}_*^{(j+1)}}(T_0', T_1') > \frac{1}{2} \cdot \Delta^{(i_0',i_1')}_{(\ell_0',\ell_1')}$$

and note that $j \geqslant 0$, since $\mathcal{G}_*^{(0)}$ is empty. We claim first that

$$\deg_{\mathcal{G}_*}(T_0', T_1') = \deg_{\mathcal{G}_*^{(j+1)}}(T_0', T_1'). \tag{2.6}$$

Indeed, observe that $(T_0', T_1') \in M^{(i_0',i_1')}_{(\ell_0',\ell_1')}\big(\mathcal{G}_*^{(j+1)}\big)$, and therefore the algorithm removes from $\mathcal{A}^{(j)}$ (when forming $\mathcal{A}^{(j+1)}$ in step (S4)) all pairs (A_0, A_1) such that $T_0' \subseteq A_0$ and $T_1' \subseteq A_1$. As a consequence, no further pairs (A_0', A_1') with $T_0' \subseteq A_0'$ and $T_1' \subseteq A_1'$ are added to $\mathcal{G}_*$ in step (S3).

We next claim that

$$\deg_{\mathcal{G}_*^{(j+1)}}(T_0', T_1') - \deg_{\mathcal{G}_*^{(j)}}(T_0', T_1') \leqslant \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}. \tag{2.7}$$

To see this, recall that when we extend $\mathcal{G}_*^{(j)}$ to $\mathcal{G}_*^{(j+1)}$ in step (S3), we only add pairs $\big(A_0 \setminus \{v_j\}, A_1 \setminus \{v_j\}\big)$ such that $(A_0, A_1) \in \mathcal{A}^{(j)} \subseteq \mathcal{G}$ and $v_j \in A_c$. Therefore, setting $T_c = T_c' \cup \{v_j\}$ and $T_{1-c} = T_{1-c}'$, we have

$$\deg_{\mathcal{G}_*^{(j+1)}}(T_0', T_1') - \deg_{\mathcal{G}_*^{(j)}}(T_0', T_1') \leqslant \deg_{\mathcal{G}}(T_0, T_1) \leqslant \Delta_{(\ell_0,\ell_1)}(\mathcal{G}) \leqslant \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)},$$

where the last inequality is by our assumption, as claimed.

Combining (2.6) and (2.7), it follows immediately that

$$\deg_{\mathcal{G}_*}(T_0', T_1') \leqslant \frac{1}{2} \cdot \Delta^{(i_0',i_1')}_{(\ell_0',\ell_1')} + \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} \leqslant \Delta^{(i_0',i_1')}_{(\ell_0',\ell_1')},$$

where the final inequality holds by Definition 2.2.3. This contradicts our choice of (T_0', T_1') and therefore the lemma follows. □

We are now ready for the final lemma, which is really the heart of the matter. We will show that if $\mathcal{G}$ has sufficiently many edges and all of the maximum degrees of $\mathcal{G}$ are appropriately bounded, then either the output hypergraph $\mathcal{G}_*$ has sufficiently many edges, or we either have a big set $W \subset V_1 \setminus J$, or we have a big set $W \subset I$. We remark that here we shall use the assumption that $|I| \leqslant m$.

Lemma 2.2.9. *Suppose that $|I| \leqslant m$ and let $\alpha > 0$. If*

(A1) $e(\mathcal{G}) \geqslant \alpha \cdot \big(\frac{b}{v_1(\mathcal{H})}\big)^{r_1 - i_1}\big(\frac{b}{m}\big)^{r_0 - i_0} e(\mathcal{H})$ *and*

(A2) $\Delta_{(\ell_0,\ell_1)}(\mathcal{G}) \leqslant \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$ *for every* $(\ell_0, \ell_1) \in \mathcal{V}(i_0, i_1)$,

then at least one of the following statements is true:

(P1) $e(\mathcal{G}_*) \geqslant 2^{-i_0-i_1-1}\alpha \cdot \left(\frac{b}{v_1(\mathcal{H})}\right)^{r_1-i_1'}\left(\frac{b}{m}\right)^{r_0-i_0'} e(\mathcal{H})$.

(P2) $c = 1$ *and* $|W| \geqslant 2^{-r_1-1}R^{-1}\alpha \cdot v_1(\mathcal{H})$.

(P3) $c = 0$ *and* $|W| \geqslant 2^{-r_0-r_1-1}R^{-1}\alpha \cdot q$.

Proof. Suppose first that $c = 0$ and observe that[3]

$$e(\mathcal{G}_*) = \sum_{j\in S}\Big(e(\mathcal{G}_*^{(j+1)}) - e(\mathcal{G}_*^{(j)})\Big) = \sum_{j\in S} \deg_{\mathcal{A}^{(j)}}(\{v_j\},\emptyset), \tag{2.8}$$

since $e(\mathcal{G}_*^{(j+1)}) - e(\mathcal{G}_*^{(j)}) = \deg_{\mathcal{A}^{(j)}}(\{v_j\},\emptyset)$ for each $j \in S$ and $\mathcal{G}_*^{(j+1)} = \mathcal{G}_*^{(j)}$ for each $j \notin S$. To bound the right-hand side of (2.8), we count the edges removed from $\mathcal{A}^{(j)}$ in (a) and (b) of step (S4), which gives

$$e(\mathcal{A}^{(j)}) - e(\mathcal{A}^{(j+1)}) \leqslant \deg_{\mathcal{A}^{(j)}}(\{v_j\},\emptyset) + \sum_{(\ell_0,\ell_1)} \big|M_{(\ell_0,\ell_1)}^{(i_0',i_1')}(\mathcal{G}_*^{(j+1)}) \setminus M_{(\ell_0,\ell_1)}^{(i_0',i_1')}(\mathcal{G}_*^{(j)})\big| \cdot \Delta_{(\ell_0,\ell_1)}(\mathcal{G}).$$

Summing over $j \in \{0,\dots,L-1\}$, it follows (using (2.8)) that

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta_{(1,0)}(\mathcal{G}) + \sum_{(\ell_0,\ell_1)} \big|M_{(\ell_0,\ell_1)}^{(i_0',i_1')}(\mathcal{G}_*)\big| \cdot \Delta_{(\ell_0,\ell_1)}^{(i_0,i_1)},$$

since $\mathcal{A} = \mathcal{A}^{(L)} \subseteq \ldots \subseteq \mathcal{A}^{(0)} = \mathcal{G}$ and $\Delta_{(\ell_0,\ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0,\ell_1)}^{(i_0,i_1)}$ by (A2). Observe also that if $c = 1$, then we obtain an identical bound, with $\Delta_{(1,0)}(\mathcal{G})$ replaced by $\Delta_{(0,1)}(\mathcal{G})$.

In order to discuss both cases simultaneously, we set $\chi(0) = (1,0)$ and $\chi(1) = (0,1)$. Observe that

$$\Delta_{\chi(c)}(\mathcal{A}) \leqslant \Delta_{\chi(c)}(\mathcal{A}^{(j)}) \leqslant \Delta_{\chi(c)}(\mathcal{G}) \leqslant \Delta_{\chi(c)}^{(i_0,i_1)}, \tag{2.9}$$

since $\mathcal{A} \subseteq \mathcal{A}^{(j)} \subseteq \mathcal{G}$ and $\mathcal{G}$ satisfies (A2). It follows that, for both $c \in \{0,1\}$,

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta_{\chi(c)}^{(i_0,i_1)} + \sum_{(\ell_0,\ell_1)} \big|M_{(\ell_0,\ell_1)}^{(i_0',i_1')}(\mathcal{G}_*)\big| \cdot \Delta_{(\ell_0,\ell_1)}^{(i_0,i_1)}. \tag{2.10}$$

Now, recall that v_j is the c-maximum vertex of $\mathcal{A}^{(j)}$ and observe that therefore, by (2.8) and (2.9),

$$e(\mathcal{G}_*) = \sum_{j\in S} \Delta_{\chi(c)}\big(\mathcal{A}^{(j)}\big) \geqslant |S| \cdot \Delta_{\chi(c)}(\mathcal{A}) = b \cdot \Delta_{\chi(c)}(\mathcal{A}), \tag{2.11}$$

where the equality is due to the fact that $|S| \neq b$ only when $\mathcal{A}$ is empty, see step (S1).

Next, to bound the sum in (2.10), observe that, by Definition 2.2.5, we have

$$\big|M_{(\ell_0,\ell_1)}^{(i_0',i_1')}(\mathcal{G}_*)\big| \cdot \frac{1}{2} \cdot \Delta_{(\ell_0,\ell_1)}^{(i_0',i_1')} \leqslant \sum_{|T_0|=l_0,|T_1|=l_1} \deg_{\mathcal{G}_*}(T_0,T_1) \leqslant \binom{i_0'}{\ell_0}\binom{i_1'}{\ell_1} \cdot e(\mathcal{G}_*)$$

[3]Recall that $\mathcal{G}_*$ (and $\mathcal{G}_*^{(j)}$ etc.) are multi-hypergraphs and that edges are counted with multiplicity.

for each $(\ell_0,\ell_1)\in\mathcal{V}(i_0',i_1')$ and therefore

$$\begin{aligned}\sum_{(\ell_0,\ell_1)\in\mathcal{V}(i_0',i_1')}\left|M^{(i_0',i_1')}_{(\ell_0,\ell_1)}(\mathcal{G}_*)\right|\cdot\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}&\leqslant 2\cdot\sum_{(\ell_0,\ell_1)}\binom{i_0'}{\ell_0}\binom{i_1'}{\ell_1}\cdot e(\mathcal{G}_*)\cdot\left(\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\right)\\&\leqslant 2\cdot\left(2^{i_0'+i_1'}-1\right)\cdot e(\mathcal{G}_*)\cdot\max_{(\ell_0,\ell_1)}\left\{\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\right\}.\end{aligned}\tag{2.12}$$

We claim that $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\leqslant m/b$ if $c=0$ and $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\leqslant v_1(\mathcal{H})/b$ if $c=1$. Indeed, both inequalities following directly from Definition 2.2.3, since if $c=0$, then $(i_0',i_1')=(i_0-1,i_1)$, and if $c=1$, then $(i_0',i_1')=(i_0,i_1-1)$. We split the remainder of the proof into two cases, depending on the value of c.

Suppose first that $c=1$ and observe that substituting (2.12) into (2.10) yields, using the bound $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\leqslant v_1(\mathcal{H})/b$,

$$e(\mathcal{G})-e(\mathcal{A})\leqslant e(\mathcal{G}_*)+|W|\cdot\Delta^{(i_0,i_1)}_{(0,1)}+2\cdot\left(2^{i_0'+i_1'}-1\right)\cdot e(\mathcal{G}_*)\cdot\frac{v_1(\mathcal{H})}{b}.\tag{2.13}$$

Moreover, by (2.11), and since $i_1\geqslant 1$ when $c=1$, we have

$$\frac{e(\mathcal{G}_*)}{b}\geqslant\Delta_{(0,1)}(\mathcal{A})\geqslant\frac{i_1\cdot e(\mathcal{A})}{v_1(\mathcal{A})}\geqslant\frac{e(\mathcal{A})}{v_1(\mathcal{H})},\tag{2.14}$$

since the maximum degree of a hypergraph is at least as large as its average degree. Combining (2.13) and (2.14), we obtain

$$\begin{aligned}e(\mathcal{G})&\leqslant e(\mathcal{G}_*)\cdot\frac{v_1(\mathcal{H})}{b}\cdot\left(\frac{b}{v_1(\mathcal{H})}+1+2^{i_0'+i_1'+1}-2\right)+|W|\cdot\Delta^{(i_0,i_1)}_{(0,1)}\\&\leqslant e(\mathcal{G}_*)\cdot\frac{v_1(\mathcal{H})}{b}\cdot 2^{i_0+i_1}+|W|\cdot\Delta^{(i_0,i_1)}_{(0,1)},\end{aligned}\tag{2.15}$$

since $b\leqslant v_1(\mathcal{H})$. Now, if the first summand on the right-hand side of (2.15) exceeds $e(\mathcal{G})/2$, then (A1) implies (P1), since $(i_0',i_1')=(i_0,i_1-1)$. Otherwise, the second summand is at least $e(\mathcal{G})/2$ and by (A1) and (2.3),

$$|W|\geqslant\frac{e(\mathcal{G})}{2\cdot\Delta^{(i_0,i_1)}_{(0,1)}}\geqslant\frac{\alpha}{2^{r_1+1}R}\cdot v_1(\mathcal{H}),$$

which is (P2).

The case $c=0$ is slightly more delicate; in particular, we will finally use our assumption that $|I|\leqslant m$. Observe first that if $c=0$, then $i_1=0$ and substituting (2.12) into (2.10) yields, using the bound $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i_0',i_1')}_{(\ell_0,\ell_1)}\leqslant m/b$,

$$e(\mathcal{G})-e(\mathcal{A})\leqslant e(\mathcal{G}_*)+|W|\cdot\Delta^{(i_0,i_1)}_{(1,0)}+\left(2^{i_0+i_1}-2\right)\cdot e(\mathcal{G}_*)\cdot\frac{m}{b},\tag{2.16}$$

cf. (2.13). We claim that

$$\frac{e(\mathcal{G}_*)}{b}\geqslant\Delta_{(1,0)}(\mathcal{A})\geqslant\frac{e(\mathcal{A})}{m}.\tag{2.17}$$

The first inequality follows from (2.11), so we only need to prove the second inequality. To do so, observe that $\mathcal{G}$ is an $(i_0, 0)$-uniform hypergraph (since $c = 0$) and therefore for each pair $(I, J) \in \mathcal{F}(\mathcal{G})$ we must have $I \cap A_0 \neq \emptyset$ for every $(A_0, \emptyset) \in \mathcal{G}$. Now, recall that $(I, J) \in \mathcal{F}(\mathcal{G})$, that $\mathcal{A} \subseteq \mathcal{G}$, and that $|I| \leqslant m$. It follows that $e(\mathcal{A}) \leqslant m \cdot \Delta_{(1,0)}(\mathcal{A})$, as claimed.

Combining (2.16) and (2.17), we obtain (cf. (2.15))

$$\begin{aligned} e(\mathcal{G}) &\leqslant e(\mathcal{G}_*) \cdot \frac{m}{b} \cdot \left(\frac{b}{m} + 1 + 2^{i_0+i_1} - 2 \right) + |W| \cdot \Delta_{(1,0)}^{(i_0,i_1)} \\ &\leqslant e(\mathcal{G}_*) \cdot \frac{m}{b} \cdot 2^{i_0+i_1} + |W| \cdot \Delta_{(1,0)}^{(i_0,i_1)}, \end{aligned} \tag{2.18}$$

since $b \leqslant m$. Now, if the first summand on the right-hand side of (2.15) exceeds $e(\mathcal{G})/2$, then (A1) implies (P1), since $(i_0', i_1') = (i_0 - 1, i_1)$. Otherwise, the second summand is at least $e(\mathcal{G})/2$ and by (A1) and (2.4),

$$|W| \geqslant \frac{e(\mathcal{G})}{2 \cdot \Delta_{(1,0)}^{(i_0,i_1)}} \geqslant \frac{\alpha}{2^{r_0+r_1+1} R} \cdot q,$$

which is (P3). □

2.2.4 Construction of the container

In this section, we present the construction of containers for pairs in $\mathcal{F}_{\leqslant m}(\mathcal{H})$ and analyse their properties, thus proving Theorem 2.2.2. For each $s \in \{0, \dots, r_0 + r_1\}$, define

$$\alpha_s = 2^{-s(r_0+r_1+1)} \qquad \text{and} \qquad \beta_s = \alpha_s \cdot \left(\frac{b}{v_1(\mathcal{H})} \right)^{\min\{r_1, s\}} \left(\frac{b}{m} \right)^{\max\{0, s-r_1\}}.$$

Given an $(I, J) \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, we construct the container (A, B) for (I, J) using the following procedure.

Construction of the container. Let $\mathcal{H}^{(r_0,r_1)} = \mathcal{H}$, let $S_0 = S_1 = \emptyset$, and let $(i_0, i_1) = (r_0, r_1)$. Do the following for $s = 0, \dots, r_0 + r_1 - 1$:

(C1) Let $c \in \{0, 1\}$ be the number that is compatible with (i_0, i_1) and let (i_0', i_1') be the pair defined by $i_c' = i_c - 1$ and $i_{1-c}' = i_{1-c}$.

(C2) Run the algorithm with $\mathcal{G} \leftarrow \mathcal{H}^{(i_0,i_1)}$ to obtain the (i_0', i_1')-uniform hypergraph $\mathcal{G}_*$, the sequence $v_0, \dots, v_{L-1} \in V(\mathcal{H})$, and the partition $\{0, 1, \dots, L-1\} = S \cup W$.

(C3) Let $S_c \leftarrow S_c \cup \{v_j : j \in S\}$.

(C4) If $e(\mathcal{G}_*) < \beta_{s+1} \cdot e(\mathcal{H})$, then define (A, B), the container for (I, J), by

$$(A, B) = (W, \emptyset)$$

if $c = 0$ and

$$(A, B) = (\emptyset, V_1 \setminus W)$$

if $c = 1$, and `STOP`.

(C5) Otherwise, let $\mathcal{H}^{(i_0',i_1')} \leftarrow \mathcal{G}_*$ and $(i_0, i_1) \leftarrow (i_0', i_1')$ and `CONTINUE`.

We will show that the above procedure indeed constructs containers for $\mathcal{F}_{\leqslant m}(\mathcal{H})$ that have the desired properties. To this end, we first claim that for each pair $(i_0, i_1) \in \mathcal{U} \cup \{(0,0)\}$, the hypergraph $\mathcal{H}^{(i_0,i_1)}$, if it was defined, satisfies:

(*i*) $(I, J) \in \mathcal{F}(\mathcal{H}^{(i_0,i_1)})$ and

(*ii*) $\Delta_{(\ell_0,\ell_1)}(\mathcal{H}^{(i_0,i_1)}) \leqslant \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$ for every $(\ell_0, \ell_1) \in \mathcal{V}(i_0, i_1)$.

Indeed, one may easily prove (*i*) and (*ii*) by induction on $(r_0 + r_1) - (i_0 + i_1)$. The basis of the induction is trivial as $\mathcal{H}^{(r_0,r_1)} = \mathcal{H}$, see Definition 2.2.3. The inductive step follows immediately from Observation 2.2.6 and Lemma 2.2.8.

Second, we claim that for each input $(I, J) \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, step (C4) is called for some s and hence the container (A, B) is defined. If this were not true, the condition in step (C5) would be met $r_0 + r_1$ times and, consequently, we would finish with a non-empty $(0, 0)$-uniform hypergraph $\mathcal{H}^{(0,0)}$, i.e., we would have $(\emptyset, \emptyset) \in \mathcal{H}^{(0,0)}$. But this contradicts (*i*), since pair satisfies the empty constraint and thus $(I, J) \notin \mathcal{F}(\mathcal{H}^{(0,0)})$.

Suppose, therefore, that step (C4) is executed when $\mathcal{G} = \mathcal{H}^{(i_0,i_1)}$ for some $(i_0, i_1) \in \mathcal{U}$, and note that $s = (r_0 + r_1) - (i_0 + i_1)$. We claim that $e(\mathcal{H}^{(i_0,i_1)}) \geqslant \beta_s e(\mathcal{H})$. Indeed, this is trivial if $s = 0$, whereas if $s > 0$ and this were not true, then we would have executed step (C4) at the previous step. We therefore have

$$e(\mathcal{G}) = e(\mathcal{H}^{(i_0,i_1)}) \geqslant \beta_s \cdot e(\mathcal{H}) \qquad \text{and} \qquad e(\mathcal{G}_*) < \beta_{s+1} \cdot e(\mathcal{H}),$$

which, by Lemma 2.2.9 and (*ii*), implies that either (P2) or (P3) of Lemma 2.2.9 holds. Note that if $c = 1$, then $r_1 \geqslant i_1 > 0$ and we have

$$|W| \geqslant 2^{-r_1-1} R^{-1} \alpha_s \cdot v_1(\mathcal{H}) \geqslant \alpha_{r_0+r_1} R^{-1} v_1(\mathcal{H}) = \delta v_1(\mathcal{H}),$$

where $\delta = 2^{-(r_0+r_1)(r_0+r_1+1)} R^{-1}$. On the other hand, if $c = 0$, then $r_0 \geqslant i_0 > 0$ and

$$|W| \geqslant 2^{-r_0-r_1-1} R^{-1} \alpha_s \cdot q \geqslant \alpha_{r_0+r_1} R^{-1} q = \delta q.$$

This verifies that (A, B) satisfies property (ii) from the statement of Theorem 2.2.2.

To complete the proof, we need to show that (A, B) can be assigned to each (I, J) by a pair of functions $f \circ g$ for some $g\colon \mathcal{F}_{\leqslant m}(\mathcal{H}) \to \binom{V_0}{\leqslant r_0 b} \times \binom{V_1}{\leqslant r_1 b}$ and to verify that properties (i) and (iii) from the statement of the theorem hold. We claim that one may take $g(I, J) = (S_0, S_1)$, where S_0 and S_1 are the sets constructed by the above procedure, see (C3). To this end, it suffices to show that if for some $(I, J), (I', J') \in \mathcal{F}(\mathcal{H})$ the above procedure produces the same pair (S_0, S_1), then $f \circ g(I, J) = f \circ g(I', J')$. To see this, observe first that the set S defined in step (C2) is precisely the set of all indices $j \in \{0, \dots, L-1\}$ that satisfy $v_j \in S_c$. Indeed, the former set is contained in the latter by construction, see (C3). The reverse inclusion holds because

$$S = \big\{ j \in \{0, \dots, L-1\} : v_j \in V_0 \setminus I \text{ or } v_j \in J \big\}$$

which is exactly the condition on $v \in S_c$. By Observation 2.2.7, it follows that the output of the algorithm depends only on the pair (S_0, S_1) and hence (A, B), as claimed.

Finally, observe that $S_0 \subseteq V_0 \setminus I$ and $S_1 \subseteq J$, by construction, $A \subset I$ and $J \subset B$. This verifies properties (i) and (iii) and hence completes the proof of Theorem 2.2.2. □

2.3 Supersaturation results

We would like to remind the reader that G will always be a fixed abelian group throughout this chapter. To apply Theorem 2.2.2 to our setting we will need, for sets $A, B \subset G$, bounds on the number of pairs $(b_1, b_2) \in B \times B$ such that $b_1 + b_2 \notin A$. In the case $G = \mathbb{Z}$, one such result is Pollard's Theorem [37], which tell us that if $|B| \geqslant (1/2 + \epsilon)|A|$ and $\epsilon < 1/2$ then at least an ϵ^2 proportion of all pairs $(b_1, b_2) \in B \times B$ are such that $b_1 + b_2 \notin A$. To prove similar results for arbitrary abelian groups one has to have some control on the structure of the group. With this in mind, we define the following quantity.

Definition 2.3.1. *Given finite sets $U, V \subset G$, we define*

$$\alpha(U, V) = \max\big\{ |V'| : V' \subset G,\ |V'| \leqslant |V|,\ |\langle V' \rangle| \leqslant |U| + |V| - |V'| \big\}.$$

Given $U, V \subset G$ and $x \in G$ we will use the notation $1_U * 1_V(x)$ to denote the number of pairs $(u, v) \in U \times V$ such that $u + v = x$. The following theorem is the generalization we want of Pollard's theorem for arbitrary abelian groups. It is a simple variant of a result of Hamidoune and Serra [24], and we present a version of their proof for completeness.

Theorem 2.3.2. *Let t be a positive integer and $U, V \subset G$ with $t \leqslant |V| \leqslant |U| < \infty$. Then*

$$\sum_{x \in G} \min(1_U * 1_V(x), t) \geqslant t\big(|U| + |V| - t - \alpha\big), \tag{2.19}$$

where $\alpha := \alpha(U, V)$

Proof. Given an abelian group G and finite subsets $A, B \subset G$, we will proceed by induction on $|B|$ to show that

$$\sum_{x\in G} \min(1_A * 1_B(x), t) \geqslant t(|A| + |B| - t - \alpha),$$

for all integers $t \leqslant |B|$, where $\alpha := \alpha(A, B)$. First, note that if $t = |B| = 1$ then we have

$$\sum_{x\in G} \min(1_A * 1_B(x), t) = |A| \geqslant t(|A| - \alpha).$$

Now take B of size $|B| \geqslant 2$, and define $B' = B - b$ for some $b \in B$ and note that $0 \in B'$. Suppose first that $B' + A \subset A$, and observe that in this case A is an union of cosets of $\langle B'\rangle$, that is, $A = \bigcup_{i=1}^{k} \langle B'\rangle + h_i$ for some $h_1, ..., h_k \in G$. It follows that $1_A * 1_{B'}(x) \geqslant t$ for all $x \in A$, since if $x \in A \cap (\langle B'\rangle + h_i)$ then there are at least $|B'| \geqslant t$ sums $a + b' = x$ with $a \in A \cap (\langle B'\rangle + h_i)$ and $b' \in B'$. Since $G = G - b$ it follows that

$$\sum_{x\in G} \min(1_A * 1_B(x), t) \geqslant t|A| \geqslant t(|A| + |B| - t - \alpha),$$

where the second inequality follows because $|\langle B'\rangle| \leqslant |A|$ and so $\alpha(A, B) \geqslant |B'| = |B|$.

On the other hand, if $B' + A \not\subset A$ then there exists $a^* \in A$ such that $B^* = a^* + B' \not\subset A$ and therefore $1 \leqslant |A \cap B^*| < |B|$. Define $C = A \cup B^*$, $D = A \cap B^*$, $A_1 = A \setminus D$ and $B_1 = B^* \setminus D$. Note that $1_A = 1_{A_1} + 1_D$ and $1_{B^*} = 1_{B_1} + 1_D$ and therefore, by the distributivity property of the convolution operation,

$$\begin{aligned} 1_A * 1_{B^*}(x) &= (1_{A_1} + 1_D) * (1_{B_1} + 1_D)(x) \\ &= 1_{A_1} * 1_{B_1}(x) + (1_{A_1} + 1_{B_1} + 1_D) * 1_D(x) = 1_{A_1} * 1_{B_1}(x) + 1_C * 1_D(x). \quad (2.20) \end{aligned}$$

In particular $1_A * 1_{B^*}(x) \geqslant 1_C * 1_D(x)$. If $|D| \geqslant t$ then by applying our induction hypothesis to C and D, we obtain

$$\sum_{x\in G} \min(1_A * 1_B(x), t) \geqslant \sum_{x\in G} \min(1_C * 1_D(x), t) \geqslant t(|A| + |B| - t - \alpha),$$

where the first step follows since $G = G + a^*$, and the last step follows from the fact that $|C| + |D| = |A| + |B|$ and $\alpha(C, D) \leqslant \alpha(A, B)$, since $|D| \leqslant |B|$.

Finally, if $|D| < t$, observe that

$$\sum_{x\in G} \min(1_A * 1_B(x), t) \geqslant \sum_{x\in G} \min(1_{A_1} * 1_{B_1}(x), t - |D|) + \sum_{x\in G} \min(1_C * 1_D(x), |D|), \quad (2.21)$$

by (2.20). Because $|B_1| < |B|$ we can apply the induction hypothesis to A_1 and B_1, so the right hand side of (2.21) is at least

$$(t - |D|)\big(|A| + |B| - |D| - t - \alpha(A_1, B_1)\big) + |C||D|.$$

Noting that $\alpha(A_1, B_1) \leqslant \alpha(A, B)$, because $|B_1| \leqslant |B|$, and that $|A| + |B| - |D| = |C|$, it follows that the last expression is at least $t(|A| + |B| - t - \alpha)$, as required. □

This implies the following corollary.

Corollary 2.3.3. *Let $A, B \subset G$ be finite and non-empty sets, let $0 < \epsilon < \frac{1}{2}$ and set $\beta := \beta((1+4\epsilon)|A|)$. If $|B| \geqslant (\frac{1}{2} + \epsilon)(|A| + \beta)$ then there are at least $\epsilon^2|B|^2$ pairs $(b_1, b_2) \in B^2$ such that $b_1 + b_2 \notin A$.*

Proof. Note first that if $|B| \geqslant (1+\epsilon)|A|$ then the result is trivial, since for each element $a \in A$ there are at most $|B|$ pairs $(b_1, b_2) \in B^2$ with $b_1 + b_2 = a$, and therefore there are at least $|B|^2 - |A||B| \geqslant \epsilon^2|B|^2$ pairs in B whose sum is not in A. When $|B| \leqslant (1+\epsilon)|A|$ we will apply Theorem 2.3.2 with $U = V = B$ and $t = \epsilon|B|$. We first observe that

$$\alpha(B, B) \leqslant \max\big(\beta, 2|B| - (1+4\epsilon)|A|\big).$$

Indeed, suppose that $B' \subset G$ satisfies $|\langle B'\rangle| \leqslant 2|B| - |B'|$. If $|\langle B'\rangle| > (1+4\epsilon)|A|$ then $|B'| \leqslant 2|B| - |\langle B'\rangle| \leqslant 2|B| - (1+4\epsilon)|A|$. Otherwise, if $|\langle B'\rangle| \leqslant (1+4\epsilon)|A|$, then by the definition (2.1) of β, we have $|B'| \leqslant |\langle B'\rangle| \leqslant \beta$.

Now by Theorem 2.3.2, we have

$$\sum_{x \in G} \min(1_B * 1_B(x), \epsilon|B|) \geqslant \epsilon|B|\Big((2-\epsilon)|B| - \max\big(\beta, 2|B| - (1+4\epsilon)|A|\big)\Big).$$

By subtracting from both sides the sum over $x \in A$, we obtain

$$\sum_{x \in G\setminus A} \min(1_B * 1_B(x), \epsilon|B|) \geqslant \epsilon|B|\Big((2-\epsilon)|B| - \max\big(\beta, 2|B| - (1+4\epsilon)|A|\big) - |A|\Big).$$

Now, if $2|B| - (1+4\epsilon)|A| \geqslant \beta$, then, using that $|B| \leqslant 2|A|$,

$$\sum_{x \in G\setminus A} 1_B * 1_B(x) \geqslant \epsilon|B|\big(4\epsilon|A| - \epsilon|B|\big) \geqslant \epsilon^2|B|^2$$

as required. Otherwise, if $\beta \geqslant 2|B| - (1+4\epsilon)|A|$, then

$$\sum_{x \in G\setminus A} 1_B * 1_B(x) \geqslant \epsilon|B|\big((2-\epsilon)|B| - \beta - |A|\big) \geqslant \epsilon^2|B|^2,$$

since $|B| \geqslant (\frac{1}{2} + \epsilon)(|A| + \beta)$ and $0 < \epsilon < \frac{1}{2}$, so $(2-\epsilon) - \frac{2}{1+2\epsilon} \geqslant \epsilon$. □

To prove a stability theorem for almost all sets with a given size and doubling constant we will also need the following result of Mazur [29].

Theorem 2.3.4. *Let l and t be positive integers, with $t \leqslant l/40$, and let $B \subset \mathbb{Z}$ be a set of size l. Suppose that*

$$\sum_{x \in \mathbb{Z}} \min(1_B * 1_B(x), t) \leqslant (2+\delta)lt,$$

for some $0 < \delta \leqslant 1/8$. *Then there is an arithmetic progression* P *of length at most* $(1+2\delta)l+6t$ *containing all but at most* $3t$ *points of* B.

From Theorem 2.3.4 we can easily deduce the following corollary:

Corollary 2.3.5. *Let* s *be an integer,* $\lambda > 0$, *and* $0 < \epsilon < 2^{-10}$. *If* $A, B \subset \mathbb{Z}$, *with* $(1-\epsilon)\frac{\lambda k}{2} \leqslant |B| \leqslant (1+2\epsilon)\frac{\lambda k}{2}$ *and* $|A| \leqslant \lambda k$ *then one of the following holds:*

(a) There are at least $4\epsilon^2\lambda^2k^2$ *pairs* $(b_1, b_2) \in B^2$ *such that* $b_1 + b_2 \notin A$.

(b) There is an arithmetic progression P *of size at most* $\frac{\lambda k}{2} + 32\epsilon\lambda k$ *containing all but at most* $8\epsilon\lambda k$ *points of* B.

Proof. Suppose first that

$$\sum_{x\in\mathbb{Z}} \min(1_B * 1_B(x), t) \leqslant (2+8\epsilon)2\epsilon|B|\lambda k. \tag{2.22}$$

In this case we apply Theorem 2.3.4 with $l := |B|$, $\delta := 8\epsilon$, and $t = 2\epsilon\lambda k \leqslant l/40$, and deduce that (b) holds. Therefore suppose (2.22) doesn't hold, in this case

$$\sum_{x\in\mathbb{Z}\setminus A} \min(1_B * 1_B(x), t) \geqslant (2+8\epsilon)(1-\epsilon)\epsilon\lambda^2k^2 - t|A|,$$

since $|B| \geqslant (1-\epsilon)\frac{1}{2}\lambda k$. Noting that $t|A| \leqslant 2\epsilon\lambda^2k^2$ it follows that

$$\sum_{x\in\mathbb{Z}\setminus A} 1_B * 1_B(x) \geqslant \Big((2+8\epsilon)(1-\epsilon) - 2\Big)\epsilon\lambda^2k^2 \geqslant 4\epsilon^2\lambda^2k^2,$$

since $\epsilon < 2^{-10}$, so (a) holds as required. □

2.4 Number of sets with a given doubling

In this section we prove the following statement which implies Theorems 2.1.1 and 2.1.3.

Theorem 2.4.1. *Let* k, n *be integers, let* $2 \leqslant \lambda < 2^{-36}\frac{k}{(\log n)^3}$, *and let* $Y \subset G$ *with* $|Y| = n$. *The number of sets* $J \subset Y$ *with* $|J| = k$ *such that* $|J+J| \leqslant \lambda|J|$ *is at most*

$$\exp\left(2^9\lambda\lambda^{1/6}k^{5/6}\sqrt{\log n}\right)\binom{\frac{1}{2}(\lambda k+\beta)}{k},$$

where $\beta := \beta\big(\lambda k + 2^6\lambda^{7/6}k^{5/6}\sqrt{\log n}\big)$ *and* $\lambda := \min\big\{\frac{\lambda}{\lambda-2}, \log k\big\}$.

Theorem 2.4.1 will follow easily from the following container theorem combined with Corollary 2.3.3. We will also use it together with Corollary 2.3.5 to prove Theorem 2.5.1.

Theorem 2.4.2. *Let m, n be integers with $m \geqslant (\log n)^2$, let $Y \subset G$ with $|Y| = n$, and let $0 < \epsilon < \frac{1}{4}$. There is a family $\mathcal{A} \subset 2^{Y+Y} \times 2^Y$ of pairs of sets (A, B), of size*

$$|\mathcal{A}| \leqslant \exp\left(2^{16}\frac{1}{\epsilon^2}\sqrt{m}(\log n)^{3/2}\right) \tag{2.23}$$

such that:

(i) For every pair of sets $J \subset Y$, $I \subset Y + Y$, with $J + J \subset I$ and $|I| \leqslant m$ there is $(A, B) \in \mathcal{A}$ such that $A \subset I$ and $J \subset B$.

(ii) For every $(A, B) \in \mathcal{A}$, $|A| \leqslant m$ and either $|B| \leqslant \frac{m}{\log n}$ or there are at most $\epsilon^2|B|^2$ pairs $(b_1, b_2) \in B \times B$ such that $b_1 + b_2 \notin A$.

Proof that Theorem 2.4.2 implies Theorem 2.4.1. Let $\mathcal{A}$ be a family given by Theorem 2.4.2 applied with $m := \lambda k$ and $\epsilon > 0$ to be chosen later. Then by condition (i), for every k-set J with doubling constant λ there is a pair $(A, B) \in \mathcal{A}$ such that $J \subset B$ and $A \subset J + J$. Define $\mathcal{B}$ to be the family of all sets B that are in some container pair, that is

$$\mathcal{B} = \{B \subset Y : \exists A \text{ such that } (A, B) \in \mathcal{A}\}.$$

Observe that, by Corollary 2.3.3 and condition (ii) on $\mathcal{A}$, for every $B \in \mathcal{B}$ we have $|B| \leqslant (\frac{1}{2} + \epsilon)(m + \beta)$, where $\beta := \beta((1 + 4\epsilon)m)$, since the number of pairs $(b_1, b_2) \in B^2$ such that $b_1 + b_2 \notin A$ is at most $\epsilon^2|B|^2$ and $\frac{m}{\log n} \leqslant (\frac{1}{2} + \epsilon)(m + \beta)$. Therefore the number of sets of size k with doubling constant λ is at most

$$|\mathcal{B}| \max_{B \in \mathcal{B}} \binom{|B|}{k} \leqslant \exp\left(2^{16}\frac{1}{\epsilon^2}\sqrt{\lambda k}(\log n)^{3/2}\right)\binom{(\frac{1+2\epsilon}{2})(\lambda k + \beta)}{k}. \tag{2.24}$$

Let $\lambda := \min\{\frac{\lambda}{\lambda - 2}, \log k\}$, suppose first that $\frac{\lambda}{\lambda-2} \leqslant \log k$. By applying the inequality $\binom{cn}{k} \leqslant (\frac{cn-k}{n-k})^k \binom{n}{k}$ with $k = k$, $c = 1 + 2\epsilon$ and $n = \frac{\lambda k + \beta}{2}$, it follows that in this case (2.24) is at most

$$\exp\left(2^{16}\frac{1}{\epsilon^2}\sqrt{\lambda k}(\log n)^{3/2} + 2\epsilon\lambda k\right)\binom{\frac{\lambda k+\beta}{2}}{k}.$$

Now choosing $\epsilon := 2^4\left(\frac{\lambda}{k}\right)^{1/6}\sqrt{\log n}$, by our restrictions on λ we see that

$$\epsilon < 2^4\left(\frac{1}{2^{36}(\log n)^3}\right)^{1/6}\sqrt{\log n} = \frac{1}{4}.$$

It follows that there are at most $\exp\left(2^9\lambda\lambda^{1/6}k^{5/6}\sqrt{\log n}\right)\binom{\frac{1}{2}(\lambda k+\beta)}{k}$ sets of size k with doubling constant λ, when $\frac{\lambda}{\lambda-2} \leqslant \log k$. If $\log k \leqslant \frac{\lambda}{\lambda-2}$ we use the binomial estimate

$$\binom{(\frac{1+2\epsilon}{2})(\lambda k + \beta)}{k} \leqslant \exp\left(4\epsilon k \log\frac{1}{\epsilon}\right)\binom{\frac{\lambda k+\beta}{2}}{k}$$

and the result follows by a similar calculation. Since $\beta(m + 4\epsilon m) = \beta(\lambda k + 2^6\lambda^{7/6}k^{5/6}\sqrt{\log n})$, this proves the theorem. □

Before we proceed with the proof of Theorem 2.4.2, let us give a brief overview of how we will deduce it from Theorem 2.2.2. We fix from now on a finite subset $Y \subset G$ with $|Y| = n$, and recall that the $(1,2)$-bounded hypergraph $\mathcal{H}(A,B)$ in Definition 2.2.1 was defined to have as edges pairs $(\{c\},\{a,b\})$ where $a+b=c$, with $a,b \in B$ and $c \notin A$. Note that condition (ii) in Theorem 2.4.2 implies that $\mathcal{H}(A,B)$ has at most $\frac{\epsilon^2}{2}|B|^2$ edges, as long as $|B| > \frac{m}{\log n}$. We remind the reader that a pair of sets $I \subset Y+Y$ and $J \subset Y$ with $J+J \subset I$ correspond to an independent set in $\mathcal{H}(A,B)$ for any $A \subset Y+Y$ and $B \subset Y$, since there are no $c \notin I$ and $a,b \in J$ such that $a+b=c$. If we additionally assume that $(I,J) \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, then we know that every J that is in such an independent pair satisfies $|J+J| \leqslant m$.

Our strategy will be to iteratively apply the container lemma until either there are few edges in the hypergraph $\mathcal{H}(A,B)$, or $|A| > m$, in which case the container doesn't contain any elements of $\mathcal{F}_{\leqslant m}(\mathcal{H})$. More precisely we will build a rooted tree $\mathcal{T}$ with root $\mathcal{H}(\emptyset, Y)$ whose vertices correspond to hypergraphs $\mathcal{H}(A,B)$ and whose leaves correspond to a family $\mathcal{A}$ satisfying the conclusion of Theorem 2.4.2. Given a vertex $\mathcal{H}(A,B)$ of the tree, such that $|A| \leqslant m$, $|B| > \frac{m}{\log n}$ and

$$e(\mathcal{H}(A,B)) > \frac{\epsilon^2}{2}|B|^2, \tag{2.25}$$

we will generate its children by applying the following procedure:

(a) Apply the asymmetric container lemma (Theorem 2.2.2) to $\mathcal{H} := \mathcal{H}(A,B)$ setting

$$R := \frac{2}{\epsilon^2}, \quad q := \frac{m}{\log n}, \quad b := \sqrt{\frac{m}{\log n}}.$$

Notice that the co-degrees of $\mathcal{H}$ satisfy

$$\max\left\{\Delta_{(1,0)}(\mathcal{H}), \Delta_{(0,1)}(\mathcal{H})\right\} \leqslant |B| = \frac{2}{\epsilon^2}\frac{\epsilon^2|B|^2}{2|B|} \leqslant R\frac{e(\mathcal{H})}{|B|}$$

and

$$\Delta_{(0,2)}(\mathcal{H}) = \Delta_{(1,1)}(\mathcal{H}) = \Delta_{(1,2)}(\mathcal{H}) = 1 = \frac{2}{\epsilon^2}\frac{b^2}{q|B|^2}\frac{\epsilon^2}{2}|B|^2 \leqslant R\frac{b^2}{q|B|^2}e(\mathcal{H}),$$

since (2.25) holds. Since $b < q < |B|$, it follows that

$$\Delta_{(0,2)}(\mathcal{H}) \leqslant R\frac{b^2}{q|B|^2}e(\mathcal{H}) \leqslant R\frac{b}{|B|^2}e(\mathcal{H}),$$

$$\Delta_{(1,1)}(\mathcal{H}) \leqslant R\frac{b^2}{q|B|^2}e(\mathcal{H}) \leqslant R\frac{b}{q|B|}e(\mathcal{H})$$

and

$$\Delta_{(1,0)}(\mathcal{H}) \leqslant R\frac{e(\mathcal{H})}{|B|} \leqslant R\frac{e(\mathcal{H})}{q},$$

as required.

(b) By Theorem 2.2.2, there exists a family $\mathcal{C} \subset 2^{(Y+Y)\setminus A} \times 2^B$ of at most

$$\binom{n^2}{b}\binom{|B|}{2b} \leqslant n^{4b} \leqslant e^{4\sqrt{m \log n}}, \tag{2.26}$$

pairs of sets (C, D) that satisfies the conditions of the container lemma. That is for each independent pair $(I, J) \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, with $I \subset Y + Y$ and $J \subset Y$, there is $(C, D) \in \mathcal{C}$ such that $C \subset I$ and $J \subset D$, and either $|C| \geqslant \delta \frac{m}{\log n}$, or $D \leqslant (1-\delta)|B|$.

(c) For each $(C, D) \in \mathcal{C}$, let $\mathcal{H}(A \cup C, D)$ be a child of $\mathcal{H}(A, B)$ in the tree $\mathcal{T}$.

Now to count the number of leaves of $\mathcal{T}$ we will first bound its depth.

Lemma 2.4.3. *The tree $\mathcal{T}$ has depth at most $d = 2^{14}\epsilon^{-2}\log n$.*

Proof. We will prove that after d iterations either $|A| > m$, $|B| \leqslant \frac{m}{\log n}$ $e(\mathcal{H}(A, B)) \leqslant \frac{\epsilon^2}{2}|B|^2$. Notice that the δ provided by Theorem 2.2.2 in this application is $2^{-13}\epsilon^2$ and in each iteration either we increase the size of A by δq or we decrease the size of B by $\delta|B|$. After d iterations, either we would have increased the size of A more than $\frac{d}{2}$ times, in which case

$$|A| > \frac{d}{2}\delta q = \frac{2^{13}\log n}{\epsilon^2} 2^{-13}\epsilon^2 \frac{m}{\log n} = m,$$

or we would have reduced the size of B at least $\frac{d}{2}$ times, in which case

$$|B| \leqslant (1-\delta)^{\frac{d}{2}} n < e^{-\frac{\delta d}{2}} n \leqslant e^{-\log n} n = 1.$$

In either case, we would have stopped already by this point because we only generate children of $\mathcal{H}(A, B)$ if $|A| \leqslant m$, $|B| > \frac{m}{\log n}$ and (2.25) holds. □

Proof of Theorem 2.4.2. Let $\mathcal{L}$ be the set of leaves of the tree $\mathcal{T}$ constructed above, and define

$$\mathcal{A} := \{(A, B) : A \subset Y + Y,\ B \subset Y,\ \mathcal{H}(A, B) \in \mathcal{L},\ |A| \leqslant m\}.$$

Notice that for every $(A, B) \in \mathcal{A}$, we have either the bound $e(\mathcal{H}(A, B)) \leqslant \frac{\epsilon^2}{2}|B|^2$ or $|B| \leqslant \frac{m}{\log n}$, since they come from the leaves of $\mathcal{T}$ and $|A| \leqslant m$. Since the edges of $\mathcal{H}(A, B)$ correspond exactly to pairs $a, b \in B$ such that $a + b \notin A$, it follows that $\mathcal{A}$ has property (ii).

To bound the size of $\mathcal{A}$, notice that the number of leaves of the tree $\mathcal{T}$ is at most Z^d where Z denotes the maximum number of children of a vertex of the tree and d denotes its depth. Thus, by (2.26) and Lemma 2.4.3,

$$|\mathcal{A}| \leqslant |\mathcal{L}| \leqslant Z^d \leqslant \exp\left(2^{16}\frac{1}{\epsilon^2}\sqrt{m}(\log n)^{3/2}\right),$$

so $\mathcal{A}$ satisfies (2.23), as required.

Finally, observe that for every pair of sets $J \subset Y$, $I \subset Y + Y$ with $J + J \subset I$ and $|I| \leqslant m$, there is $(A, B) \in \mathcal{A}$ such that $A \subset I$ and $J \subset B$. Indeed $(I, J) \in \mathcal{F}_{\leqslant m}\big(\mathcal{H}(\emptyset, Y)\big)$ and therefore, by

property (b) of our containers, there exists a path from the root to a leaf of $\mathcal{T}$ such that $A \subset I$ and $J \subset B$ for every vertex $\mathcal{H}(A,B)$ of the path, so (i) holds. □

2.5 Typical structure result

In this section we use Theorem 2.4.2 to determine the typical structure of a set $J \subset [n]$ of a given size with doubling constant λ.

Theorem 2.5.1. *Let $n, k \in \mathbb{N}$ and $2 \leqslant \lambda \leqslant 2^{-120}\frac{k}{(\log n)^3}$, and let $2^8\lambda^{1/6}k^{-1/6}\sqrt{\log n} \leqslant \gamma < 2^{-8}$. For all but at most*

$$e^{-\gamma k}\binom{\lambda k/2}{k}$$

sets $J \subset [n]$ with $|J| = k$ and $|J+J| \leqslant \lambda k$, the following holds: there exists $T \subset J$, with $|T| \leqslant 2^9\gamma k$, such that $J \setminus T$ is contained in an arithmetic progression of size $\lambda k/2 + 2^7\gamma\lambda k$.

Let us say that a set $B \subset [n]$ is (ε, m)-*close to an arithmetic progression* if there is an arithmetic progression P with $|P| \leqslant m/2 + 2^5\varepsilon m$, and a set $T \subset B$ with $|T| \leqslant 8\varepsilon m$ such that $B \setminus T \subset P$. Recall also from (3.19) the definition of Λ.

Proof of Theorem 2.5.1. Set $Y := [n]$, $\varepsilon := 4\gamma$ and $m := \lambda k \geqslant 2^{120}\lambda^2(\log n)^3$, and let $\mathcal{A}$ be the family of sets given by Theorem 2.4.2. We prove the theorem via three simple claims.

Claim 1: For every pair $(A,B) \in \mathcal{A}$, either

(*a*) $|B| \leqslant (1-\varepsilon)\lambda k/2$ or

(*b*) $|B| \leqslant (1+2\varepsilon)\lambda k/2$, and B is $(\varepsilon, \lambda k)$-close to an arithmetic progression.

Proof of Claim 1. To see this, let $(A,B) \in \mathcal{A}$ and suppose that $|B| \geqslant (1-\varepsilon)\lambda k/2$. By Theorem 3.3.1(*ii*), there are at most $\varepsilon^2|B|^2$ pairs $b_1, b_2 \in B$ with $b_1 + b_2 \notin A$. By Lemma 2.3.3, it follows that $|B| \leqslant (1+2\varepsilon)\lambda k/2$. Now, by Lemma 2.3.5, and noting that $\varepsilon^2|B|^2 < 4\varepsilon^2\lambda^2k^2$, it follows that there is an arithmetic progression P with $|P| \leqslant \lambda k/2 + 2^5\varepsilon\lambda k$, and a set $T \subset B$ with $|T| \leqslant 8\varepsilon\lambda k$ such that $B \setminus T \subset P$, as required. □

Now, recall from Theorem 2.4.2(*i*) that for each set J, with $|J| = k$ and $|J+J| \leqslant \lambda k$, there exists $(A,B) \in \mathcal{A}$ such that $A \subset J+J$ and $J \subset B$. We first consider the pairs $(A,B) \in \mathcal{A}$ with $|B| \leqslant (1-\varepsilon)\lambda k/2$.

Claim 2: There are at most

$$e^{-\varepsilon k/2}\binom{\lambda k/2}{k}$$

sets J, with $|J| = k$ and $|J+J| \leqslant \lambda k$, such that $J \subset B$ for some $(A,B) \in \mathcal{A}$ with $|B| \leqslant (1-\varepsilon)\lambda k/2$.

Proof of Claim 2. Recalling the bound (3.4) on the size of $\mathcal{A}$, it follows (using (3.17)) that the number of sets J is at most

$$|\mathcal{A}| \cdot \binom{(1-\varepsilon)\lambda k/2}{k} \leqslant \exp\left(2^{16}\varepsilon^{-2}\sqrt{\lambda k}(\log n)^{3/2} - \varepsilon k\right)\binom{\lambda k/2}{k}. \tag{2.27}$$

Now, recalling that $\varepsilon \geqslant 2^6\lambda^{1/6}k^{-1/6}(\log n)^{1/2}$, it follows that the right-hand side is at most $e^{-\varepsilon k/2}\binom{\lambda k/2}{k}$, as claimed. □

Finally, we consider the pairs $(A, B) \in \mathcal{A}$ that satisfy property (b) of Claim 1. Let us write Λ' for the family of sets J, with $|J| = k$ and $|J + J| \leqslant \lambda k$, such that $J \setminus T$ is contained in an arithmetic progression of size $\lambda k/2 + 2^5\varepsilon\lambda k$ for some $T \subset J$ with $|T| \leqslant 2^7\varepsilon k$.

Claim 3: There are at most

$$e^{-\varepsilon k}\binom{\lambda k/2}{k}$$

k-sets $J \notin \Lambda'$ with $|J + J| \leqslant \lambda k$ and $J \subset B$ for some $(A, B) \in \mathcal{A}$ such that $|B| \leqslant (1 + 2\varepsilon)\lambda k/2$, and B is $(\varepsilon, \lambda k)$-close to an arithmetic progression.

Proof of Claim 3. Let $(A, B) \in \mathcal{A}$, and suppose that $B \setminus T \subset P$ for some arithmetic progression P and set $T \subset B$ with

$$|P| \leqslant \lambda k/2 + 2^5\varepsilon\lambda k \qquad \text{and} \qquad |T| \leqslant 8\varepsilon\lambda k.$$

Observe that there at most

$$\sum_{s \geqslant 2^7\varepsilon k} \binom{(1+2\varepsilon)\lambda k/2}{k-s}\binom{8\varepsilon\lambda k}{s} \tag{2.28}$$

k-sets $J \notin \Lambda'$ with $|J + J| \leqslant \lambda k$ and $J \subset B$. Indeed, $J \setminus T \subset P$, so if $|J \cap T| \leqslant 2^7\varepsilon k$ then $J \in \Lambda'$.

Note that the right-hand side of (2.28) is zero if $\lambda < 2^4$, so we may assume that $\lambda \geqslant 2^4$. Now, observe that

$$\binom{(1+2\varepsilon)\lambda k/2}{k-s}\binom{8\varepsilon\lambda k}{s} \leqslant (1+2\varepsilon)^k\left(\frac{2}{\lambda - 2} \cdot \frac{8e\varepsilon\lambda k}{s}\right)^s\binom{\lambda k/2}{k}.$$

Hence, summing (2.28) over $(A, B) \in \mathcal{A}$, and noting that $|\mathcal{A}| \leqslant e^{\varepsilon k}$ (cf. (2.27)), it follows that there are at most

$$e^{3\varepsilon k}\binom{\lambda k/2}{k}\sum_{s \geqslant 2^7\varepsilon k}\left(\frac{2^6\varepsilon k}{s}\right)^s \leqslant e^{-\varepsilon k}\binom{\lambda k/2}{k},$$

as claimed. □

Now, recall that, by Theorem 2.4.2 (i), for every J, with $|J| = k$ and $|J + J| \leqslant \lambda k$, there exists $(A, B) \in \mathcal{A}$ such that $A \subset J + J$ and $J \subset B$. Combining Claims 1, 2 and 3, it follows that there are at most

$$\left(e^{-\varepsilon k/2} + e^{-\varepsilon k}\right)\binom{\lambda k/2}{k} \leqslant e^{-\gamma k}\binom{\lambda k/2}{k},$$

k-sets $J \notin \Lambda'$ with $|J + J| \leqslant \lambda k$, as required. □

2.6 Lower bound on number of sets with large doubling

In this section we present the details for the construction showing that Conjecture 1.1.1 isn't true if $\lambda \gg k(\log n)^{-1}$

Proposition 2.6.1. *Let n and k be positive integers, and let $\lambda, \epsilon > 0$ and $C \geqslant 2$ satisfy $\min\{k, n^{1/2-\epsilon}\} \geqslant \lambda \geqslant 4\frac{\log(24C)k}{\epsilon \log n}$. There are at least*

$$\binom{C\lambda k}{k}$$

sets $J \subset [n]$ with $|J| = k$ and $|J + J| \leqslant \lambda k$.

Proof. Choose P to be an arithmetic progression of length $\frac{\lambda k}{8}$ and let $J = J_0 \cup J_1$, with $J_0 \subset P$ of size $k - \frac{\lambda}{4}$ and $J_1 \subset [n] \setminus P$ of size $\frac{\lambda}{4}$. Then J has doubling constant λ since

$$\begin{aligned} |J + J| &\leqslant |J_0 + J_0| + |J_0 + J_1| + |J_1 + J_1| \\ &\leqslant 2|P| + |J_0||J_1| + |J_1|^2 \leqslant \frac{\lambda k}{4} + \frac{\lambda k}{4} + \frac{\lambda^2}{16} \leqslant \lambda k. \end{aligned}$$

Finally, by using that $\log(\frac{n}{\lambda^2}) \geqslant \epsilon \log n$ and the bounds

$$\binom{b}{d}\binom{a}{c-d} \geqslant \left(\frac{bc}{4ad}\right)^d \binom{a}{c} \qquad \text{and} \qquad a\binom{b}{c} \geqslant \binom{\frac{a^{1/c}b}{e}}{c}$$

valid for any positive integers a, b, c, d, such that $4d \leqslant c$, we have at least

$$\binom{\frac{n}{2}}{\frac{\lambda}{4}}\binom{\frac{\lambda k}{8}}{k-\frac{\lambda}{4}} \geqslant \left(\frac{n}{\lambda^2}\right)^{\lambda/4}\binom{\frac{\lambda k}{8}}{k} \geqslant \binom{\exp(\frac{\epsilon\lambda \log n}{4k})\frac{\lambda k}{8e}}{k}$$

choices for J. In particular if $\lambda \geqslant \frac{4\log(24C)k}{\epsilon \log n}$ this is at least $\binom{C\lambda k}{k}$. □

Chapter 3

The Typical Structure of Sets with Small Sumset

3.1 Introduction

The work in this chapter is was done jointly with Maurício Collares, Robert Morris, Natasha Morrison, and Victor Souza. In this chapter we will build on Chapter 2 and obtain a significantly more precise description of the typical structure of sets with bounded doubling.[1] For each $\lambda \geqslant 3$ and $\varepsilon > 0$, define

$$c(\lambda, \varepsilon) := 2^{18}\lambda^2 \log \lambda \cdot \log(1/\varepsilon) + 2^{480}\lambda^{30}. \tag{3.1}$$

The main theorem of this chapter, which determines (up to an additive constant) the typical length of the smallest arithmetic progression containing a set with bounded doubling, is as follows.

Theorem 3.1.1. *Fix $\lambda \geqslant 3$ and $\varepsilon > 0$, let $n \in \mathbb{N}$ be sufficiently large, and let $k \geqslant (\log n)^4$. Let $A \subset [n]$ be chosen uniformly at random from the sets with $|A| = k$ and $|A + A| \leqslant \lambda k$. Then there exists an arithmetic progression P with*

$$A \subset P \qquad \text{and} \qquad |P| \leqslant \frac{\lambda k}{2} + c(\lambda, \varepsilon)$$

with probability at least $1 - \varepsilon$.

When λ is large and ε is very small the constant $c(\lambda, \varepsilon)$ is not far from best possible. Indeed, a simple construction (see Section 3.10) shows that with probability at least ε the smallest arithmetic progression containing A has size $\lambda k/2 + \Omega\big(\lambda^2 \log(1/\varepsilon)\big)$.

[1] When $|A + A| = O(|A|)$, then we (informally) say that A has *bounded doubling*.

We will use Theorem 3.1.1 to deduce the following counting result.

Corollary 3.1.2. *For every $\lambda \geqslant 3$, and every $n, k \in \mathbb{N}$ with $(\log n)^4 \leqslant k = o(n)$, we have*

$$\big|\big\{A \subset [n] : |A| = k,\, |A+A| \leqslant \lambda k\big\}\big| = \Theta_\lambda(1) \cdot \frac{n^2}{k}\binom{\lambda k/2}{k},$$

The upper bound in Corollary 3.1.2 is an almost immediate consequence of Theorem 3.1.1, and our lower bound follows from a straightforward calculation (see Sections 3.9 and 3.10). For both bounds we obtain a constant of the form $\exp\big(\lambda^{\Theta(1)}\big)$ for λ large, and it would be interesting to determine the correct exponent of λ.

We remark that similar results can be deduced from our proof for all $2 < \lambda < k^{o(1)}$ (see Section 3.9), but the constant given by our method tends to infinity as $\lambda \to 2$. In order to keep the calculations as simple as possible, we have chosen to focus on the case $\lambda \geqslant 3$.

To see why the much more precise structure given by Theorem 3.1.1 should typically occur, it is perhaps instructive to consider a random k-subset $A \subset [\lambda k/2 + r]$ for some $r \geqslant 0$. The number of such sets is $\binom{\lambda k/2+r}{k} \approx \exp(2r/(\lambda-2))\binom{\lambda k/2}{k}$, and we will be able to show (see Lemma 3.4.1 and [22, Theorem 1.3]) that (very roughly)

$$\mathbb{P}\big(|A+A| \leqslant \lambda k\big) \approx \mathbb{P}\big(A \cap \{1, \dots, r\} = \emptyset\big) \approx (1 - 2/\lambda)^r.$$

Multiplying these bounds, we already see that the number of sets $A \subset [\lambda k/2 + r]$ with $|A| = k$ and $|A + A| \leqslant \lambda k$ does not grow too quickly with r. Unfortunately, the bound given by Lemma 3.4.1 is not strong enough to deduce the result via such a simple argument, and our proof will be significantly more complicated. However, we would like to emphasize that our approach (while somewhat technical in places) is entirely combinatorial.

The main tool in the proof of Theorem 3.1.1 is the 'container theorem' for sets with small doubling (see Theorem 2.4.2), which was proved in the previous chapter 2. We will use this container theorem in three different ways: first, to control the rough structure of a set with bounded doubling (see Theorem 3.3.3 and Lemma 3.5.2); then to prove a variant of a probabilistic lemma of Green and Morris [22] (see Lemma 3.4.1); and finally to control the fine structure of the set near the ends of the progression containing it (see Section 3.8). We consider this last step to be the most interesting aspect of the proof, since we are not aware of any previous application of containers to the task of 'cleaning up' a set, that is, replacing a rough structural result with a precise one. We hope that our proof will inspire further applications of this type in other combinatorial settings.

3.2 An overview of the proof

In this section we will prepare the reader for the details of the proof by giving a rough outline of the main ideas. Let us fix $\lambda \geqslant 3$, and let $k \in \mathbb{N}$ be sufficiently large. We will mostly work with sets of integers that are 'close' to being a subset of the interval $[\lambda k/2]$, since the stability theorem proved in the previous chapter 2 (see Theorem 3.3.3, below) implies that almost all of the sets that we need to count are close to an arithmetic progression of length $\lambda k/2$, and any such progression can be mapped into $[\lambda k/2]$ (see Section 3.5 for the details).

Given a set $A \subset \mathbb{Z}$, let us write

$$b(A) := |A \setminus [\lambda k/2]| \qquad \text{and} \qquad r(A) := \max(A) - \min(A) - \lambda k/2. \tag{3.2}$$

Let us also fix $\varepsilon > 0$ and set $\delta := 2^{-18}\lambda^{-2}$. By Lemma 3.5.1, below, the problem will reduce to bounding the size of the following family of sets.

Definition 3.2.1. *Let $\mathcal{I}$ denote the family of sets $A \subset \{-\lambda k/2, \dots, \lambda k\}$ with $|A| = k$ and $|A + A| \leqslant \lambda k$, such that*

$$b(A) \leqslant \delta k \qquad \text{and} \qquad r(A) \geqslant c(\lambda, \varepsilon),$$

and the sets $\big\{x \in A : x \leqslant 0\big\}$ and $\big\{x \in A : x > \lambda k/2\big\}$ are non-empty.

We will partition the family $\mathcal{I}$ according to the 'density' of the set $B := A \setminus [\lambda k/2]$. To be precise, set

$$f(\lambda) := 2^{10}\lambda^3, \tag{3.3}$$

and say that B is *sparse* if $r(A) > f(\lambda)b(A)$. The following lemma, which is proved in Section 3.6, bounds the number of sets $A \in \mathcal{I}$ such that B is sparse.

Lemma 3.2.2. *For every $\lambda \geqslant 3$ and $\varepsilon \in (0,1)$, and every $k \in \mathbb{N}$, we have*

$$\Big|\Big\{A \in \mathcal{I} \,:\, r(A) > f(\lambda)b(A)\Big\}\Big| \leqslant \frac{\varepsilon}{\lambda^3}\binom{\lambda k/2}{k}.$$

In order to motivate the proof of Lemma 3.2.2, it is instructive to consider the following (very simple) construction, which shows that the bound in Theorem 3.1.1 is close to best possible. Set $r := 2^{-6}\lambda^2 \log(1/\varepsilon)$, and consider the family of sets $A = A' \cup \{v\}$, where $1 \in A' \subset [\lambda k/2 - 8r/\lambda]$ with $|A'| = k-1$, and $v = \lambda k/2 + r$. The number of such sets is

$$\binom{\lambda k/2 - 8r/\lambda - 1}{k-2} \geqslant \frac{4}{\lambda^2}\exp\left(-\frac{2^5 r}{\lambda^2}\right)\binom{\lambda k/2}{k} \geqslant \frac{\varepsilon}{\lambda^2}\binom{\lambda k/2}{k},$$

and most such sets satisfy $|A + A| \leqslant \lambda k$ (for the details, see Section 3.10).

The reason that we cannot take r significantly larger than $\lambda^2 \log(1/\varepsilon)$ in the construction above is that the set $(A' + \max(B)) \setminus [\lambda k]$ typically contains about $2r/\lambda$ elements, and this restricts the size of the set $A' + A'$, and hence the number of choices for $A' := A \cap [\lambda k/2]$. In the proof of Lemma 3.2.2 we will use this simple idea to bound the number of choices for A' (using a straightforward counting argument when $(A' + \max(B)) \setminus [\lambda k]$ is much smaller than r/λ, and an application of the container theorem (via Lemma 3.4.1) when it is larger). We will then use the inequality $r(A) > f(\lambda)b(A)$ to (trivially) bound the number of choices for the set B (that is, the remaining elements of A).

Let us note here that the key tool in the proof of Lemma 3.2.2 outlined above is a probabilistic lemma (Lemma 3.4.1), which is a variant of a result of Green and Morris [22]. This lemma gives a (close to tight) upper bound on the number of k-subsets of $[n]$ whose sumset missed many elements of $\{2, \dots, 2n\}$, and is proved in Section 3.4, using the container theorem 2.4.2 from Chapter 2.

When $r(A) \leqslant f(\lambda)b(A)$, we will say that the set is *dense*. In Sections 3.7 and 3.8 we will prove the following lemma, which bounds the number of dense sets in $\mathcal{I}$.

Lemma 3.2.3. *For every $\lambda \geqslant 3$ and $\varepsilon \in (0,1)$, and every $k \in \mathbb{N}$, we have*

$$\left|\left\{A \in \mathcal{I} : r(A) \leqslant f(\lambda)b(A)\right\}\right| \leqslant \frac{\varepsilon}{\lambda^3}\binom{\lambda k/2}{k}.$$

The proof of Lemma 3.2.3 is significantly more difficult than that of Lemma 3.2.2, and is the most interesting and novel part of the argument, involving a surprising and unusual application of the container method. Set $A' := A \cap [\lambda k/2]$ and $B := A \setminus [\lambda k/2]$, as above, and suppose that $|B| = b$ and $|(B+B)\setminus[\lambda k]| = \mu b$. The main difficulties arise when $r = O(\mu b)$ and $\mu = \Theta(\lambda)$, and we first take care of the remaining cases in Section 3.7. For these 'easy' cases (see Lemmas 3.7.2 and 3.7.5) we use similar ideas to those used to prove Lemma 3.2.2, except that we will apply Theorem 3.3.2 to bound the number of choices for the set B (see Lemma 3.7.3), and the calculations are significantly more delicate. In particular, we will need to use our bounds on the size of both $(A' + \max(B)) \setminus [\lambda k]$ (as in Section 3.6) and $(B + B) \setminus [\lambda k]$ to bound the size of $A' + A'$, and thus the number of choices for A'.

When $r = O(\mu b)$ and $\mu = \Theta(\lambda)$, the first step is to apply the container theorem 2.4.2, to show that for each $b \in \mathbb{N}$, there exists a family $\mathcal{B}(b)$ of size $2^{o(b)}$, such that for each set A that we would like to count (with $|B| = b$), there exists an element $(C, D) \in \mathcal{B}(b)$ that 'contains' A in a suitable sense (see Corollary 3.8.1). The properties of these 'containers' are sufficiently restrictive that we can bound (see Lemmas 3.8.3 and 3.8.4) the number of sets A that are 'contained' in a given element of $\mathcal{B}(b)$ by (roughly) $\exp\big(-b/\lambda \log \lambda\big)\binom{\lambda k/2}{k}$. Hence, summing

over $b, r \in \mathbb{N}$ with $r \geqslant c(\lambda, \varepsilon)$ and $b \leqslant r = O(\lambda b)$, and also over containers $(C, D) \in \mathcal{B}(b)$, we obtain the bound in Lemma 3.2.3.

The rest of the chapter is organised as follows. First, in Section 3.3, we recall the main results from the previous chapter, and deduce the container theorem we will use in the proof (Corollary 3.3.4). In Section 3.4 we use this container theorem to prove the probabilistic lemma mentioned above (Lemma 3.4.1), and in Section 3.5 we will use the results of the previous chapter to reduce the problem to that of bounding the size of the set $\mathcal{I}$. In Section 3.6 we prove Lemma 3.2.2, in Sections 3.7 and 3.8 we prove Lemma 3.2.3, and in Section 3.9 we put the pieces together and prove Theorem 3.1.1. Finally, in Section 3.10, we provide two simple constructions that show that the upper bounds in Theorem 3.1.1 and Corollary 3.1.2 are not far from best possible.

3.3 The container theorem

In this section we will recall for convience the main results from the previous chapter, which will play an important role in the proofs of the main theorems of this chapter. We begin by stating the main container theorem.

Theorem 3.3.1 (Theorem 2.4.2)**.** *Let $m \geqslant (\log n)^2$, let $Y \subset \mathbb{Z}$ with $|Y| = n$, and let $0 < \gamma < 1/4$. There is a family $\mathcal{A} \subset 2^{Y+Y} \times 2^Y$ of pairs of sets (A, B), of size*

$$|\mathcal{A}| \leqslant \exp\left(2^{16}\gamma^{-2}\sqrt{m}\,(\log n)^{3/2}\right), \tag{3.4}$$

such that:

(i) For every pair of sets $J \subset Y$, $I \subset Y + Y$, with $J + J \subset I$ and $|I| \leqslant m$, there is $(A, B) \in \mathcal{A}$ such that $A \subset I$ and $J \subset B$.

(ii) For every $(A, B) \in \mathcal{A}$, $|A| \leqslant m$ and either $|B| \leqslant \frac{m}{\log n}$ or there are at most $\gamma^2|B|^2$ pairs $(b_1, b_2) \in B \times B$ such that $b_1 + b_2 \notin A$.

In order to understand the statement of Theorem 3.3.1, it is useful to consider the case $I = J + J$ and $|B| > m/\log n$. In this case the conditions imply that there exists a 'container' $(A, B) \in \mathcal{A}$ for the pair (I, J) such that $J \subset B$, $B + B \approx A$, and $A \subset J + J$.

We will also use the other main results from the previous chapter.

Theorem 3.3.2 (Theorem 2.4.1)**.** *Let $n, k \in \mathbb{N}$, and let $2 < \lambda < 2^{-36}\frac{k}{(\log n)^3}$. The number of sets $A \subset [n]$ with $|A| = k$ such that $|A + A| \leqslant \lambda k$ is at most*

$$\exp\left(2^9\mu\lambda^{1/6}k^{5/6}\sqrt{\log n}\right)\binom{\lambda k/2}{k},$$

where $\mu := \min\left\{\frac{\lambda}{\lambda-2}, \log k\right\}$.

The second determines the typical structure of a set with small doubling; we will use it in Section 3.5.

Theorem 3.3.3 (Theorem 2.5.1 of Chapter 2)**.** *Let* $n, k \in \mathbb{N}$ *and* $2 \leqslant \lambda \leqslant 2^{-120}\frac{k}{(\log n)^3}$*, and let* $2^8\lambda^{1/6}k^{-1/6}\sqrt{\log n} \leqslant \gamma < 2^{-8}$*. For all but at most*

$$e^{-\gamma k}\binom{\lambda k/2}{k}$$

sets $A \subset [n]$ *with* $|A| = k$ *and* $|A + A| \leqslant \lambda k$*, the following holds: there exists* $T \subset A$*, with* $|T| \leqslant 2^9\gamma k$*, such that* $A \setminus T$ *is contained in an arithmetic progression of size* $\lambda k/2 + 2^7\gamma\lambda k$*.*

The upper bounds on λ in Theorems 3.3.2 and 3.3.3 are the reason why we require the bound $k \geqslant (\log n)^4$ in Theorem 3.1.1 and Corollary 3.1.2. We remark that some log-factor is necessary here, since we show in Appendix 2.6 that the conclusions of the theorems fail to hold if $k = o\big(\lambda \log n\big)$. However, it seems plausible that these theorems (and also Theorem 3.1.1 and Corollary 3.1.2) could hold (for λ fixed) whenever $k/\log n \to \infty$.

We will apply Theorem 3.3.1 (in Sections 3.4 and 3.8) via the following corollary.

Corollary 3.3.4. *Let* $0 < \gamma < 1/4$*, let* $S_1, S_2 \subset \mathbb{Z}$ *be intervals, and set*

$$Y := S_1 \cup S_2 \qquad \text{and} \qquad X := (S_1 + S_1) \cup (S_2 + S_2). \tag{3.5}$$

Then there is a family $\mathcal{B} \subset 2^X \times 2^Y$ *of size at most*

$$\exp\left(2^{18}\gamma^{-2}\sqrt{|Y|}\left(\log|Y|\right)^{3/2}\right) \tag{3.6}$$

such that:

(*a*) *For every pair of sets* $U \subset Y$ *and* $W \subset X \setminus (U + U)$*, there exists* $(C, D) \in \mathcal{B}$ *such that* $W \subset C$ *and* $U \subset D$*.*

(*b*) *For every* $(C, D) \in \mathcal{B}$*,*

$$|D| \leqslant \max\left\{(1 + 4\gamma)|Y| - \frac{|C|}{2}, \frac{3|Y|}{\log|Y|}\right\}. \tag{3.7}$$

To deduce Corollary 3.3.4 from Theorem 3.3.1, we will need the following easy lemma, cf.2.3.3.

Lemma 3.3.5. *Let* $\gamma > 0$*, let* $S_1, S_2 \subset \mathbb{Z}$ *be intervals, and set*

$$Y := S_1 \cup S_2 \qquad \text{and} \qquad X := (S_1 + S_1) \cup (S_2 + S_2).$$

Let $C \subset X$ and $D \subset Y$. If

$$|D| \geqslant (1+4\gamma)|Y| - |C|/2$$

then there are at least $\gamma^2|D|^2$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$.

Proof. Suppose first that $S_1 \cap S_2$ is non-empty, so $X = Y + Y$, and let the elements of D be $d_1 < \cdots < d_\ell$. Then $D + D \subset X$ contains the $2\ell - 1$ elements

$$d_1 + d_1 < d_1 + d_2 < \cdots < d_1 + d_\ell < d_2 + d_\ell < \cdots < d_\ell + d_\ell,$$

and $2\ell - 1 \geqslant (2+8\gamma)|Y| - |C| - 1 = |X| - |C| + 8\gamma|Y|$, since $|X| = 2|Y| - 1$. Since $C \subset X$, it follows that there are at least $8\gamma|Y|$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$ and $\{b_1, b_2\} \cap \{d_1, d_\ell\}$ is non-empty. Removing d_1 and d_ℓ from D, and repeating the argument $\gamma|Y|$ times, we obtain $\gamma^2|Y|^2$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$.

When S_1 and S_2 are disjoint, we simply apply the argument above for the two sets $D_1 := D \cap S_1$ and $D_2 := D \cap S_2$. To spell out the details, for each $i \in \{1, 2\}$ there are $2|D_i| - 1$ pairs $(b_1, b_2) \in D_i \times D_i$ with distinct sums such that either $b_1 = \min(D_i)$ or $b_2 = \max(D_i)$. Moreover, $D_1 + D_1$ and $D_2 + D_2$ are disjoint subsets of X, and

$$2|D| - 2 \geqslant (2 + 8\gamma)|Y| - |C| - 2 = |X| - |C| + 8\gamma|Y|,$$

since $|X| = 2|Y| - 2$. As before, it follows that there are at least $8\gamma|Y|$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$ and either $b_1 \in \{\min(D_1), \min(D_2)\}$ or $b_2 \in \{\max(D_1), \max(D_2)\}$. Removing the minimum and maximum elements of D_1 and D_2, and repeating the argument $\gamma|Y|$ times, we obtain $\gamma^2|Y|^2$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$, as claimed. □

Proof of Corollary 3.3.4. Applying Theorem 3.3.1 with $n := |Y|$ and $m := 3|Y|$, we obtain a family $\mathcal{A} \subset 2^{Y+Y} \times 2^Y$, with

$$|\mathcal{A}| \leqslant \exp\left(2^{18}\gamma^{-2}\sqrt{|Y|}\big(\log|Y|\big)^{3/2}\right),$$

satisfying properties (i) and (ii) of the theorem. We claim that

$$\mathcal{B} := \big\{(X \setminus A, B) : (A, B) \in \mathcal{A}\big\} \subset 2^X \times 2^Y$$

satisfies properties (a) and (b) of Corollary 3.3.4.

To show that property (a) holds, let $U \subset Y$ and $W \subset X \setminus (U + U)$, and set $I := (Y + Y) \setminus W$ and $J := U$. Noting that $J \subset Y$ and $J + J \subset I \subset Y + Y$, and that

$$|I| = |(Y + Y) \setminus W| \leqslant 3|Y| = m,$$

it follows from Theorem 3.3.1(i) that there exists $(A, B) \in \mathcal{A}$ with $A \subset I$ and $J \subset B$, and hence there exists $(C, D) = (X \setminus A, B) \in \mathcal{B}$ such that $W \subset C$ and $U \subset D$.

For property (b), let $(C, D) \in \mathcal{B}$, and observe that, by Theorem 3.3.1(ii), either $|D| \leqslant \frac{3|Y|}{\log |Y|}$, or there are at most $\gamma^2|D|^2$ pairs $(b_1, b_2) \in D \times D$ such that $b_1 + b_2 \in C$. In the latter case, we have $|D| \leqslant (1 + 4\gamma)|Y| - |C|/2$, by Lemma 3.3.5. Since $|\mathcal{B}| \leqslant |\mathcal{A}|$, the corollary follows. □

3.4 A probabilistic lemma

Green and Morris [22, Theorem 1.3] used their bounds on the number of sets with small sumset to prove that if S is a random subset of $\mathbb{N}$, with each element included in S independently with probability $1/2$, then

$$\mathbb{P}\Big(\big|\mathbb{N} \setminus (S + S)\big| \geqslant m\Big) = 2^{-m/2+o(m)}.$$

We will use Corollary 3.3.4 to prove the following generalisation of their theorem.

Lemma 3.4.1. *Let $n \in \mathbb{N}$ and $k \in [n]$, set $p := k/n$, and let $m \geqslant 2^{80}p^{-8}$. If S is a uniformly-chosen random subset of $[n]$ of size k, then*

$$\mathbb{P}\Big(\big|\{2, \dots, 2n\} \setminus (S + S)\big| \geqslant m\Big) \leqslant \exp\Big(2^{14}m^{5/6}p^{-7/6}(\log m)^{1/2}\Big) \cdot (1 - p)^{m/2}. \tag{3.8}$$

In the proof of Lemma 3.4.1 we will also use the following well-known inequality (see, e.g., [2, Lemma 5.2]).

Lemma 3.4.2 (Pittel's inequality)**.** *Let $n, k \in \mathbb{N}$ with $k \leqslant n$, and set $p := k/n$. If $\mathcal{I}$ is a monotone decreasing property on $[n]$, then*

$$\mathbb{P}\big(\mathcal{I} \text{ holds for a random } k\text{-subset of } [n]\big) \leqslant 2 \cdot \mathbb{P}\big(\mathcal{I} \text{ holds for a } p\text{-random subset of } [n]\big).$$

Proof. Following the proof in [2], recall that $\mathrm{Bin}(n, p) \leqslant \lceil pn \rceil = k$ holds with probability at least $1/2$. Since $\mathcal{I}$ is monotone decreasing, the claimed bound follows. □

We first prove a simple lemma that will also be useful in Section 3.8.

Lemma 3.4.3. *Let $n \in \mathbb{N}$ and $k \in [n]$, set $p := k/n$, and let $M \in \mathbb{N}$. If S is a uniformly-chosen random subset of $[n]$ of size k, then*

$$\mathbb{P}\Big(\big\{M + 1, \dots, 2n - M + 1\big\} \not\subset S + S\Big) \leqslant \frac{8}{p^2} \cdot (1 - p^2)^{M/2}.$$

Proof. Observe that the left-hand side is at most

$$\sum_{x=M+1}^{2n-M+1} \mathbb{P}\big(x \notin S+S\big) \leqslant 2 \sum_{x=M+1}^{n+1} \mathbb{P}\big(x \notin S+S\big),$$

since, by symmetry, $\mathbb{P}\big(x \notin S+S\big) = \mathbb{P}\big(2n+2-x \notin S+S\big)$. Now, for $x \leqslant n+1$, we can use Pittel's inequality to bound

$$\mathbb{P}\big(x \notin S+S\big) = \mathbb{P}\Big(\bigcap_{i=1}^{\lfloor x/2 \rfloor} \big(\{i \notin S\} \cup \{x-i \notin S\} \big) \Big) \leqslant 2\big(1-p^2\big)^{(x-1)/2}.$$

It follows that

$$\mathbb{P}\Big(\big\{M+1, \dots, 2n-M+1\big\} \not\subset S+S \Big) \leqslant 4 \sum_{x=M+1}^{\infty} \big(1-p^2\big)^{(x-1)/2} \leqslant \frac{8}{p^2}\big(1-p^2\big)^{M/2},$$

as claimed. □

We are now ready to deduce Lemma 3.4.1 from Corollary 3.3.4.

Proof of Lemma 3.4.1. We first use Lemma 3.4.3 to deal with the case that the 'middle' is not covered by $S+S$. To be precise, set $M := \lfloor 4m/p \rfloor$ and let us write $\mathcal{E}$ for the event that $\big\{2M+1, \dots, 2n-2M+1\big\} \subset S+S$. Note that if $\mathcal{E}$ holds, then

$$\{2, \dots, 2n\} \setminus (S+S) \subset X := \big\{2, \dots, 2M\big\} \cup \big\{2n-2M+2, \dots, 2n\big\}.$$

Setting $W := X \setminus (S+S)$, it follows that

$$\mathbb{P}\Big(\big|\big\{2, \dots, 2n\big\} \setminus \big(S+S\big)\big| \geqslant m \Big) \leqslant \mathbb{P}\big(|W| \geqslant m\big) + \mathbb{P}(\mathcal{E}^c).$$

By Lemma 3.4.3, we have

$$\mathbb{P}(\mathcal{E}^c) \leqslant \frac{8}{p^2}\big(1-p^2\big)^M \leqslant \frac{8}{p^2}\big(1-p\big)^m,$$

where the second inequality follows since $1-x^2 \leqslant (1-x)^{x/2}$ for all $0 \leqslant x \leqslant 1$.

To complete the proof, we will use Corollary 3.3.4 to bound the probability that $|W| \geqslant m$. Indeed, applying the corollary to the set

$$Y := \big\{1, \dots, M\big\} \cup \big\{n-M+1, \dots, n\big\},$$

we obtain a family $\mathcal{B} \subset 2^X \times 2^Y$ of containers of size at most

$$\exp\Big(2^{18}\gamma^{-2}\sqrt{M}(\log M)^{3/2}\Big) = \big(1-p\big)^{-\gamma M}, \tag{3.9}$$

where $\gamma > 0$ is chosen so that the equality holds. (Note that the set X defined above is the same as that defined in (3.5).) Using the bounds $1 - p \leqslant e^{-p}$ and $M \geqslant m/p$, and noting that the function $x \mapsto (\log x)^{3/2}/\sqrt{x}$ is decreasing for $x > 2^5$, it follows that

$$\gamma^3 \leqslant \frac{2^{18}(\log M)^{3/2}}{p\sqrt{M}} \leqslant \frac{2^{18}}{\sqrt{pm}} \left(\log \frac{m}{p}\right)^{3/2},$$

and hence, since $M \leqslant 8m/p$,

$$\gamma M \leqslant \frac{8\gamma m}{p} \leqslant \frac{2^9 m^{5/6}}{p^{7/6}} \left(\log \frac{m}{p}\right)^{1/2} < m, \tag{3.10}$$

where the final inequality follows from the assumption that $m \geqslant 2^{80}p^{-8}$. Since $M \geqslant 4m$, it follows from (3.10) that $\gamma < 1/4$, and so this is a valid choice of γ in Corollary 3.3.4.

We next claim that

$$\mathbb{P}\big(|W| \geqslant m\big) \leqslant \sum_{(C,D)\in\mathcal{B}} \mathbb{P}\Big(\big(W \subset C\big) \cap \big(S \cap Y \subset D\big)\Big). \tag{3.11}$$

To see this, observe first that

$$W = X \setminus (S + S) \subset X \setminus \big((S \cap Y) + (S \cap Y)\big)$$

since $S \cap Y \subset S$. By Corollary 3.3.4(a), applied to the pair $U := S \cap Y$ and W, it follows that there exists a pair $(C, D) \in \mathcal{B}$ with $W \subset C$ and $S \cap Y \subset D$.

To bound the right-hand side of (3.11), observe first that

$$\mathbb{P}\big(S \cap Y \subset D\big) \leqslant \binom{n - |Y \setminus D|}{pn} \binom{n}{pn}^{-1} \tag{3.12}$$

for every $(C, D) \in \mathcal{B}$, since S is a uniformly-chosen set of size $k = pn$, and if $S \cap Y \subset D$ then $S \cap (Y \setminus D) = \emptyset$. Moreover, by Corollary 3.3.4(b), if $|W| \geqslant m$ then

$$|Y \setminus D| \geqslant |Y| - |D| \geqslant \frac{m}{2} - 8\gamma M \tag{3.13}$$

for every $(C, D) \in \mathcal{B}$ with $W \subset C$. It follows from (3.9), (3.11), (3.12) and (3.13) that

$$\mathbb{P}\big(|W| \geqslant m\big) \;\leqslant\; (1-p)^{-\gamma M} \binom{n - m/2 + 8\gamma M}{pn} \binom{n}{pn}^{-1} \;\leqslant\; (1-p)^{m/2 - 9\gamma M}, \tag{3.14}$$

where the second inequality follows from the standard binomial inequality

$$\binom{a-c}{b} \leqslant \left(\frac{a-b}{a}\right)^c \binom{a}{b}. \tag{3.15}$$

Finally, combining (3.10) and (3.14), it follows that

$$\mathbb{P}\big(|W| \geqslant m\big) \leqslant \exp\Big(2^{13}m^{5/6}p^{-7/6}(\log m)^{1/2}\Big) \cdot \big(1-p\big)^{m/2},$$

as required. □

We will usually apply Lemma 3.4.1 in the following form. Recall that $\delta = 2^{-18}\lambda^{-2}$.

Corollary 3.4.4. *Let $\lambda \geqslant 3$ and $k, m, b \in \mathbb{N}$, with $m \geqslant 2^{230}\lambda^{20}$ and $b \leqslant \delta k$. There are at most*

$$e^{2\delta m}\left(\frac{\lambda-2}{\lambda}\right)^{m/2}\binom{\lambda k/2}{k-b}$$

sets $A' \subset [\lambda k/2]$ of size $k-b$ such that $\big|[\lambda k] \setminus (A'+A')\big| \geqslant m$.

Proof. We simply apply Lemma 3.4.1 with $p = 2(k-b)/\lambda k$, and observe that

$$\exp\Big(2^{14}m^{5/6}p^{-7/6}(\log m)^{1/2}\Big)\big(1-p\big)^{m/2} \leqslant e^{2\delta m}\left(\frac{\lambda-2}{\lambda}\right)^{m/2},$$

by our bounds on b and m. To spell out the details, note that $p \geqslant 1/\lambda$, and hence

$$2^{14}m^{5/6}p^{-7/6}(\log m)^{1/2} \leqslant \delta m$$

since $\delta = 2^{-18}\lambda^{-2}$ and $m \geqslant 2^{192}\lambda^{19}(\log m)^3$. Now, observe that

$$\big(1-p\big)^{m/2} \leqslant \left(\frac{\lambda-2+2\delta}{\lambda}\right)^{m/2} \leqslant \exp\left(\frac{\delta m}{\lambda-2}\right)\left(\frac{\lambda-2}{\lambda}\right)^{m/2}.$$

Since $\lambda \geqslant 3$, the claimed bound follows. □

Since we will often only need a weaker bound, let us note here, for convenience, that

$$e^{2\delta m}\left(\frac{\lambda-2}{\lambda}\right)^{m/2} \leqslant \left(\frac{\lambda-1}{\lambda}\right)^{m/2}, \tag{3.16}$$

since $\delta < 1/4\lambda$.

3.4.1 Tools and inequalities

To finish this section, let us state some standard tools that we will use in the proof of Theorem 3.1.1. The first is known as Ruzsa's covering lemma (see, e.g., [48, Lemma 2.14]), and was first proved in [44]. For completeness, we give the proof.

Lemma 3.4.5 (Ruzsa's covering lemma)**.** *Let $A, B \subset \mathbb{Z}$ be non-empty sets of integers, and suppose that $|A+B| \leqslant \mu|A|$. Then there exists a set $X \subset B$ with $|X| \leqslant \mu$ such that*

$$B \subset A - A + X.$$

Proof. Let $X \subset B$ be maximal such that the sets $A + x$ for $x \in X$ are disjoint. Observe that $|A + B| \geqslant |A||X|$, and therefore $|X| \leqslant \mu$. Now, since X is maximal, $A + b$ intersects $A + X$ for every $b \in B \setminus X$, and hence $B \subset A - A + X$, as claimed. □

We will also use the following special case of the Plünnecke–Ruzsa inequalities [35, 36, 42], which is also an immediate consequence of Ruzsa's triangle inequality [41].

Lemma 3.4.6 (Plünnecke–Ruzsa inequality)**.** *If $|A + A| \leqslant \lambda|A|$, then $|A - A| \leqslant \lambda^2|A|$.*

Proof. To prove that $|A - A| \cdot |A| \leqslant |A + A|^2$, it suffices to construct an injective map $\varphi\colon (A - A) \times A \to (A + A)^2$. To do so, choose an arbitrary function $f\colon A - A \to A^2$ such that if $f(x) = (a, b)$ then $a - b = x$, and define $\varphi(x, c) \mapsto (a + c, b + c)$, where $f(x) = (a, b)$. To see that φ is injective, observe that $x = (a + c) - (b + c)$ and that $(a, b) = f(x)$. □

In Section 3.7 we will use a simple special case of the following result of Freĭman [16].

Lemma 3.4.7 (Freĭman's $3k - 4$ theorem)**.** *If $|A + A| \leqslant 3|A| - 4$, then $A \subset P$ for some arithmetic progression P of size $|A + A| - |A| + 1$.*

We will also make frequent use of the following standard inequality in the calculations below:

$$\binom{a-c}{b-d} \leqslant \left(\frac{a-c}{a}\right)^{b-d} \left(\frac{b}{a-b}\right)^{d} \binom{a}{b}. \tag{3.17}$$

In particular, note that

$$\binom{\lambda k/2}{k-b} \leqslant \left(\frac{2}{\lambda-2}\right)^{b} \binom{\lambda k/2}{k}. \tag{3.18}$$

We will also use the following inequality once, in Section 3.7.

Observation 3.4.8.

$$\binom{ca}{a} \leqslant \left(\frac{c^c}{(c-1)^{c-1}}\right)^{a},$$

for every $a \in \mathbb{N}$ and $1 < c \in \mathbb{R}$.

Proof. Set $y = (c-1)^{1/c}$, and note that $y/(c-1) = y^{1-c}$. It follows that

$$\left(\frac{c^c}{(c-1)^{c-1}}\right)^a = \left(\left(1+\frac{1}{c-1}\right)(c-1)^{1/c}\right)^{ca}$$
$$= \left(y+y^{1-c}\right)^{ca} = \sum_{i=0}^{ca}\binom{ca}{i}y^{ca-i}\cdot y^{(1-c)i} \geqslant \binom{ca}{a},$$

where the last step follows by considering the term $i = a$. □

3.5 Reducing to an interval

Let us fix $\lambda \geqslant 3$, and for each $n, k \in \mathbb{N}$ define

$$\Lambda = \Lambda(n,k) := \big\{A \subset [n] \,:\, |A| = k,\, |A+A| \leqslant \lambda k\big\}. \tag{3.19}$$

Let us also fix $\varepsilon \in (0,1)$ (since Theorem 3.1.1 holds trivially for $\varepsilon \geqslant 1$) and, writing $\ell(A)$ for the length of the smallest arithmetic progression containing A, define

$$\Lambda^* = \Lambda^*(n,k) := \big\{A \in \Lambda : \ell(A) \leqslant \lambda k/2 + c(\lambda,\varepsilon)\big\}. \tag{3.20}$$

In this section we will prove the following lemma, which reduces the problem of bounding $|\Lambda\setminus\Lambda^*|$ to that of bounding $|\mathcal{I}|$ (see Definition 3.2.1). Recall that $\delta = 2^{-18}\lambda^{-2}$.

Lemma 3.5.1. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $k \geqslant (\log n)^4$ and $k \geqslant 2^{480}\lambda^{20}$. We have*

$$|\Lambda \setminus \Lambda^*| \leqslant \frac{n^2}{k}\cdot|\mathcal{I}| + \exp\left(-\frac{\delta k}{2^{10}\lambda}\right)\binom{\lambda k/2}{k}.$$

To prove Lemma 3.5.1, we will successively refine $\Lambda \setminus \Lambda^*$, at each step showing that some subset with a particular property is small. The first step in the proof of Lemma 3.5.1 is the following stability lemma, which is an almost immediate consequence of Theorem 3.3.3.

Lemma 3.5.2. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $k \geqslant (\log n)^4$ and $k \geqslant 2^{480}\lambda^4$. There are at most*

$$\exp\left(-\frac{\delta k}{2^9\lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \Lambda$ such that

$$|A \setminus P| > \delta k$$

for every arithmetic progression P of size $\lambda k/2$.

Proof. Note first that if $k \geqslant (\log n)^4$ and $k \geqslant 2^{480}\lambda^4$ then $\frac{k}{(\log n)^3} \geqslant k^{1/4} \geqslant 2^{120}\lambda$. Therefore, applying Theorem 3.3.3 with $\gamma = 2^{-9}\lambda^{-1}\delta$, it follows that for all but at most

$$\exp\left(-\frac{\delta k}{2^9\lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \Lambda$, there exists $T \subset A$, with $|T| \leqslant (2^9 + 2^7\lambda)\gamma k \leqslant \delta k$, such that $A \setminus T$ is contained in an arithmetic progression of size $\lambda k/2$, as required. □

The next step is to show that almost all sets $A \in \Lambda$ are contained in an arithmetic progression of length $3\lambda k/2$. Let us write $\mathcal{F}$ for the family of sets $A \in \Lambda$ such that

$$A \subset \{a + jd : -\lambda k/2 \leqslant j \leqslant \lambda k\} \qquad \text{and} \qquad \big|A \setminus \{a + jd : 1 \leqslant j \leqslant \lambda k/2\}\big| \leqslant \delta k$$

for some $a, d \in \mathbb{Z}$.

Lemma 3.5.3. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $k \geqslant (\log n)^4$ and $k \geqslant 2^{480}\lambda^{20}$. Then*

$$|\Lambda \setminus \mathcal{F}| \leqslant \exp\left(-\frac{\delta k}{2^{10}\lambda}\right)\binom{\lambda k/2}{k}.$$

Proof. By Lemma 3.5.2, we may restrict our attention to sets $A \in \Lambda$ such that there exists an arithmetic progression $P = \{a + jd : 0 \leqslant j \leqslant \lambda k/2\}$ such that $|A \setminus P| \leqslant \delta k$. We need to bound the number of sets $A \in \Lambda$ such that

$$A \not\subset \{a + jd : -\lambda k/2 \leqslant j \leqslant \lambda k\} = P + P - P,$$

so let $Z := A \setminus (P + P - P)$ and choose an element $x \in Z$. We will first count the possible sets $A' := A \cap P$, and then (given A') the choices for $B := A \setminus P$. Observe that

$$(x + A') \cap (A' + A') = \emptyset,$$

since $A' \subset P$, and that $|x + A'| = |A'| \geqslant k - \delta k$. Since $A \in \Lambda$, it follows that

$$|A' + A'| \leqslant \lambda k - (k - \delta k) \leqslant \lambda k - k/2.$$

Hence, by Corollary 3.4.4 (applied with $m = k/2 \geqslant 2^{230}\lambda^{20}$), and using (3.16) and (3.18), it follows that, for each $b \leqslant \delta k$, there are at most

$$\left(\frac{\lambda - 1}{\lambda}\right)^{k/4}\binom{\lambda k/2}{k - b} \leqslant \exp\left(-\frac{k}{8\lambda}\right)\binom{\lambda k/2}{k}$$

choices for the set $A' = A \cap P$ such that $|A'| = k - b$.

To count the sets B (given A'), we apply Ruzsa's covering lemma (Lemma 3.4.5) to the pair (A', B) to obtain a set $X \subset B$, with $|X| \leqslant |A'+B|/|A'| \leqslant \lambda k/(k-b) \leqslant 2\lambda$, such that $B \subset A' - A' + X$. Now, by the Plünnecke–Ruzsa inequality (Lemma 3.4.6), we have $|A' - A' + X| \leqslant 2\lambda^3 k$, and hence (choosing X first, and then $B \setminus X$, and recalling that $b \leqslant \delta k$, and that $k \geqslant (\log n)^4$, $k \geqslant 2^{480}\lambda^{20}$ and $\delta = 2^{-18}\lambda^{-2}$), there are at most

$$n^{2\lambda}\binom{2\lambda^3 k}{b-2\lambda} \leqslant \exp\left(\delta k \log\left(2e\lambda^3/\delta\right) + 2\lambda \log n\right) \leqslant \exp\left(\delta^{1/2}k\right)$$

choices for the set B, given a set A' with $|A'| = k - b$.

Combining the bounds above on the number of choices for A' and B, it follows that the number of sets $A \in \Lambda$ with Z non-empty is at most

$$\sum_{b=1}^{\delta k} \exp\left(\delta^{1/2}k - \frac{k}{8\lambda}\right)\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{k}{2^4\lambda}\right)\binom{\lambda k/2}{k},$$

as required. □

Finally, to bound $|\Lambda \setminus \Lambda^*|$ in terms of $|\mathcal{I}|$, we need to map our arithmetic progression P into the interval $[\lambda k/2]$. Lemma 3.5.1 will follow from Lemma 3.5.3 and the following bound.

Lemma 3.5.4. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$. Then*

$$|\mathcal{F} \setminus \Lambda^*| \leqslant \frac{n^2}{k} \cdot |\mathcal{I}|.$$

Proof. We will define a function $\varphi\colon \mathcal{F} \setminus \Lambda^* \to \mathcal{I}$ such that $|\varphi^{-1}(S)| \leqslant n^2/k$ for every $S \in \mathcal{I}$, which will suffice to prove the lemma. To do so, let $A \in \mathcal{F} \setminus \Lambda^*$, and choose $a, d \in \mathbb{N}$ such that

$$A \subset \{a + jd : -\lambda k/2 \leqslant j \leqslant \lambda k\}$$

and such that the sets

$$\{x \in A : x \leqslant a\} \qquad \text{and} \qquad \{x \in A : x > a + \lambda k d/2\} \tag{3.21}$$

are both non-empty and together contain at most δk elements. Indeed, to obtain such a pair, take the arithmetic progression given by the definition of $\mathcal{F}$, and (recalling the definition (3.20) of Λ^*) translate it if necessary so that the sets in (3.21) are both non-empty. Now define

$$\varphi(A) := \{j \in \mathbb{Z} \,:\, a + jd \in A\},$$

and observe that $\varphi(A) \subset \{-\lambda k/2, \dots, \lambda k\}$, and that

$$b\big(\varphi(A)\big) = \big|\{x \in \varphi(A) : x \leqslant 0\}\big| + \big|\{x \in \varphi(A) : x > \lambda k/2\}\big| \leqslant \delta k.$$

Moreover, we have

$$r\big(\varphi(A)\big) = \max\big(\varphi(A)\big) - \min\big(\varphi(A)\big) - \frac{\lambda k}{2} > c(\lambda,\varepsilon),$$

since $A \notin \Lambda^*$, and hence $\varphi(A) \in \mathcal{I}$, as required.

Finally, observe that $|\varphi^{-1}(S)|$ is bounded from above by the number of pairs $(a,d) \in \mathbb{Z}^2$ such that $A := \{a + jd : j \in S\} \subset [n]$. For each set S of size k there are at most

$$\sum_{a=1}^{n} \frac{n-a}{k-1} \leqslant \frac{n^2}{k}$$

such pairs (a,d). Hence $|\varphi^{-1}(S)| \leqslant n^2/k$, as claimed, and the lemma follows. □

We are now ready to prove Lemma 3.5.1.

Proof of Lemma 3.5.1. By Lemmas 3.5.3 and 3.5.4, we have

$$|\Lambda \setminus \Lambda^*| \leqslant |\Lambda \setminus \mathcal{F}| + |\mathcal{F} \setminus \Lambda^*| \leqslant \exp\left(-\frac{\delta k}{2^{10}\lambda}\right)\binom{\lambda k/2}{k} + \frac{n^2}{k}\cdot|\mathcal{I}|,$$

as claimed. □

3.6 Counting the sparse sets in $\mathcal{I}$

Recall that, for any $A \subset \mathbb{Z}$,

$$b(A) = |A \setminus [\lambda k/2]| \qquad \text{and} \qquad r(A) = \max(A) - \min(A) - \lambda k/2,$$

and that $f(\lambda) = 2^{10}\lambda^3$, and (recalling Definition 3.2.1) let us write

$$\mathcal{S} := \Big\{A \in \mathcal{I} \,:\, r(A) > f(\lambda)b(A)\Big\}$$

for the family of 'sparse' sets in $\mathcal{I}$. In this section we will bound the size of $\mathcal{S}$, and hence prove the following quantitative version of Lemma 3.2.2.

Lemma 3.6.1. *Let $\lambda \geqslant 3$ and $\varepsilon \in (0,1)$, and let $k \in \mathbb{N}$. Then*

$$|\mathcal{S}| \leqslant \exp\left(-\frac{c(\lambda,\varepsilon)}{2^9\lambda^2}\right)\binom{\lambda k/2}{k}.$$

For each $B \subset \{-\lambda k/2, \dots, \lambda k\} \setminus [\lambda k/2]$, let us define[2]

$$\mathcal{G}(B) := \big\{A \in \mathcal{I} : A \setminus [\lambda k/2] = B\big\}. \tag{3.22}$$

Recalling Definition 3.2.1, observe that $\mathcal{G}(B) = \emptyset$ if either $\min(B) > 0$ or $\max(B) \leqslant \lambda k/2$, and also if either $|B| > \delta k$ or $r(B) < c(\lambda, \varepsilon)$. We will deduce Lemma 3.6.1 from the following bound on the size of $\mathcal{G}(B)$ by summing over $r \geqslant c(\lambda, \varepsilon)$ and sets B with $|B| < r/f(\lambda)$.

Lemma 3.6.2. *If $B \subset \{-\lambda k/2, \dots, \lambda k\} \setminus [\lambda k/2]$, then*

$$|\mathcal{G}(B)| \leqslant \exp\left(-\frac{r}{2^6\lambda^2}\right)\binom{\lambda k/2}{k-b}$$

where $b = |B|$ and $r = r(B)$.

For each $A \in \mathcal{G}(B)$, set $A' := A \setminus B$. The idea of the proof is simple: if A' contains many elements close to its ends, then we can add these to $\min(B)$ and $\max(B)$, and obtain many elements of $A+A$ outside $[\lambda k]$. Therefore, either $A'+A'$ misses many elements of $[\lambda k]$, in which case we can apply Corollary 3.4.4 to bound the number of choices, or it has few elements close to its ends, and it is straightforward to count sets A' with this property.

To be precise, define

$$Y := \big\{x \leqslant 0 : x - \min(B) \in A'\big\} \cup \big\{x > \lambda k : x - \max(B) \in A'\big\}, \tag{3.23}$$

and set $m(B) := r(B)/8\lambda$. The following bound follows from some simple counting.

Lemma 3.6.3. *If $B \subset \{-\lambda k/2, \dots, \lambda k\} \setminus [\lambda k/2]$, then there are at most*

$$e^{-m(B)}\binom{\lambda k/2}{k-b}$$

sets $A \in \mathcal{G}(B)$ with $|Y| \leqslant m(B)$.

Proof. We claim first that if $r := r(B) \geqslant \lambda k/2$, then there are no such sets $A \in \mathcal{G}(B)$. Indeed, if $A \in \mathcal{G}(B)$ with $|Y| \leqslant m := m(B)$, then $m \geqslant |Y| \geqslant |A'| = k - b \geqslant k/4$, since $b(A) \leqslant \delta k$ for every $A \in \mathcal{I}$. But this implies that $r(B) = 8\lambda m > \lambda k$, which is impossible. Let us therefore assume that $r < \lambda k$, and that $b \leqslant k/4$ and $m \leqslant k/4$.

Now, the number of sets $A \in \mathcal{G}(B)$ with $|Y| \leqslant m$ is at most

$$\sum_{\ell=0}^{m}\binom{r}{\ell}\binom{\lambda k/2 - r}{k-b-\ell} \leqslant \sum_{\ell=0}^{m}\left(\frac{er}{\ell}\right)^{\ell}\left(1-\frac{2r}{\lambda k}\right)^{k-b-\ell}\left(\frac{2}{\lambda-2}\right)^{\ell}\binom{\lambda k/2}{k-b}, \tag{3.24}$$

[2]Note that we include sets of $\mathcal{I}\setminus\mathcal{S}$ in $\mathcal{G}(B)$; we will not need to use the bound $r(A) > f(\lambda)b(A)$ when bounding the size of $\mathcal{G}(B)$ (we use it only when counting the choices for the set B), and we shall also want to reuse our bounds on $|\mathcal{G}(B)|$ in Section 3.7, below, where we will be dealing with dense sets.

where the inequality holds by (3.17). Now, observe that

$$\left(1-\frac{2r}{\lambda k}\right)^{k-b-\ell} \leqslant \left(1-\frac{2r}{\lambda k}\right)^{k/2} \leqslant \exp\left(-\frac{r}{\lambda}\right) = e^{-8m},$$

since $b+m \leqslant k/2$, and that

$$\sum_{\ell=0}^{m}\left(\frac{er}{\ell}\cdot\frac{2}{\lambda-2}\right)^{\ell} \leqslant \sum_{\ell=0}^{m}\left(\frac{2^4e\lambda}{\lambda-2}\cdot\frac{m}{\ell}\right)^{\ell} \leqslant (m+1)\left(\frac{2^4e\lambda}{\lambda-2}\right)^{m} \leqslant \left(2^7e\right)^m,$$

since $r = 8\lambda m$ and $\lambda \geqslant 3$, and since $(C/x)^x$ is increasing for $x < C/e$. It follows that the right-hand side of (3.24) (and hence the number of sets $A \in \mathcal{G}(B)$ with $|Y| \leqslant m$) is at most

$$\binom{2}{e}^{7m}\binom{\lambda k/2}{k-b} \leqslant e^{-m}\binom{\lambda k/2}{k-b},$$

as claimed. □

It remains to count sets $A \in \mathcal{G}(B)$ with $|Y| > m$. To do so, set $X := A' + A'$, and observe that X and Y are disjoint subsets of $A + A$. Since $|A + A| \leqslant \lambda k$, it follows that

$$\big|[\lambda k] \setminus X\big| \geqslant |Y| > m(B). \tag{3.25}$$

We will use Corollary 3.4.4 to count the sets with $|[\lambda k] \setminus X| \geqslant m(B)$.

Lemma 3.6.4. *If $B \subset \{-\lambda k/2, \dots, \lambda k\} \setminus [\lambda k/2]$, then there are at most*

$$\left(\frac{\lambda-1}{\lambda}\right)^{m(B)/2}\binom{\lambda k/2}{k-b}$$

sets $A \in \mathcal{G}(B)$ with $|[\lambda k] \setminus X| \geqslant m(B)$.

Proof. We want to bound the number of sets $A' \subset [\lambda k/2]$, with $|A'| = k - b$, such that $|[\lambda k] \setminus (A' + A')| \geqslant m := m(B)$. Recall that $|B| \leqslant \delta k$ and $r(B) \geqslant c(\lambda, \varepsilon)$ (otherwise $\mathcal{G}(B)$ is empty), and note that therefore $m = r(B)/8\lambda \geqslant 2^{230}\lambda^{20}$. It follows, by Corollary 3.4.4 and (3.16), that there are at most

$$\left(\frac{\lambda-1}{\lambda}\right)^{m/2}\binom{\lambda k/2}{k-b}$$

sets $A \in \mathcal{G}(B)$ such that $|[\lambda k] \setminus (A' + A')| \geqslant m$, as claimed. □

We can now easily deduce the claimed upper bound on the size of $\mathcal{G}(B)$.

Proof of Lemma 3.6.2. By (3.25), $|\mathcal{G}(B)|$ is at most the sum of the bounds in Lemmas 3.6.3 and 3.6.4. Recalling that $m(B) = r(B)/8\lambda$, this gives

$$|\mathcal{G}(B)| \leqslant \big(e^{-m(B)} + e^{-m(B)/2\lambda}\big)\binom{\lambda k/2}{k-b} \leqslant \exp\bigg(-\frac{r(B)}{2^5\lambda^2}\bigg)\binom{\lambda k/2}{k-b},$$

as required. □

Lemma 3.6.1 is a straightforward consequence.

Proof of Lemma 3.6.1. Fix b and r, and consider the sets $B \subset \{-\lambda k/2, \dots, \lambda k\} \setminus [\lambda k/2]$ with $|B| = b$ and $r(B) = r$. We may assume that $r > f(\lambda)b$ and $r \geqslant c(\lambda, \varepsilon)$, since otherwise $\mathcal{G}(B) \cap \mathcal{S} = \emptyset$. The number of choices for B (given b and r) is therefore at most

$$\binom{r}{b} \leqslant \exp\bigg(\frac{r}{2^7\lambda^2}\bigg)$$

since $r/b > f(\lambda) = 2^{10}\lambda^3$. By Lemma 3.6.2, it follows that

$$\big|\big\{A \in \mathcal{S} : b(A) = b,\, r(A) = r\big\}\big| \leqslant \exp\bigg(-\frac{r}{2^7\lambda^2}\bigg)\binom{\lambda k/2}{k-b} \leqslant \exp\bigg(-\frac{r}{2^8\lambda^2}\bigg)\binom{\lambda k/2}{k},$$

where the second inequality follows from (3.18), since $r/b > f(\lambda)$.

Summing over choices of $r \geqslant c(\lambda, \varepsilon)$ and $b < r/f(\lambda)$, it follows that

$$|\mathcal{S}| \leqslant \sum_{r \geqslant c(\lambda,\varepsilon)} \frac{r}{f(\lambda)} \exp\bigg(-\frac{r}{2^8\lambda^2}\bigg)\binom{\lambda k/2}{k} \leqslant \exp\bigg(-\frac{c(\lambda,\varepsilon)}{2^9\lambda^2}\bigg)\binom{\lambda k/2}{k},$$

as required. □

3.7 Counting the moderately dense sets

Recall from Definition 3.2.1 and (3.2) the definitions of $b(A)$, $r(A)$ and $\mathcal{I}$, and let us write

$$\mathcal{D} := \Big\{A \in \mathcal{I} \,:\, r(A) \leqslant f(\lambda)b(A)\Big\}$$

for the family of 'dense' sets in $\mathcal{I}$, where $f(\lambda) = 2^{10}\lambda^3$. In the next two sections we will prove the following quantitative version of Lemma 3.2.3.

Lemma 3.7.1. *Let $\lambda \geqslant 3$ and $\varepsilon \in (0,1)$, and let $k \in \mathbb{N}$. Then*

$$|\mathcal{D}| \leqslant \exp\bigg(-\frac{c(\lambda,\varepsilon)}{2^{18}\lambda^2\log\lambda}\bigg)\binom{\lambda k/2}{k}.$$

Let us fix $\lambda \geqslant 3$, $\varepsilon \in (0,1)$ and $k \in \mathbb{N}$ until the end of the proof of Lemma 3.7.1. In this section, we will deal with some relatively easy cases using the method of the previous section. Observe that

$$b(A) \geqslant \frac{c(\lambda,\varepsilon)}{f(\lambda)} \geqslant 2^{470}\lambda^{27} \tag{3.26}$$

for every $A \in \mathcal{D}$, since $r(A) \geqslant c(\lambda,\varepsilon)$ for every $A \in \mathcal{I}$, and by the definition (3.1) of $c(\lambda,\varepsilon)$.

For convenience, let us define, for each $b \in \mathbb{N}$ and $\mu \geqslant 1$,

$$\mathcal{D}(b,\mu) := \Big\{ A \in \mathcal{D} : |B| = b \text{ and } |(B+B)\setminus[\lambda k]| = \mu b, \text{ where } B = A \setminus [\lambda k/2] \Big\}.$$

We begin by bounding the number of sets $A \in \mathcal{D}(b,\mu)$ such that $r(A) \geqslant 2^{11}\mu b$.

Lemma 3.7.2. *Let $b \in \mathbb{N}$ and $\mu \geqslant 1$. If $r \geqslant 2^{11}\mu b$, then there are at most*

$$\exp\left(-\frac{r}{2^7\lambda^2}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}(b,\mu)$ with $r(A) = r$.

The first step is to use Theorem 3.3.2 to bound the number of choices for $B = A \setminus [\lambda k/2]$. We will use the following lemma several times in the proof of Lemma 3.7.1.

Lemma 3.7.3. *Let $b \in \mathbb{N}$ and $\mu > 2$. There are at most*

$$e^{2\delta b}\left(\frac{\mu-2}{2}\right)^b \left(\frac{\mu}{\mu-2}\right)^{\mu b/2} \tag{3.27}$$

sets B such that $B = A \setminus [\lambda k/2]$ for some $A \in \mathcal{D}(b,\mu)$.

We will use the following observation in the proof of Lemma 3.7.3, and then again (several times) in the applications below.

Observation 3.7.4.

$$(x-2)\cdot\left(\frac{x}{x-2}\right)^{x/2} \leqslant (y-2)\left(\frac{y}{y-2}\right)^{x/2}$$

for every $x, y > 2$.

Proof. Set $q(x,y) := (x/y)^{x/2}\cdot\big((y-2)/(x-2)\big)^{(x-2)/2}$, and observe that

$$\log\big(q(x,y)^{2/x}\big) = \frac{2}{x}\cdot\log\frac{x}{y} + \frac{x-2}{x}\cdot\log\left(\frac{x(y-2)}{y(x-2)}\right) \leqslant \log\left(\frac{2}{x}\cdot\frac{x}{y} + \frac{x-2}{x}\cdot\frac{x(y-2)}{y(x-2)}\right) = 0,$$

using the concavity of the log function. □

Proof of Lemma 3.7.3. Set $B_1 := \{x \in B : x \leqslant 0\}$ and $B_2 := \{x \in B : x > \lambda k/2\}$, and recall from (3.26) that $b \geqslant 2^{470}\lambda^{27}$, and that $\delta = 2^{-18}\lambda^{-2}$. Observe first that, since $r(A) \leqslant f(\lambda)b$ for

each $A \in \mathcal{D}(b,\mu)$, for each $i \in \{1,2\}$ there are at most

$$\binom{f(\lambda)b}{b^{3/4}} \leqslant \exp\left(b^{3/4}\log b\right) \leqslant e^{\delta b} \tag{3.28}$$

choices for the set B_i with $|B_i| \leqslant b^{3/4}$. Moreover, by Lemma 3.4.7, if $|B_i + B_i| \leqslant 2|B_i|$, then B_i is contained in an arithmetic progression of size $|B_i|+1$, and so in this case there are at most $r^3 \leqslant 2^{30}\lambda^9 b^3 \leqslant e^{\delta b}$ choices for B_i.

Now, set $b_i = |B_i|$ and $\mu_i b_i = |B_i + B_i|$, and suppose that $b_i \geqslant b^{3/4}$, and $\mu_i > 2$. Observe that

$$\mu_i \leqslant \frac{2f(\lambda)b}{b_i} \leqslant 2^{-36}\frac{b_i}{\left(\log(f(\lambda)b)\right)^3}, \tag{3.29}$$

since $b \geqslant 2^{470}\lambda^{27}$ implies $b \geqslant 2^{74}f(\lambda)^2\left(\log(f(\lambda)b)\right)^6$. Hence, by Theorem 3.3.2, the number of choices for B_i (given b_i and μ_i) is at most

$$\exp\left(2^9\mu_i^{1/6}b_i^{5/6}\log b_i\sqrt{\log(f(\lambda)b)}\right)\binom{\mu_i b_i/2}{b_i} \leqslant e^{\delta b}\binom{\mu_i b_i/2}{b_i}, \tag{3.30}$$

where the inequality holds since $\mu_i^{1/6}b_i^{5/6} \leqslant 4\lambda \cdot b^{5/6}$, by (3.29), and $b \geqslant 2^{470}\lambda^{27}$.

Now, by Observations 3.4.8 and 3.7.4, it follows that

$$\binom{\mu_i b_i/2}{b_i} \leqslant \left(\frac{\mu_i - 2}{2}\cdot\left(\frac{\mu_i}{\mu_i-2}\right)^{\mu_i/2}\right)^{b_i} \leqslant \left(\frac{\mu-2}{2}\right)^{b_i}\left(\frac{\mu}{\mu-2}\right)^{\mu_i b_i/2}. \tag{3.31}$$

Since $\mu b = \mu_1 b_1 + \mu_2 b_2$, the lemma follows from (3.28), (3.30) and (3.31). □

We are now ready to prove Lemma 3.7.2.

Proof of Lemma 3.7.2. Observe first that if $\mu \leqslant 2$, then B is contained in two arithmetic progressions of combined size at most $|B|+2$, by Lemma 3.4.7, and so in this case there are at most r^6 choices for B. By Lemma 3.6.2, it follows that there are at most

$$r^6\exp\left(-\frac{r}{2^6\lambda^2}\right)\left(\frac{2}{\lambda-2}\right)^b\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{r}{2^7\lambda^2}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}(b,\mu)$ with $r(A) = r \geqslant 2^{11}b$, where we used (3.18) and (3.26).

Now, if $\mu > 2$, then by Lemma 3.7.3 and Observation 3.7.4 there are at most

$$e^{2\delta b}\left(\frac{\lambda-2}{2}\right)^b\left(\frac{\lambda}{\lambda-2}\right)^{\mu b/2} \tag{3.32}$$

sets B such that $B = A \setminus [\lambda k/2]$ for some $A \in \mathcal{D}(b,\mu)$. Moreover, by Lemma 3.6.3 and (3.18), for each such set B there are at most

$$e^{-m}\binom{\lambda k/2}{k-b} \leqslant e^{-m}\left(\frac{2}{\lambda-2}\right)^b\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{G}(B)$ with $|Y| \leqslant m := m(B)$, where $m(B) = r(B)/8\lambda$, and Y is as defined in (3.23). Noting that if $r(B) \geqslant 2^{11}\mu b$ then $\mu b \leqslant 2^{-8}\lambda m$, it follows that there are at most

$$e^{2\delta b}\left(\frac{\lambda}{\lambda-2}\right)^{\mu b/2} e^{-m}\binom{\lambda k/2}{k} \leqslant e^{-m/2}\binom{\lambda k/2}{k}$$

choices for A with $|Y| \leqslant m$.

Suppose next that $|Y| \geqslant m$ and $|Y \cap (B+B) \setminus [\lambda k]| \leqslant m/2$. Since

$$|A'+A'| + |Y \cup (B+B) \setminus [\lambda k]| \leqslant |A+A| \leqslant \lambda k,$$

it follows that $|[\lambda k] \setminus (A'+A')| \geqslant \mu b + m/2 \geqslant 2^{230}\lambda^{20}$. Therefore, by Corollary 3.4.4 and (3.18), for each set B such that $B = A \setminus [\lambda k/2]$ for some $A \in \mathcal{D}(b,\mu)$, there are at most

$$\exp\left(2\delta \cdot (\mu b + m/2)\right)\left(\frac{\lambda-2}{\lambda}\right)^{\mu b/2+m/4}\left(\frac{2}{\lambda-2}\right)^b\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{G}(B)$ such that $|Y| \geqslant m$ and $|Y \cap (B+B) \setminus [\lambda k]| \leqslant m/2$. By (3.32), and recalling that $\mu b \leqslant 2^{-8}\lambda m$ and $\delta = 2^{-18}\lambda^{-2}$, it follows that there are at most

$$e^{\lambda\delta m}\left(\frac{\lambda-2}{\lambda}\right)^{m/4}\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{m}{4\lambda}\right)\binom{\lambda k/2}{k}$$

choices for A in this case.

Finally, suppose that $|Y| \geqslant m$ and $|Y \cap (B+B) \setminus [\lambda k]| > m/2$, and consider the set

$$Z := \big\{x \in [\lambda k/2] : x + \min(B) \in (B+B) \setminus [\lambda k] \ \text{ or } \ x + \max(B) \in (B+B) \setminus [\lambda k]\big\}.$$

Observe that $|A' \cap Z| > m/2$ and $|Z| \leqslant |(B+B) \setminus [\lambda k]|$. It follows that, given B such that $B = A \setminus [\lambda k/2]$ for some $A \in \mathcal{D}(b,\mu)$, the number of choices for A' is at most

$$\sum_{\ell > m/2}\binom{\mu b}{\ell}\binom{\lambda k/2}{k-b-\ell} \leqslant \sum_{\ell > m/2}\left(\frac{e\mu b}{\ell}\cdot\frac{2}{\lambda-2}\right)^{\ell}\binom{\lambda k/2}{k-b} \leqslant 2^{-m}\left(\frac{2}{\lambda-2}\right)^b\binom{\lambda k/2}{k},$$

where the inequalities follow from (3.18) and the bounds $\mu b \leqslant 2^{-8}\lambda m$ and $\lambda \geqslant 3$, which together imply that

$$\frac{2e\mu b}{m}\cdot\frac{2}{\lambda-2} \leqslant \frac{e\lambda}{2^5(\lambda-2)} \leqslant \frac{1}{4}.$$

By (3.32), and recalling again that $\mu b \leqslant 2^{-8}\lambda m$, it follows that there are at most

$$e^{2\delta b}\left(\frac{\lambda}{\lambda-2}\right)^{\mu b/2} 2^{-m}\binom{\lambda k/2}{k} \leqslant 2^{-m/2}\binom{\lambda k/2}{k}$$

choices for A in this case, as required. □

It will be useful in the next section (which deals with the case $r \leqslant 2^{11}\mu b$) to be able to assume that $\mu = \Theta(\lambda)$. The next lemma, which follows from Corollary 3.4.4, provides a suitable bound on the size of $\mathcal{D}(b,\mu)$ when this is not the case.

Lemma 3.7.5. *Let $b \in \mathbb{N}$. If $r \leqslant 2^{11}\mu b$ and $\mu \notin (\lambda/2, 2\lambda-2)$, then there are at most*

$$\exp\left(-\frac{r}{2^{16}\lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}(b,\mu)$ with $r(A) = r$.

Proof. For each $A \in \mathcal{D}(b,\mu)$, set $A' := A \cap [\lambda k/2]$ and $B := A \setminus [\lambda k/2]$, and observe that $\big|[\lambda k] \setminus (A'+A')\big| \geqslant |(B+B)\setminus[\lambda k]| = \mu b$, since $|A+A| \leqslant \lambda k$. Hence, by Corollary 3.4.4 applied with $m = \mu b \geqslant 2^{230}\lambda^{20}$, and using (3.18), there are at most

$$\exp\big(2\delta\cdot\mu b\big)\left(\frac{\lambda-2}{\lambda}\right)^{\mu b/2}\left(\frac{2}{\lambda-2}\right)^{b}\binom{\lambda k/2}{k}, \tag{3.33}$$

choices for the set A'. Next, by Lemma 3.7.3, for each $\mu > 2$ there are at most

$$e^{2\delta b}\left(\frac{\mu-2}{2}\right)^{b}\left(\frac{\mu}{\mu-2}\right)^{\mu b/2} \tag{3.34}$$

sets B with $B = A \setminus [\lambda k/2]$ for some $A \in \mathcal{D}(b,\mu)$.

Now, if $\mu \geqslant 2\lambda-2$, then by Observation 3.7.4 (applied with with $x=\mu$ and $y = 2\lambda-2$), and recalling that $\lambda \geqslant 3$ and $\delta = 2^{-18}\lambda^{-2}$, the product of (3.33) and (3.34) is at most

$$\exp\big(3\delta\cdot\mu b\big)\cdot 2^{b}\left(\frac{\lambda-1}{\lambda}\right)^{\mu b/2}\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{\mu b}{2^5\lambda}\right)\binom{\lambda k/2}{k}.$$

Alternatively, if $2 < \mu \leqslant \lambda/2$, then by Observation 3.7.4 (applied with $x=\mu$ and $y=\lambda/2$), and noting that in this case $\lambda > 4$, the product of (3.33) and (3.34) is at most

$$\exp\big(3\delta\cdot\mu b\big)\cdot\left(\frac{\lambda-2}{\lambda-4}\right)^{\lambda b/4}\left(\frac{\lambda-4}{2\lambda-4}\right)^{b}\binom{\lambda k/2}{k} \leqslant e^{-b/16}\binom{\lambda k/2}{k}.$$

Finally, if $\mu \leqslant 2$ then B is contained in two arithmetic progressions of combined size at most $|B|+2$, by Lemma 3.4.7, and so in this case there are at most $r^6 \leqslant 2^{60}\lambda^{18}b^6 \leqslant e^{\delta b}$ choices for B.

Noting that $\mu b \in \{2b-1, 2b\}$, it follows from (3.33) that there are

$$e^{7\delta b}\left(\frac{\lambda-2}{\lambda}\right)^b\left(\frac{2}{\lambda-2}\right)^b\binom{\lambda k/2}{k} \leqslant e^{-b/4}\binom{\lambda k/2}{k}$$

choices for A. Since $r \leqslant 2^{11}\mu b$, in each case the claimed bound follows. □

3.8 Counting the very dense sets with containers

It remains to bound the size of the family

$$\mathcal{D}^*(b,\mu) := \Big\{A \in \mathcal{D}(b,\mu) : r(A) \leqslant 2^{11}\mu b\Big\}$$

of *very dense* sets, for each $\lambda/2 \leqslant \mu \leqslant 2\lambda - 2$. To do so, we will once again use the container theorem (Theorem 3.3.1), but this time our application of it will be rather different.

To state the version of Corollary 3.3.4 we will use, we need a little additional notation. First, for each $b \in \mathbb{N}$, set $Y(b) := Y_1 \cup Y_2$ and $X(b) := (Y_1 + Y_1) \cup (Y_2 + Y_2)$, where

$$Y_1 := \big\{0, \dots, g(\lambda)b\big\}, \qquad \text{and} \qquad Y_2 := \big\{\lambda k/2 - g(\lambda)b, \dots, \lambda k/2\big\},$$

where $g(\lambda) := 2^{15}\lambda^2$. Moreover, define $M(A) := [\lambda k] \setminus (A + A)$ and

$$\mathcal{T}(b) := \big\{A \in \mathcal{I} : b(A) = b \text{ and } M(A) \subset X(b)\big\}.$$

Our key tool in this section will be the following immediate consequence of Corollary 3.3.4.

Corollary 3.8.1. *For each $b \in \mathbb{N}$, there exists a family $\mathcal{B}(b) \subset 2^{X(b)} \times 2^{Y(b)}$ of size at most*

$$\exp\left(2^{50}\lambda^2 b^{5/6}\,(\log \lambda b)^{3/2}\right)$$

such that:

(a) *For each $A \in \mathcal{T}(b)$, there exists $(C, D) \in \mathcal{B}(b)$ with $M(A) \subset C$ and $A \cap Y(b) \subset D$.*

(b) *For every $(C, D) \in \mathcal{B}(b)$,*

$$|D| \leqslant \max\left\{|Y(b)| + |Y(b)|^{5/6} - \frac{|C|}{2}, \frac{3|Y(b)|}{\log|Y(b)|}\right\}.$$

Proof. We apply Corollary 3.3.4 with $\varepsilon = |Y(b)|^{-1/6}/4$, $S_1 = Y_1$ and $S_2 = Y_2$. The bound on the size of $\mathcal{B}(b)$ follows from (4.2) since $|Y(b)| = 2g(\lambda)b + 2 \leqslant 2^{17}\lambda^2 b$ and

$$2^{22}\big(2^{17}\lambda^2 b\big)^{5/6}\,\big(\log 2^{17}\lambda^2 b\big)^{3/2} \leqslant 2^{50}\lambda^2 b^{5/6}\,(\log \lambda b)^{3/2},$$

and the bound on $|D|$ for each $(C,D) \in \mathcal{B}(b)$ follows from (3.7). Finally, for each $A \in \mathcal{T}(b)$ we apply Corollary 3.3.4(a) with $U := A \cap Y(b)$ and $W := M(A) \subset X(b) \setminus (U+U)$. It follows that there exists $(C,D) \in \mathcal{B}(b)$ such that $M(A) \subset C$ and $A \cap Y(b) \subset D$, as claimed. □

Before bounding the number of sets in each container, let's quickly observe that, by our choice of $g(\lambda)$, most sets of $\mathcal{D}^*(b,\mu)$ are also in $\mathcal{T}(b)$. Recall that $\delta = 2^{-18}\lambda^{-2}$.

Lemma 3.8.2. *For each $b \leqslant \delta k$ and $\lambda/2 \leqslant \mu \leqslant 2\lambda - 2$, there are at most*

$$e^{-b}\binom{\lambda k/2}{k} \tag{3.35}$$

sets $A \in \mathcal{D}^*(b,\mu)$ such that $M(A) \not\subset X(b)$.

Proof. Let A be a uniformly random k-subset of $[-2^{11}\mu b, \lambda k/2 + 2^{11}\mu b]$, and observe that

$$\mathbb{P}\big(M(A) \not\subset X(b)\big) \leqslant \mathbb{P}\Big(\big\{M'+1,\dots,\lambda k - M' + 1\big\} \not\subset A + A\Big),$$

where $M' := 2g(\lambda)b$. By Lemma 3.4.3 (applied with $n = \lambda k/2 + 2^{12}\mu b + 1$ and $M = M' + 2^{12}\mu b + 2$), it follows that

$$\mathbb{P}\big(M(A) \not\subset X(b)\big) \leqslant \frac{8}{p^2} \cdot \big(1-p^2\big)^{M/2} \leqslant \exp\Big(-g(\lambda)b/\lambda^2\Big),$$

where $p = k\big(\lambda k/2 + 2^{12}\mu b + 1\big)^{-1} \geqslant 1/\lambda$, since $\mu b \leqslant 2\delta\lambda k$ and $\delta \leqslant 2^{-15}$. Now, observe that

$$\binom{\lambda k/2 + 2^{12}\mu b + 1}{k} \leqslant \exp\big(2^{14}b\big)\binom{\lambda k/2}{k}.$$

Hence, recalling that $g(\lambda) = 2^{15}\lambda^2$, there are at most

$$\exp\big(-2^{15}b + 2^{14}b\big)\binom{\lambda k/2}{k} \leqslant e^{-b}\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}^*(b,\mu)$ with $M(A) \not\subset X(b)$, as claimed. □

To deduce Lemma 3.2.3 from Corollary 3.8.1, we will need to bound the size of the containers in $\mathcal{B}(b)$. To do so, we will partition the containers according to the size of C; we first bound those containers with C large. Set $\alpha := \big(2^4\lambda\log\lambda\big)^{-1}$.

Lemma 3.8.3. *Let $b \leqslant \delta k$ and $\lambda/2 \leqslant \mu \leqslant 2\lambda - 2$. For each $(C,D) \in \mathcal{B}(b)$ with*

$$|C| \geqslant (1+2\alpha)\mu b,$$

there are at most

$$e^{-\alpha b/4}\binom{\lambda k/2}{k},$$

sets $A \in \mathcal{T}(b) \cap \mathcal{D}^*(b,\mu)$ *such that* $A \cap Y(b) \subset D$.

Proof. Recall that $|Y(b)| = 2^{16}\lambda^2 b + 2$, and that $b \geqslant 2^{470}\lambda^{27}$, by (3.26), and observe that therefore $|Y(b)|^{5/6} \leqslant \alpha\mu b/2$. By Corollary 3.8.1($b$) and our assumption on $|C|$, it follows that

$$|D| \leqslant |Y(b)| - \frac{(1+\alpha)\mu b}{2},$$

and therefore if $A \cap Y(b) \subset D$ then A' misses the set $Y(b) \setminus D \subset [\lambda k/2]$, which has size at least $(1+\alpha)\mu b/2$. Hence, using (3.15) and (3.18), it follows[3] that there are at most

$$\binom{\lambda k/2 - (1+\alpha)\mu b/2}{k-b} \leqslant e^{5\delta b}\left(\frac{\lambda-2}{\lambda}\right)^{(1+\alpha)\mu b/2}\left(\frac{2}{\lambda-2}\right)^b \binom{\lambda k/2}{k} \tag{3.36}$$

choices for $A' = A \cap [\lambda k/2]$ such that $A \cap Y(b) \subset D$.

Now, choose the set $B = A \setminus [\lambda k/2]$, using Lemma 3.7.3 and Observation 3.7.4 to bound the number of choices. It follows from (3.27) and (3.36) that the number of sets A is at most

$$e^{5\delta b}\left(\frac{\lambda-2}{\lambda}\right)^{\alpha\mu b/2}\binom{\lambda k/2}{k} \leqslant e^{-\alpha b/4}\binom{\lambda k/2}{k},$$

as claimed, where in the final step we used the bounds $\delta \leqslant 2^{-10}\lambda^{-2}$ and $\mu \geqslant \lambda/2$. □

When C is small, we will prove the following bound.

Lemma 3.8.4. *Let* $b \leqslant \delta k$ *and* $\lambda/2 \leqslant \mu \leqslant 2\lambda - 2$. *For each* $(C,D) \in \mathcal{B}(b)$ *with*

$$|C| \leqslant (1+2\alpha)\mu b,$$

there are at most

$$e^{-b/8}\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{T}(b) \cap \mathcal{D}^*(b,\mu)$ *such that* $M(A) \subset C$.

Proof. Let us first count the choices for the set $A' = A \cap [\lambda k/2]$, given sets $B = A \setminus [\lambda k/2]$ and $M(A) \subset C$. Recall that $M(A) = [\lambda k] \setminus (A + A)$, so $|M(A)| \geqslant \mu b$ (since $|A + A| \leqslant \lambda k$ and $|(B + B) \setminus [\lambda k]| = \mu b$). We set $F(A) := M(A) - \{\min(B), \max(B)\}$, and claim that

$$\big|F(A) \cap [\lambda k/2]\big| \geqslant |M(A)| \qquad \text{and} \qquad F(A) \cap A' = \emptyset.$$

Indeed, note first that $F(A) \cap A' = \emptyset$ holds because $M(A)$ and $A + A$ are disjoint. Now, recall that $M(A) \subset X(b)$ and $r(A) \leqslant 2^{11}\mu b \leqslant k/4$ for every $A \in \mathcal{T}(b) \cap \mathcal{D}^*(b,\mu)$, and observe that

[3] Here we use the bounds $b \leqslant \delta k$ and $\mu \leqslant 2\lambda - 2$, which imply that $\big(1 + \frac{2b}{(\lambda-2)k}\big)^{(1+\alpha)\mu b/2} \leqslant \exp\big(5\delta b\big)$.

$2g(\lambda)b \leqslant k/4$, since $b \leqslant \delta k$ and $\delta = 2^{-18}\lambda^{-2} = (8g(\lambda))^{-1}$. It follows that the sets $M(A)-\min(B)$ and $M(A) - \max(B)$ do not overlap, which implies the inequality.

It follows, using (3.15) and (3.18), that we have at most (cf. (3.36))

$$\binom{\lambda k/2 - \mu b}{k-b} \leqslant e^{8\delta b}\left(\frac{\lambda-2}{\lambda}\right)^{\mu b}\left(\frac{2}{\lambda-2}\right)^{b}\binom{\lambda k/2}{k}$$

choices for A', given B and $M(A)$.

Now, observe that there are at most

$$\binom{(1+2\alpha)\mu b}{\geqslant \mu b} = \binom{(1+2\alpha)\mu b}{\leqslant 2\alpha\mu b} \leqslant \exp\left(\frac{\mu b}{2\lambda}\right)$$

ways of choosing $M(A) \subset C$, since $\alpha = (2^4\lambda\log\lambda)^{-1}$. Finally, we again use Lemma 3.7.3 and Observation 3.7.4 to bound the number of choices for the set $B = A \setminus [\lambda k/2]$. Combining this with the bounds above, and recalling that $\mu \geqslant \lambda/2$ and $\delta \leqslant 2^{-8}\lambda^{-1}$, it follows that there are at most

$$e^{9\delta b}\exp\left(\frac{\mu b}{2\lambda}\right)\left(\frac{\lambda-2}{\lambda}\right)^{\mu b/2}\binom{\lambda k/2}{k} \leqslant e^{-b/8}\binom{\lambda k/2}{k},$$

$A \in \mathcal{T}(b) \cap \mathcal{D}^*(b,\mu)$ with $M(A) \subset C$, as claimed. □

We are finally ready to prove Lemma 3.7.1.

Proof of Lemma 3.7.1. Let us fix $b, r \in \mathbb{N}$ and $\mu \geqslant 1$, and bound the number of sets $A \in \mathcal{D}(b,\mu)$ with $r(A) = r$. Recall first that if $r \geqslant 2^{11}\mu b$ then, by Lemma 3.7.2, there are at most

$$\exp\left(-\frac{r}{2^7\lambda^2}\right)\binom{\lambda k/2}{k}$$

such sets, and if $r \leqslant 2^{11}\mu b$ and either $\mu \leqslant \lambda/2$ or $\mu \geqslant 2\lambda - 2$, then by Lemma 3.7.5 there are at most

$$\exp\left(-\frac{r}{2^{16}\lambda}\right)\binom{\lambda k/2}{k}$$

such sets. Now, if $r \leqslant 2^{11}\mu b$ and $\lambda/2 \leqslant \mu \leqslant 2\lambda - 2$, then by Lemma 3.8.2 there are at most

$$e^{-b}\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{r}{2^{12}\lambda}\right)\binom{\lambda k/2}{k}$$

such sets that are not in $\mathcal{T}(b)$. Moreover, by Corollary 3.8.1, there exists a family $\mathcal{B}(b)$ of size at most

$$\exp\left(2^{50}\lambda^2 b^{5/6}(\log\lambda b)^{3/2}\right)$$

such that for every $A \in \mathcal{T}(b)$, there exists $(C, D) \in \mathcal{B}(b)$ with $M(A) \subset C$ and $A \cap Y(b) \subset D$. Finally, by Lemmas 3.8.3 and 3.8.4, for each $(C, D) \in \mathcal{B}(b)$ there are at most

$$e^{-\alpha b/4}\binom{\lambda k/2}{k} \leqslant \exp\left(-\frac{r}{2^{12}\lambda^2 \log \lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{T}(b) \cap \mathcal{D}^*(b, \mu)$ such that $M(A) \subset C$ and $A \cap Y(b) \subset D$.

Combining these bounds, it follows that there are at most

$$\exp\left(2^{50}\lambda^2 b^{5/6}(\log \lambda b)^{3/2}\right)\exp\left(-\frac{r}{2^{15}\lambda^2 \log \lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}(b, \mu)$ with $r(A) = r$. Now, summing over choices of $b \leqslant r$ and $\mu \leqslant 2r/b$ such that $\mu b \in \mathbb{N}$, and recalling that $r \geqslant 2^{480}\lambda^{30}$, it follows that there are at most

$$\exp\left(-\frac{r}{2^{16}\lambda^2 \log \lambda}\right)\binom{\lambda k/2}{k}$$

sets $A \in \mathcal{D}$ with $r(A) = r$.

Finally, summing over $r \geqslant c(\lambda, \varepsilon)$, we deduce that

$$|\mathcal{D}| \leqslant \exp\left(-\frac{c(\lambda, \varepsilon)}{2^{18}\lambda^2 \log \lambda}\right)\binom{\lambda k/2}{k},$$

as claimed. □

3.9 The proof of Theorem 3.1.1

In this section we will prove the following quantitative version of Theorem 3.1.1, which allows us to control the typical structure of A when $\lambda = k^{o(1)}$. Recall that $\delta = 2^{-18}\lambda^{-2}$.

Theorem 3.9.1. *Let $\lambda \geqslant 3$, let $n, k \in \mathbb{N}$ with $k \geqslant (\log n)^4$ and $k \geqslant 2^{480}\lambda^{20}$, and let $\varepsilon > e^{-\delta^2 k}$. Let $A \subset [n]$ be chosen uniformly at random from the sets with $|A| = k$ and $|A + A| \leqslant \lambda k$. Then there exists an arithmetic progression P with*

$$A \subset P \qquad \text{and} \qquad |P| \leqslant \frac{\lambda k}{2} + c(\lambda, \varepsilon)$$

with probability at least $1 - \varepsilon$.

There is only one piece still missing in the proof of Theorem 3.9.1: a lower bound on $|\Lambda|$. The following very simple bound will suffice for our current purposes; a stronger lower bound (at least, for large λ) will be proved in Section 3.10.

Lemma 3.9.2. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $\lambda k \leqslant n$. Then*

$$\big|\big\{A \subset [n] : |A| = k,\, |A + A| \leqslant \lambda k\big\}\big| \geqslant \frac{1}{\lambda^3} \cdot \frac{n^2}{k} \binom{\lambda k/2}{k}.$$

Proof. We consider, for each arithmetic progression P of length $\lambda k/2$ in $[n]$, all subsets $A \subset P$ of size k containing both endpoints of P. All of these sets are distinct, and all satisfy $|A+A| \leqslant \lambda k$. There are at least $n^2/2\lambda k$ choices for the arithmetic progression, and therefore

$$|\Lambda| \geqslant \frac{n^2}{2\lambda k} \binom{\lambda k/2 - 2}{k-2} \geqslant \frac{n^2}{\lambda^3 k} \binom{\lambda k/2}{k},$$

as claimed. □

We can now deduce Theorem 3.9.1 from Lemmas 3.5.1, 3.6.1, 3.7.1 and 3.9.2.

Proof of Theorem 3.9.1. For simplicity, we will assume that $\lambda k \leqslant n$; the case $\lambda k > n$ is dealt with in Appendix A. By Lemma 3.5.1, since $\varepsilon > e^{-\delta^2 k}$, we have

$$|\Lambda \setminus \Lambda^*| \leqslant \frac{n^2}{k} \cdot |\mathcal{I}| + \exp\left(-\frac{\delta k}{2^{10}\lambda}\right) \binom{\lambda k/2}{k} \leqslant \frac{n^2}{k} \cdot |\mathcal{I}| + \frac{\varepsilon}{2\lambda^3} \binom{\lambda k/2}{k}.$$

Now, by Lemmas 3.6.1 and 3.7.1, and recalling that $\mathcal{S} \cup \mathcal{D} = \mathcal{I}$, we have

$$|\mathcal{I}| = |\mathcal{S}| + |\mathcal{D}| \leqslant 2 \cdot \exp\left(-\frac{c(\lambda,\varepsilon)}{2^{18}\lambda^2 \log \lambda}\right) \binom{\lambda k/2}{k} \leqslant \frac{\varepsilon}{2\lambda^3} \binom{\lambda k/2}{k}$$

since $c(\lambda, \varepsilon) = 2^{18}\lambda^2 \log \lambda \cdot \log(1/\varepsilon) + 2^{480}\lambda^{30}$. By Lemma 3.9.2, it follows that

$$|\Lambda \setminus \Lambda^*| \leqslant \frac{\varepsilon}{\lambda^3} \cdot \frac{n^2}{k} \binom{\lambda k/2}{k} \leqslant \varepsilon |\Lambda|,$$

as required. □

When $\lambda \in (2, 3)$, the proof of Theorem 3.9.1 implies the following weaker bound.

Theorem 3.9.3. *For each $\gamma > 0$, there exists a constant $C(\gamma) > 0$ such that the following holds. Let $2 + \gamma \leqslant \lambda \leqslant 3$ and $\varepsilon > 0$ be fixed, let n be sufficiently large, and let $k \geqslant (\log n)^4$. If $A \subset [n]$ is chosen uniformly at random from the sets with $|A| = k$ and $|A + A| \leqslant \lambda k$, then there exists an arithmetic progression P with*

$$A \subset P \qquad \text{and} \qquad |P| \leqslant \frac{\lambda k}{2} + C(\gamma) \log(1/\varepsilon)$$

with probability at least $1 - 2\varepsilon$.

Theorem 3.9.3 follows by repeating the proof of Theorem 3.9.1, replacing the condition $\lambda \geqslant 3$ by the condition $\lambda \geqslant 2+\gamma$, and the conditions $r(A) \geqslant c(\lambda,\varepsilon)$ and $k \geqslant 2^{480}\lambda^{20}$ by the conditions $r(A) \geqslant C(\gamma)$ and k is sufficiently large. We leave the (straightforward, though somewhat tedious) details to the reader.

To finish the section, let us quickly deduce Corollary 3.1.2.

Proof of Corollary 3.1.2. The lower bound follows from Lemma 3.9.2 (see also Proposition 3.10.2, below), so it remains to prove the upper bound. To do so, note that (by increasing the implicit constant in the upper bound if necessary) we may assume that $k \geqslant 2^{480}\lambda^{20}$, and hence we may apply Theorem 3.9.1 with $\varepsilon := 1/2$. Since there are at most n^2/k arithmetic progressions of length $\lambda k/2 + c(\lambda,\varepsilon)$, it follows that

$$|\Lambda| \leqslant \frac{2n^2}{k}\binom{\lambda k/2 + c(\lambda,\varepsilon)}{k} \leqslant \exp\left(\frac{2c(\lambda,\varepsilon)}{\lambda}\right)\frac{n^2}{k}\binom{\lambda k/2}{k} \leqslant \exp\big(c(\lambda,\varepsilon)\big)\cdot\frac{n^2}{k}\binom{\lambda k/2}{k},$$

as required. □

3.10 The lower bounds

In this section, we prove lower bounds for the size of Λ, and for the typical size of the smallest arithmetic progression containing a set $A \in \Lambda$. The bounds we obtain indicate that the upper bounds in Theorem 3.1.1 and Corollary 3.1.2 are not far from best possible. We begin with the construction for the typical structure, which is very simple.

Proposition 3.10.1. *Given $\lambda \geqslant 4$, let $\varepsilon > 0$ be sufficiently small, and let $n, k \in \mathbb{N}$ be sufficiently large. If $A \subset [n]$ is chosen uniformly at random from the sets with $|A| = k$ and $|A+A| \leqslant \lambda k$, then with probability at least ε,*

$$|P| \geqslant \frac{\lambda k}{2} + 2^{-6}\lambda^2\log(1/\varepsilon)$$

for every arithmetic progression P containing A.

Proof. Set $r := 2^{-6}\lambda^2\log(1/\varepsilon)$, and consider the family of sets A of the form $A' \cup \{v\}$, where $1 \in A' \subset [\lambda k/2 - 8r/\lambda]$ with $|A'| = k-1$, and $v = \lambda k/2 + r$. We claim that most such sets satisfy $|A+A| \leqslant \lambda k$. Indeed, since $A'+A' \subset [\lambda k - 16r/\lambda]$, this holds as long as the set $\{x \in A' : x > \lambda k/2 - r - 16r/\lambda\}$ has at most $16r/\lambda$ elements. If $k \geqslant 16r/\lambda$, then the expected number of elements of this set is

$$\frac{k-2}{\lambda k/2 - 8r/\lambda - 1}\cdot\left(r + \frac{8r}{\lambda}\right) \leqslant \frac{2(\lambda+8)}{\lambda-1}\cdot\frac{r}{\lambda} \leqslant \frac{8r}{\lambda},$$

it follows by Markov's inequality that $|A+A| \leqslant \lambda k$ with probability at least $1/2$, as claimed.

The number of sets A as above is

$$\binom{\lambda k/2 - 8r/\lambda - 1}{k-2} \geqslant \frac{4}{\lambda^2} \exp\left(-\frac{16r}{\lambda(\lambda-1)}\right)\binom{\lambda k/2}{k} \geqslant \frac{\sqrt{\varepsilon}}{\lambda^2}\binom{\lambda k/2}{k},$$

since $k \geqslant 16r/\lambda$ and $r \leqslant 2^{-5}\lambda(\lambda-1)\log(1/\varepsilon)$. Now, for each $a \in [n/\lambda k]$ and $b \in [n/4]$, and each set A as above, we apply the linear map $x \mapsto ax + b$ to A. We obtain at least

$$\frac{n^2}{4\lambda k} \cdot \frac{1}{2} \cdot \frac{\sqrt{\varepsilon}}{\lambda^2}\binom{\lambda k/2}{k} \geqslant \varepsilon^{2/3} \cdot \frac{n^2}{k}\binom{\lambda k/2}{k} \tag{3.37}$$

distinct sets $A \subset [n]$ with $|A| = k$ and $|A+A| \leqslant \lambda k$. Finally, recalling the upper bound on $|\Lambda|$ given by Corollary 3.1.2, and that ε was chosen sufficiently small, it follows that the right-hand side of (3.37) is at least $\varepsilon|\Lambda|$, as required. □

Obtaining our lower bound on the size of $|\Lambda|$ will be slightly more delicate.

Proposition 3.10.2. *If $\lambda \geqslant 2^{24}$ and $n, k \in \mathbb{N}$ are sufficiently large, then*

$$\left|\left\{A \subset [n] : |A| = k,\, |A+A| \leqslant \lambda k\right\}\right| \geqslant \exp\left(2^{-8}\lambda^{1/2}\right)\frac{n^2}{k}\binom{\lambda k/2}{k}. \tag{3.38}$$

We will use the following easy application of the FKG inequality for the hypergeometric distribution, see, e.g., [7, Lemma 3.2].

Lemma 3.10.3. *Let G be a graph with n vertices, m edges and ℓ loops. Let R be a uniformly chosen random subset of k vertices, where $k \leqslant \lfloor n/2 \rfloor$. If $\mathcal{B}$ is the event that R is an independent set, then*

$$\mathbb{P}(\mathcal{B}) \geqslant \exp\left(-\frac{9mk^2}{2n^2} - \frac{3\ell k}{n}\right) - \exp\left(-\frac{k}{16}\right).$$

Proof. This follows immediately from [7, Lemma 3.2], applied with (in the notation of [7]) $m = k$ and $\eta = 1/2$, and the sets B_i being the edges and loops of G, and using the fact that $1 - x \geqslant e^{-2x}$ for $0 \leqslant x \leqslant 3/4$. □

Proof of Proposition 3.10.2. Set $c := 2^{-8}$ and $r := 2c\lambda^{3/2}$. We will first prove that there are at least $\exp\left(2c\lambda^{1/2}\right)\binom{\lambda k/2}{k}$ subsets $A \subset [\lambda k/2 + r]$ of size k with $|A+A| \leqslant \lambda k$, each containing the endpoints 1 and $\lambda k/2 + r$. Since this bound can be applied in each of the (at least) $n^2/4\lambda k$ arithmetic progressions of length $\lambda k/2 + r$ in $[n]$, and since the sets A obtained for different arithmetic progressions are distinct, it will follow that

$$|\Lambda| \geqslant \frac{n^2}{4\lambda k} \cdot \exp\left(2c\lambda^{1/2}\right)\binom{\lambda k/2}{k} \geqslant \exp\left(c\lambda^{1/2}\right)\frac{n^2}{k}\binom{\lambda k/2}{k},$$

as required.

To prove the claimed bound, let R be a uniformly chosen subset of $[2, \lambda k/2 + r - 1]$ with exactly $k-2$ elements, and set $A := R \cup \{1, \lambda k/2 + r\}$. Observe first that (using (3.17))

$$\binom{\lambda k/2 + r - 2}{k-2} \geqslant \frac{1}{\lambda^2}\left(\frac{\lambda k + 2r}{\lambda k}\right)^k \binom{\lambda k/2}{k} \geqslant \exp\left(3c\lambda^{1/2}\right)\binom{\lambda k/2}{k}, \tag{3.39}$$

since $r = 2c\lambda^{3/2}$ and λ and k were chosen sufficiently large. It will therefore suffice to prove that $|A + A| \leqslant \lambda k$ with probability at least $\exp\big(-c\lambda^{1/2}\big)$. To do so, define

$$A' := \big\{x \in A : x \leqslant \lambda k/2 - r\big\} \qquad \text{and} \qquad B := \big\{x \in A : x > \lambda k/2 - r\big\},$$

and set $b := 16c\lambda^{1/2}$. Observe that $\mathbb{E}[|B|] \leqslant 4r/\lambda = b/2$, and hence

$$\mathbb{P}\big(|B + B| \geqslant b^2\big) \leqslant \mathbb{P}\big(|B| \geqslant b\big) \leqslant \exp\big(-c\lambda^{1/2}\big), \tag{3.40}$$

by Hoeffding's inequality. We claim that, setting $X := [\lambda k - 2r + 1, \lambda k - 2r + b^2]$, we have

$$\mathbb{P}\big((A' + B) \cap X = \emptyset\big) \geqslant 2 \cdot \exp\big(-c\lambda^{1/2}\big). \tag{3.41}$$

Before proving (3.41), observe that, together with (3.39) and (3.40), it will suffice to deduce the proposition. Indeed, if $(A' + B) \cap X = \emptyset$ and $|B + B| \leqslant b^2 = |X|$, then

$$|A + A| \leqslant \lambda k - 2r + |(A' + B) \setminus [\lambda k - 2r]| + |B + B| \leqslant \lambda k,$$

since $A' + A' \subset [\lambda k - 2r]$ and $A' + B \subset [\lambda k]$, and noting that $b^2 = 2^8 c^2 \lambda \leqslant 4c\lambda^{3/2} = 2r$.

To prove (3.41) we will use Lemma 3.10.3. To do so, we define a graph G with vertex set $[\lambda k/2 + r]$ and edge set

$$E(G) = \big\{xy \,:\, x \leqslant \lambda k/2 - r,\, y > \lambda k/2 - r \text{ and } x + y \in X\big\} \cup \big\{x \,:\, x + \lambda k/2 + r \in X\big\}.$$

Observe that if R is an independent set in G, then $(A' + B) \cap X = \emptyset$. Note that G has at most $2rb^2 \leqslant 2^{10}c^3\lambda^{5/2}$ edges and at most $b^2 = 2^8c^2\lambda$ loops, and that

$$\frac{9 \cdot 2^{10}c^3\lambda^{5/2}k^2}{2(\lambda k/2 + r)^2} + \frac{3 \cdot 2^8 c^2 \lambda k}{\lambda k/2 + r} \leqslant 2^{15}c^3\lambda^{1/2} + 2^{11}c^2 \leqslant c\lambda^{1/2} - 1,$$

since $c = 2^{-8}$ and $\lambda \geqslant 2^{30}$. It follows by Lemma 3.10.3 that

$$\mathbb{P}\big((A' + B) \cap X = \emptyset\big) \geqslant \exp\big(-c\lambda^{1/2} + 1\big) - \exp\big(-k/16\big) \geqslant 2 \cdot \exp\big(-c\lambda^{1/2}\big)$$

as required, since k is sufficiently large. This completes the proof of Proposition 3.10.2. □

Chapter 4

On the singularity of random symmetric matrices

4.1 Introduction

The work in this chapter was done jointly with Letícia Mattos, Robert Morris and Natasha Morrison. In this chapter we prove the following theorem

Theorem 4.1.1. *There exists $c > 0$ such that if M_n is a uniformly-chosen random $n \times n$ symmetric matrix with entries in the set $\{-1, 1\}$, then*

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant \exp\big(-c\sqrt{n}\big) \tag{4.1}$$

for all sufficiently large $n \in \mathbb{N}$.

The main new ingredient in our approach is an inverse Littlewood–Offord theorem (see Theorem 4.1.2, below) which applies to vectors $v \in \mathbb{Z}_p^n$ that exhibit a very mild amount of 'structure'. In order to motivate this theorem, let us begin by recalling the problem of Littlewood and Offord [28], introduced in 1943 during their study of random polynomials. For any abelian group G, integer $n \in \mathbb{N}$, and vector $v \in G^n$, define

$$\rho(v) := \max_{a \in G} \mathbb{P}\bigg(\sum_{i=1}^{n} u_i v_i = a\bigg),$$

where u is a uniformly-chosen random element of $\{-1, 1\}^n$. Littlewood and Offord [28] proved that $\rho(v) = O\big(n^{-1/2} \log n\big)$ when $G = \mathbb{Z}$, and Erdős [12] improved this to $\rho(v) = O\big(n^{-1/2}\big)$, which is best possible, using Sperner's theorem. The problem of proving upper bounds on $\rho(v)$ (under various assumptions) has become known as the 'Littlewood–Offord problem', and has been extensively studied, perhaps most notably by Frankl and Füredi [15] and by Halász [23].

Costello, Tao and Vu [9] proved a 'quadratic' Littlewood–Offord inequality, and used it to deduce their bound (1.1) on the probability that M_n is singular.

Inverse Littlewood–Offord theory, the study of the structure of vectors $v \in G^n$ such that $\rho(v)$ is (relatively) large, was initiated by Tao and Vu [50], and has since played an important role in the study of random matrices, see for example the work of Rudelson and Vershynin [38, 39], Tao and Vu [51], Nguyen and Vu [31, 32], and the surveys [33, 40, 54]. Our inverse Littlewood–Offord theorem differs from these earlier results in several important ways: it is designed for $\mathbb{Z}_p$, rather than $\mathbb{Z}$; it gives (weak) structural information about every vector $v \in \mathbb{Z}_p^n$ such that $\rho(v) \geqslant 4/p$ (most earlier results gave stronger structural information, but required a condition of the form $\rho(v) \geqslant n^{-C}$ for some $C > 0$); and it is designed to facilitate iteration. We remark that the statement of Theorem 4.1.2 was inspired by the method of hypergraph containers, a technique that was introduced several years ago by Balogh, Morris and Samotij [5] and (independently) Saxton and Thomason [47], and which has turned out to have a large number of applications in extremal and probabilistic combinatorics. We refer the interested reader to the survey [6] for more details.

Given a vector $v \in \mathbb{Z}_p^n$, let $|v| := \big|\{i \in [n] : v_i \neq 0\}\big|$ denote the size of the support of v, and for each subset $Y \subset [n]$, let us write v_Y for the restriction of v to the coordinates of Y. Our inverse Littlewood–Offord theorem is as follows.

Theorem 4.1.2. *Let p be a prime. There exists a family $\mathcal{C}$ of subsets of $\mathbb{Z}_p$, with*

$$|\mathcal{C}| \leqslant \exp\Big(2^{12}\big(\log p\big)^2\Big), \tag{4.2}$$

such that for each $n \in \mathbb{N}$, and every $v \in \mathbb{Z}_p^n$ with $\rho(v) \geqslant 4/p$ and $|v| \geqslant 2^{18}\log p$, there exist sets $B(v) \in \mathcal{C}$ and $Y = Y(v) \subset [n]$, with $n/4 \leqslant |Y| \leqslant n/2$, such that

$$\big|\{i \in [n] : v_i \notin B(v)\}\big| \leqslant \frac{n}{4} \qquad \text{and} \qquad |B(v)| \leqslant \frac{2^{16}}{\rho(v_Y)\sqrt{|v|}}. \tag{4.3}$$

In order to motivate the statement of the theorem above, it is instructive to consider the example of a vector whose entries are chosen uniformly (and independently) at random from a d-dimensional generalised arithmetic progression[1] Q. For such a vector, $\rho(v)$ is typically of order $|Q|^{-1}|v|^{-d/2}$ (as long as $|v|$ is not too small), and the $p^{\Theta(d)}$ such progressions are natural 'containers' for these vectors. This example suggests that one might be able to prove a stronger version of Theorem 4.1.2, in which most 'containers' (members of the family $\mathcal{C}$) are significantly smaller than the maximum given in (4.3). However, without significant additional ideas such a strengthening would *not* imply a significant improvement over the bound in Theorem 4.1.1, see the discussion in Section 4.2.2 for more details.

[1] This is a set of the form $\{a + j_1\ell_1 + \cdots + j_d\ell_d : 1 \leqslant j_i \leqslant k_i\}$ for some $a, \ell_1, \ldots, \ell_d \in \mathbb{Z}_p$ and $k_1, \ldots, k_d \in \mathbb{N}$.

We remark that the sets $Y(v)$, whose appearance in Theorem 4.1.2 might appear somewhat unnatural at first sight, will play a vital role in our application of the theorem to prove Theorem 4.1.1. More precisely, we will use the sets $Y(v)$ to maintain independence as we reveal various rows and columns of the matrix, see Section 4.2.1 for more details. Let us also mention here that the family of containers $\mathcal{C}$ will be defined explicitly (see (4.11), below), but we will only need the properties stated in the theorem. The proof of Theorem 4.1.2 uses the probabilistic method (for those readers familiar with the container method, we choose the 'fingerprint' randomly), and a classical 'anticoncentration lemma' proved by Halász [23] (Lemma 4.3.1, below), see Section 4.3 for more details.

The rest of the chapter is organised as follows: in Section 4.2 we give an overview of the proof, in Section 4.3 we prove Halász's Theorem, in Section 4.4 we prove Theorem 4.1.2, in Section 4.5 we deduce Theorem 4.1.1, and in Section 4.6 we provide the proof of a 'reduction lemma' of Ferber and Jain [14] (whose proof was based on the method of [9, 31]).

4.2 An overview of the proof

In this section we will outline the proof of our inverse Littlewood–Offord theorem, and the deduction of Theorem 4.1.1. The first step is to apply the method of [9, 14, 31] to reduce the problem to bounding the quantity

$$q_n(\beta) := \max_{w \in \mathbb{Z}_p^n} \mathbb{P}\Big(\exists\, v \in \mathbb{Z}_p^n \setminus \{0\} : M_n \cdot v = w \text{ and } \rho(v) \geqslant \beta\Big), \tag{4.4}$$

for some suitable $\beta = \exp\big(-\Theta(\sqrt{n})\big)$ and a prime $p = \Theta\big(1/\beta\big)$. To be precise, we will use the following lemma, which was proved by Ferber and Jain [14] using techniques developed by Costello, Tao and Vu [9] and Nguyen [31]. Note that the dependence of $q_n(\beta)$ on the prime p is suppressed in the notation.

Lemma 4.2.1. *Let $n \in \mathbb{N}$, and let $p > 2$ be prime. For every $\beta > 0$,*

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant 16n \sum_{m=n-1}^{2n-3} \left(\beta^{1/8} + \frac{q_m(\beta)}{\beta}\right).$$

Since Lemma 4.2.1 was not stated explicitly in [14], for completeness we provide the proof in Appendix 4.6. Using our inverse Littlewood–Offord theorem (Theorem 4.1.2), we will prove the following bound on $q_n(\beta)$.

Lemma 4.2.2. *Let $n \in \mathbb{N}$, and let $2 < p \leqslant \exp\big(2^{-10}\sqrt{n}\big)$ be prime. If $\beta \geqslant 4/p$, then*

$$q_n(\beta) \leqslant 2^{-n/4}.$$

Theorem 4.1.1 is easily deduced from Lemmas 4.2.1 and 4.2.2.

Proof of Theorem 4.1.1, assuming Lemmas 4.2.1 and 4.2.2. Let $n \in \mathbb{N}$ be sufficiently large, let $\exp\left(2^{-11}\sqrt{n}\right) \leqslant p \leqslant 2 \cdot \exp\left(2^{-11}\sqrt{n}\right)$ be prime, and set $\beta := 4/p$. By Lemmas 4.2.1 and 4.2.2, it follows that

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant 16n \sum_{m=n-1}^{2n-3} \left((4/p)^{1/8} + \frac{p}{2^{m/4+2}} \right) \leqslant \exp\big(-c\sqrt{n}\big)$$

for some $c > 2^{-15}$, as required. □

We will prove Theorem 4.1.2 in Section 4.4, and deduce Lemma 4.2.2 in Section 4.5. Although the proofs are not especially technical, some of the definitions may initially seem somewhat surprising. In order to motivate these definitions, we will now provide a brief outline of the argument, beginning with the deduction of Lemma 4.2.2 from Theorem 4.1.2.

4.2.1 An outline of the proof of Lemma 4.2.2

We will bound $q_n(\beta)$ using the first moment method: for each $w \in \mathbb{Z}_p^n$, we will bound the expected number of vectors $v \in \mathbb{Z}_p^n \setminus \{0\}$ with $\rho(v) \geqslant \beta$ such that $M_n \cdot v = w$. In order to do so, we will use Theorem 4.1.2 to partition the collection of vectors $v \in \mathbb{Z}_p^n \setminus \{0\}$ with $|v| \geqslant \lambda\sqrt{n}$ and $\rho(v) \geqslant \beta$ into a collection $\mathcal{U}$ of at most n^{cn} 'containers' (for some $\lambda > 0$ and $c > 0$); we will then apply the union bound inside each container.[2] The bound we obtain on the probability that $M_n \cdot v = w$ will depend on the container of v, and the containers are chosen so that (for each $C \in \mathcal{U}$ and $w \in \mathbb{Z}_p^n$) the expected number of vectors $v \in C$ with $M_n \cdot v = w$ is at most $n^{-c'n}$ (for some $c' > c$). The claimed bound then follows by summing over containers, and then dealing with the vectors with small support separately.

To construct the container of a vector $v \in \mathbb{Z}_p^n \setminus \{0\}$, we repeatedly apply Theorem 4.1.2, in each step bounding the number of choices for v_X, for some set $X \subset [n]$. Revealing the rows of M_n corresponding to X, we will be able to use the probability that $M_{X\times[n]} \cdot v = w_X$, and the bound on $|B(v_Z)|$ given by (4.3), to 'beat' this number of choices. We continue this iteration until we have chosen all but $O(\sqrt{n})$ of the non-zero entries of v.

To describe a single step of this iteration, assume that we have already revealed a subset of the rows of M_n, and let $Z \subset [n]$ denote the set of rows that have not yet been revealed. By Theorem 4.1.2, we may associate, to each vector $v \in \mathbb{Z}_p^n \setminus \{0\}$ with $\rho(v) \geqslant \beta \geqslant 4/p$, sets

$$Y(v_Z) \subset Z, \quad B(v_Z) \subset \mathbb{Z}_p \quad \text{and} \quad X(v_Z) := \big\{ i \in Z \setminus Y(v_Z) : v_i \in B(v_Z) \big\}.$$

[2] We remark that this is one of several ways in which our method differs from the 'standard' container method; usually one would like to avoid using the union bound inside a container.

In this step we will 'reveal' the rows of M_n corresponding to $X = X(v_Z)$, and sum over the choices for $v_i \in B(v_Z)$ for each $i \in X$. We claim that

$$\mathbb{P}\big(M_{X\times[n]} \cdot v = w_X\big) \leqslant \rho(v_Y)^{|X|}. \tag{4.5}$$

Indeed, since X and $Y = Y(v_Z)$ are disjoint subsets of Z, the entries of $M_{X\times Y}$ are all independent (of each other, and of the previously revealed entries of M_n), so the claimed bound holds by the definition of ρ (see the proof of Lemma 4.5.4, below, for the details).

It remains to count the number of choices for the sets X, Y and $B(v_Z)$, and for the entry $v_i \in B(v_Z)$ for each $i \in X(v_Z)$. We have at most $2^{|Z|}$ choices each for X and Y, and at most

$$\exp\Big(2^{12}\big(\log p\big)^2\Big) \leqslant \exp\big(2^{-8}n\big)$$

choices for the set $B(v_Z)$, by (4.2) and our choice of p. Now, it follows from (4.3) and our bounds on $|Y|$ that $|X| \geqslant |Z|/4$, and hence the total number of choices for these sets (over all steps of the process) is at most $\exp\big(2^{-6}n\log n\big)$, see Lemma 4.5.3, below.

Finally, we have at most $|B(v_Z)|^{|X|}$ choices for the vector v_X. Multiplying this by the probability bound (4.5), and using the bound on $|B(v_Z)|$ given by (4.3), we obtain

$$|B(v_Z)|^{|X|}\rho(v_Y)^{|X|} \leqslant \left(\frac{2^{16}}{\sqrt{|v|}}\right)^{|X|} \leqslant n^{-|X|/4},$$

since $|v| \geqslant \lambda\sqrt{n}$. Since $|X| \geqslant n/4$ in the first step, this will be sufficient to prove the claimed bound on the expected number of vectors $v \in C$ with $M_n \cdot v = w$.

4.2.2 A natural barrier at $\exp\big(-\sqrt{n\log n}\big)$

In this section we explain why a simple union bound (like that described in Section 4.2.1) cannot be used to prove a significantly stronger bound than that in Theorem 4.1.1, without 'reusing' some of the randomness in M_n. Let $m \leqslant n$, and consider the family of vectors v whose entries are chosen from the set $\{-N,\dots,N\}$, where $N = n^{-1/2}2^m$. For a typical such v,

$$\rho(v_{[k]}) \geqslant \rho(v) \geqslant 2^{-m}$$

for every $k \geqslant m$, and $\rho(v_{[k]}) \geqslant 2^{-k}$ for every $k < m$.

Now, it follows that the natural bound

$$\mathbb{P}\big(M_n \cdot v = 0\big) \leqslant \prod_{k=1}^{n} \rho(v_{[k]}),$$

which uses all of the randomness in M_n, cannot give a stronger bound than

$$\mathbb{P}\big(M_n \cdot v = 0\big) \leqslant 2^{-m(n-m)} \prod_{k=1}^{m} 2^{-k} = 2^{-mn+m^2/2+O(n)}.$$

Since there are $N^n = 2^{mn} n^{-n/2}$ choices for the vector v, a union bound (over these vectors) gives (at best) a bound of $n^{-n/2} 2^{m^2/2+O(n)}$, which is small only if $m \leqslant \sqrt{n \log n}$.

It follows that our proof method only has a chance of working if $p \leqslant \exp\big(\sqrt{n \log n}\big)$. However, if we are working over $\mathbb{Z}_p$ then we cannot hope to prove a stronger bound on the singularity probability than $1/p$. Indeed, let M_{n-1} be the matrix obtained by removing the first row and column of M_n, and suppose that $\det(M_{n-1}) \neq 0$ and $\langle u,\, M_{n-1}^{-1} \cdot u \rangle = m_{11}$, where $u \in \{-1, 1\}^{n-1}$ is obtained from the first row of M_n by deleting the entry m_{11}. Then there exists a vector $w := (1, -M_{n-1}^{-1} \cdot u) \in \mathbb{Z}_p^n \setminus \{0\}$ with $M \cdot w = 0$, and hence $\det(M_n) = 0$. Since one would expect $\langle u,\, M_{n-1}^{-1} \cdot u \rangle$ to be (roughly) uniformly distributed over $\mathbb{Z}_p$, it seems reasonable to expect that $\det(M_n) = 0$ occurs with probability at least $1/p$.

4.3 Halász's inequality, and the inverse Littlewood–Offord theorem

In this section we will state the main tool we will use in the proof of Theorem 4.1.2, a classical Littlewood–Offord theorem due to Halász [23]. We will also prepare the reader for the proof in the next section by providing some motivation for the way we define our family of containers.

In order to state Halász's inequality, we need a little preparation. First, let us define multiplication on $\mathbb{Z}_p$ as follows: if $x, y \in \mathbb{Z}_p$, then the product $x \cdot y \in \mathbb{Z}$ is obtained by projecting x and y onto elements of $\{0, 1, \dots, p-1\}$ in the usual way, and then multiplying in $\mathbb{Z}$. Let $\|\cdot\|$ denote the distance to the nearest integer, and for each $n \in \mathbb{N}$, prime p and vector $v \in \mathbb{Z}_p^n$, define the *level sets* of v to be

$$T_t(v) := \left\{ k \in \mathbb{Z}_p \,:\, \sum_{i=1}^{n} \left\| \frac{k \cdot v_i}{p} \right\|^2 \leqslant t \right\}, \tag{4.6}$$

for each $t \geqslant 0$.

We can now state the lemma of Halász [23].

Lemma 4.3.1 (Halász's Anticoncentration Lemma)**.** *Let $n \in \mathbb{N}$ and p be prime, and let $v \in \mathbb{Z}_p^n \setminus \{0\}$. Then*

$$\rho(v) \leqslant \frac{3}{p} + \frac{6|T_\ell(v)|}{p\sqrt{\ell}} + 3e^{-\ell}$$

for every $1 \leqslant \ell \leqslant 2^{-6}|v|$.

Now we provide a few lemmas we need to prove Lemma 4.3.1, which is due to Halász [23]. Let us fix a prime p, and an integer $n \in \mathbb{N}$; the first step is the following bound on $\rho(v)$. Recall that $\|\cdot\|$ denotes the distance to the nearest integer.

Lemma 4.3.2. *For every $v \in \mathbb{Z}_p^n$,*

$$\rho(v) \leqslant \frac{1}{p} \cdot \sum_{k \in \mathbb{Z}_p} \exp\left(- \sum_{j=1}^{n} \left\| \frac{k \cdot v_j}{p} \right\|^2 \right). \tag{4.7}$$

Proof. We need to bound, for each $a \in \mathbb{Z}_p$, the probability that $u \cdot v = a$, where u is chosen uniformly at random from $\{-1, 1\}^n$. The first step is to rewrite this probability as

$$\mathbb{P}\big(u \cdot v = a\big) = \frac{1}{p} \cdot \sum_{k \in \mathbb{Z}_p} \mathbb{E}\left[\exp\left(\frac{2\pi i \cdot (u \cdot v - a)k}{p} \right) \right],$$

using the fact that $\sum_{k \in \mathbb{Z}_p} \exp\big(2\pi i \cdot xk/p\big) = 0$ for every $x \in \mathbb{Z}_p \setminus \{0\}$. Now, noting that

$$\mathbb{E}\left[\exp\left(\frac{2\pi i \cdot u_j v_j k}{p} \right) \right] = \frac{1}{2}\Big(e^{2\pi i k v_j/p} + e^{-2\pi i k v_j/p} \Big) = \cos\left(\frac{2\pi k \cdot v_j}{p} \right)$$

for each $k \in \mathbb{Z}_p$ and $j \in [n]$, and recalling that the u_j are independent, it follows that

$$\begin{aligned} \mathbb{P}\big(u \cdot v = a\big) \;&=\; \frac{1}{p} \cdot \sum_{k \in \mathbb{Z}_p} \exp\left(- \frac{2\pi i \cdot a \cdot k}{p} \right) \prod_{j=1}^{n} \cos\left(\frac{2\pi k \cdot v_j}{p} \right) \\ &\leqslant \frac{1}{p} \cdot \sum_{k \in \mathbb{Z}_p} \prod_{j=1}^{n} \left| \cos\left(\frac{\pi k \cdot v_j}{p} \right) \right|, \end{aligned}$$

where we used the fact that $\{2k : k \in \mathbb{Z}_p\} = \mathbb{Z}_p$.

Finally, using the inequality $\big|\cos(\pi x/p)\big| \leqslant \exp\big(-\|x/p\|^2\big)$, we obtain

$$\rho(v) = \max_{a \in \mathbb{Z}_p} \mathbb{P}\big(u \cdot v = a\big) \leqslant \frac{1}{p} \cdot \sum_{k \in \mathbb{Z}_p} \exp\left(- \sum_{j=1}^{n} \left\| \frac{k \cdot v_j}{p} \right\|^2 \right),$$

as claimed. □

We next rewrite the right-hand side of (4.7) in terms of the level sets $T_t(v)$.

Lemma 4.3.3. *For every $v \in \mathbb{Z}_p^n \setminus \{0\}$ and $k \geqslant 1$,*

$$\rho(v) \leqslant \frac{1}{p} + \frac{e}{p} \sum_{t=1}^{k} e^{-t} \big|T_t(v)\big| + 3e^{-k}. \tag{4.8}$$

Proof. By Lemma 4.3.2 and the definition (4.6) of $T_t(v)$, we have

$$\rho(v) \leqslant \frac{1}{p}\Big(|T_0(v)| + \sum_{t=1}^{n} \big|T_t(v) \setminus T_{t-1}(v)\big| \cdot e^{-(t-1)}\Big).$$

Now observe that $T_0(v) = \{0\}$, since $v \neq 0$, and therefore

$$\rho(v) \;\leqslant\; \frac{1}{p} + \frac{e}{p}\sum_{t=1}^{k} e^{-t}\big|T_t(v)\big| + 3e^{-k}$$

for any $k \geqslant 1$, as required. □

In order to deduce Lemma 4.3.1 from Lemma 4.3.3, we will need the following simple lemma.

Lemma 4.3.4. *For any $m \in \mathbb{N}$ and $t \geqslant 0$, and any vector $v \in \mathbb{Z}_p^n$,*

$$m \cdot T_t(v) \subset T_{m^2 t}(v)$$

where $m \cdot T$ denotes the m-fold sumset of a set T.

Proof. For each $a_1, \dots a_m \in T_t(v)$, we have

$$\sum_{k=1}^{n}\Big\|\sum_{j=1}^{m}\frac{a_j \cdot v_k}{p}\Big\|^2 \leqslant \sum_{k=1}^{n}\Big(\sum_{j=1}^{m}\Big\|\frac{a_j \cdot v_k}{p}\Big\|\Big)^2 \leqslant m\sum_{j=1}^{m}\sum_{k=1}^{n}\Big\|\frac{a_j \cdot v_k}{p}\Big\|^2 \leqslant m^2 t$$

by the triangle inequality for $\|\cdot\|$, convexity, and the definition of $T_t(v)$. □

Finally, we will need the Cauchy–Davenport theorem.

Lemma 4.3.5. *Let $m \in \mathbb{N}$, let p be a prime, and let $A \subset \mathbb{Z}_p$ be such that $m \cdot A \neq \mathbb{Z}_p$. Then*

$$|m \cdot A| \geqslant m|A| - m + 1.$$

We are now ready to prove Halász's Anticoncentration Lemma.

Proof of Lemma 4.3.1. Let $v \in \mathbb{Z}_p^n \setminus \{0\}$, and let $1 \leqslant t \leqslant \ell \leqslant 2^{-6}|v|$. We claim first that $|T_\ell(v)| < p$. To see this, let a be a uniformly-chosen random element of $\mathbb{Z}_p$, and note that for each fixed $k \in \mathbb{Z}_p \setminus \{0\}$ we have $\mathbb{P}\big(\|a \cdot k/p\| \geqslant 1/4\big) > 1/4$, and therefore

$$\mathbb{E}\Big[\sum_{i=1}^{n}\Big\|\frac{a \cdot v_i}{p}\Big\|^2\Big] > \frac{|v|}{2^6}. \tag{4.9}$$

Since $\ell \leqslant 2^{-6}|v|$, it follows that there exists $k \in \mathbb{Z}_p$ with $k \notin T_\ell(v)$, as claimed.

Now, by Lemma 4.3.4, applied with $m := \lfloor\sqrt{\ell/t}\rfloor \geqslant \sqrt{\ell}/(2\sqrt{t})$, and by the definitions of $T_t(v)$ and $|v|$, we have

$$|m \cdot T_t(v)| \leqslant |T_{m^2 t}(v)| \leqslant |T_\ell(v)|.$$

By the Cauchy–Davenport theorem, it follows that $|m \cdot T_t(v)| \geqslant m\big(|T_t(v)| - 1\big)$, and hence

$$|T_t(v)| \leqslant 1 + \frac{|T_\ell(v)|}{m} \leqslant 1 + 2\sqrt{\frac{t}{\ell}} \cdot |T_\ell(v)|.$$

Combining this with Lemma 4.3.3, we obtain

$$\rho(v) \leqslant \frac{3}{p} + \frac{2e}{p} \cdot \frac{|T_\ell(v)|}{\sqrt{\ell}} \sum_{t=1}^{\ell} \sqrt{t} e^{-t} + 3e^{-\ell} \leqslant \frac{3}{p} + \frac{6|T_\ell(v)|}{p\sqrt{\ell}} + 3e^{-\ell},$$

as claimed. □

Let us now motivate the way we choose our family of containers, see (4.11), below. The basic intuition, first suggested by Tao and Vu [49, 50], is that if $\rho(v)$ is large, then v should have some arithmetic structure. We think of the elements of the level sets $T_t(v)$ as 'frequencies' that correlate with the entries of v, and thus encode this arithmetic structure. Following the strategy of Tao and Vu [49] and Nguyen and Vu [32], we would therefore like to define the container of each 'structured' vector using its level sets.

The problem is that we would like a relatively small family of containers, whereas the number of level sets could potentially be very large. The solution is very simple: we consider a random subset U of the coordinates of v. We will show that if $|U| \geqslant 2^{12} \log p$, then v_U still correlates with the frequencies of the level sets of v, and we will choose the container of v to be (roughly speaking) the elements of $\mathbb{Z}_p$ that correlate with these frequencies. We then choose $|U|$ as small as possible (subject to the above argument working), which implies that there are few choices for the vector v_U, and hence few containers.

4.4 Proof of the inverse Littlewood–Offord theorem

In this section we will prove Theorem 4.1.2. Let $n \in \mathbb{N}$ and a prime p be fixed[3] throughout the section, and assume that $n \geqslant 2^{18} \log p$ (since otherwise the statement is vacuous). For each $m \in \mathbb{N}$ and $w \in \mathbb{Z}_p^m$, define (cf. [49, Section 7] and [32, Section 5]) the set of 'frequencies' of w to be

$$F(w) := \left\{ k \in \mathbb{Z}_p \,:\, \sum_{i=1}^{m} \left\| \frac{k \cdot w_i}{p} \right\|^2 \leqslant \log p \right\},$$

and note (recalling (4.6)) that $F(w) = T_{\log p}(w)$. Now, for each $S \subset \mathbb{Z}_p$, define

$$C(S) := \left\{ a \in \mathbb{Z}_p \,:\, \sum_{k \in S} \left\| \frac{a \cdot k}{p} \right\|^2 \leqslant \frac{|S|}{2^5} \right\}. \tag{4.10}$$

[3] We may assume that $p \geqslant 2^{10}$, since otherwise the conclusion of Theorem 4.1.2 holds trivially with $\mathcal{C}$ equal to the collection of all subsets of $\mathbb{Z}_p$.

Now set $m := \lfloor 2^{12} \log p \rfloor$, and define

$$\mathcal{C} := \big\{ C\big(F(w)\big) : w \in \mathbb{Z}_p^m \big\}, \tag{4.11}$$

and observe that $|\mathcal{C}| \leqslant p^m$, as required. We will show that $\mathcal{C}$ has the desired properties.

The following simple lemma motivates our choice of containers (cf. [32, Section 5]).

Lemma 4.4.1. *Let $v \in \mathbb{Z}_p^n$, and let $t \leqslant 2^{-7}n$. If $S \subset T_t(v)$, then*

$$\big|\big\{ i \in [n] : v_i \notin C(S) \big\}\big| \leqslant \frac{n}{4}.$$

Proof. Let $R = \big\{ i \in [n] : v_i \notin C(S) \big\}$, and observe that, by (4.6) and (4.10),

$$\frac{|R||S|}{2^5} \leqslant \sum_{i \in R} \sum_{k \in S} \left\| \frac{k \cdot v_i}{p} \right\|^2 \leqslant \sum_{k \in S} \sum_{i=1}^{n} \left\| \frac{k \cdot v_i}{p} \right\|^2 \leqslant \frac{n|S|}{2^7},$$

so $|R| \leqslant n/4$, as required. □

Later in the proof, we will define $B(v) := C\big(F(v_U)\big)$ for some set $U \subset [m]$ with $|U| \leqslant \lambda$ such that $F(v_U) \subset T_t(v)$ for $t = 2^{-7}n$ (see Lemma 4.4.6, below). We next turn to bounding the size of our containers; the following lemma (cf. [32, Section 5]) provides a first step.

Lemma 4.4.2. *For any set $S \subset \mathbb{Z}_p$, we have*

$$|C(S)| \leqslant \frac{4p}{|S|}. \tag{4.12}$$

Proof. We will instead bound the size of the larger set

$$C'(S) := \left\{ a \in \mathbb{Z}_p \, : \, \sum_{k \in S} \cos\left(\frac{2\pi a k}{p}\right) \geqslant \frac{|S|}{2} \right\}.$$

Indeed, observe that $C(S) \subset C'(S)$, since we have $1 - 2^4\|x\|^2 \leqslant \cos(2\pi x)$ for every $x \in \mathbb{R}$.

Now, let a be a uniformly-chosen random element of $\mathbb{Z}_p$, and observe that, by Markov's inequality,

$$\begin{aligned} \mathbb{P}\big(a \in C'(S)\big) \; &= \; \mathbb{P}\left(\left(\sum_{k \in S} \cos\left(\frac{2\pi a k}{p}\right) \right)^2 \geqslant \frac{|S|^2}{4} \right) \\ &\leqslant \; \frac{4}{|S|^2} \cdot \frac{1}{p} \sum_{a \in \mathbb{Z}_p} \left(\sum_{k \in S} \cos\left(\frac{2\pi a k}{p}\right) \right)^2, \end{aligned}$$

Now, since $2\cos(x) = e^{ix} + e^{-ix}$, we have

$$4\sum_{a\in\mathbb{Z}_p}\left(\sum_{k\in S}\cos\left(\frac{2\pi ak}{p}\right)\right)^2 = \sum_{k_1\in\pm S}\sum_{k_2\in\pm S}\sum_{a\in\mathbb{Z}_p}\exp\left(\frac{2\pi ia(k_1+k_2)}{p}\right) \leqslant 4p|S|,$$

where $\pm S$ is the multi-set obtained by taking the union of S and $-S$, counting elements in both twice. For the second step, simply note that the roots of unity sum to zero, so the only terms that contribute are those with $k_1 + k_2 = 0$. It follows that

$$\frac{4}{|S|^2}\cdot\frac{1}{p}\sum_{a\in\mathbb{Z}_p}\left(\sum_{k\in S}\cos\left(\frac{2\pi ak}{p}\right)\right)^2 \leqslant \frac{4}{|S|},$$

and hence $|C(S)| \leqslant |C'(S)| \leqslant 4p/|S|$, as claimed. □

We will use Halász's Anticoncentration Lemma (Lemma 4.3.1) to bound the right-hand side of (4.12) in terms of $\rho(v_Y)$ (for some set Y that will be chosen in Lemma 4.4.5, below). The following lemma is a straightforward application of Lemma 4.3.1.

Lemma 4.4.3. *Let $v \in \mathbb{Z}_p^n$ with $\rho(v) \geqslant 4/p$ and $|v| \geqslant 2^{18}\log p$, and let $Y \subset [n]$ be such that $|v_Y| \geqslant |v|/4$. Then*

$$\rho(v_Y) \leqslant \frac{2^{13}|T_\ell(v_Y)|}{p\sqrt{|v|}},$$

where $\ell := 2^{-16}|v|$.

In the proof of Lemma 4.4.3, and also later in the section, we will need the following simple observation (see Lemma 4.6.10 or [14, Lemma 2.8]).

Observation 4.4.4 (Lemma 2.8 of [14])**.** *$\rho(v_Y) \geqslant \rho(v)$ for every $v \in \mathbb{Z}_p^n$ and every $Y \subset [n]$.*

Proof of Lemma 4.4.3. Applying Lemma 4.3.1 to v_Y, with $\ell = 2^{-16}|v| \leqslant 2^{-14}|v_Y| < |v_Y|$, gives

$$\rho(v_Y) \leqslant \frac{3}{p} + \frac{6|T_\ell(v_Y)|}{p\sqrt{\ell}} + 3e^{-\ell}.$$

Now, by Observation 4.4.4 and our assumption on $\rho(v)$, we have $\rho(v_Y) \geqslant \rho(v) \geqslant 4/p$. Since $\ell \geqslant 4\log p$, it follows that

$$\rho(v_Y) \leqslant \frac{2^5|T_\ell(v_Y)|}{p\sqrt{\ell}} = \frac{2^{13}|T_\ell(v_Y)|}{p\sqrt{|v|}},$$

as claimed. □

To complete the proof, it will now suffice to choose sets $Y \subset [n]$, with $n/4 \leqslant |Y| \leqslant n/2$, and $U \subset [n]$, with $|U| \leqslant \lambda$, such that

$$F(v_U) \subset T_t(v), \qquad |v_Y| \geqslant \frac{|v|}{4} \qquad \text{and} \qquad |T_\ell(v_Y)| \leqslant 2 \cdot |F(v_U)|, \tag{4.13}$$

where $\ell = 2^{-16}|v|$ and $t = 2^{-7}n$. Indeed, for any such sets we have, by Lemmas 4.4.2 and 4.4.3,

$$\big|C\big(F(v_U)\big)\big| \leqslant \frac{4p}{|F(v_U)|} \leqslant \frac{2^{15}}{\rho(v_Y)\sqrt{|v|}} \cdot \frac{|T_\ell(v_Y)|}{|F(v_U)|} \leqslant \frac{2^{16}}{\rho(v_Y)\sqrt{|v|}},$$

and, by Lemma 4.4.1, we have

$$\big|\big\{i \in [n] : v_i \notin C\big(F(v_U)\big)\big\}\big| \leqslant \frac{n}{4}.$$

Thus, setting $B(v) := C\big(F(v_U)\big)$, we obtain a set in $\mathcal{C}$ for which the properties (4.3) hold.

We will choose the sets Y and U in the next two lemmas. In each case we simply choose a random set of the correct density. We will say that R is a *q-random* subset of a set S if each element of S is included in R independently at random with probability q.

Lemma 4.4.5. *Let $v \in \mathbb{Z}_p^n$ with $|v| \geqslant 2^{18} \log p$. There exists $Y \subset [n]$, with $n/4 \leqslant |Y| \leqslant n/2$, such that*

$$|v_Y| \geqslant \frac{|v|}{4} \qquad \textit{and} \qquad T_\ell(v_Y) \subset T_{8\ell}(v),$$

where $\ell = 2^{-16}|v|$.

Proof. Let Y be a $(3/8)$-random subset of $[n]$; we will prove that with positive probability Y has all of the required properties. Since $n \geqslant |v| \geqslant 2^{18} \log p \geqslant 2^{18}$, the properties

$$\frac{n}{4} \leqslant |Y| \leqslant \frac{n}{2} \qquad \text{and} \qquad |v_Y| \geqslant \frac{|v|}{4}$$

each hold with probability at least $3/4$, by Chernoff's inequality. To bound the probability that $T_\ell(v_Y) \setminus T_{8\ell}(v)$ is non-empty, define a random variable

$$W(k) := \sum_{i \in Y} \left\| \frac{k \cdot v_i}{p} \right\|^2$$

for each $k \in \mathbb{Z}_p$, and observe that, by (4.6),

$$k \in T_\ell(v_Y) \;\Leftrightarrow\; W(k) \leqslant \ell \qquad \text{and} \qquad k \notin T_{8\ell}(v) \;\Rightarrow\; \mathbb{E}[W(k)] \geqslant 3\ell.$$

Moreover, by Chernoff's inequality,[4]

$$\mathbb{P}\big(k \in T_\ell(v_Y)\big) = \mathbb{P}\big(W(k) \leqslant \ell\big) \leqslant e^{-\ell/2} \leqslant \frac{1}{p^2}$$

[4]Here we use the following variant of the standard Chernoff inequality: if $X_1, \dots, X_N$ are iid Bernoulli random variables, and $t_1, \dots, t_N \in [0,1]$, then $\mathbb{P}\big(\sum_{i=1}^N t_i X_i \leqslant s\big) \leqslant \exp\big(-\mathbb{E}[X]/2 + s\big)$.

for every $k \notin T_{8\ell}(v)$, since $\ell \geqslant 4 \log p$. It follows that

$$\mathbb{E}\big[|T_\ell(v_Y) \setminus T_{8\ell}(v)|\big] \leqslant \frac{1}{p},$$

and hence $T_\ell(v_Y) \subset T_{8\ell}(v)$ with probability at least $3/4$, as required. □

Finally, we need to show that a suitable set U exists.

Lemma 4.4.6. *Let $v \in \mathbb{Z}_p^n$. There exists $U \subset [n]$, with $|U| \leqslant m$, such that*

$$|T_{8\ell}(v)| \leqslant 2 \cdot |F(v_U)| \qquad \text{and} \qquad F(v_U) \subset T_t(v),$$

where $\ell = 2^{-16}|v|$ and $t = 2^{-7}n$.

Proof. Let U be a $(m/2n)$-random subset of $[n]$. We will prove that the claimed properties hold simultaneously with positive probability. Note first that $|U| \leqslant m$ with probability at least $3/4$, by Chernoff's inequality, since $m = \lfloor 2^{12} \log p \rfloor \geqslant 2^{12}$.

Next, we show that $|T_{8\ell}(v) \setminus F(v_U)| \leqslant |T_{8\ell}(v)|/2$ with probability at least $1/2$. Observe first that, for every $k \in \mathbb{Z}_p$,

$$\mathbb{P}\big(k \notin F(v_U)\big) = \mathbb{P}\left(\sum_{i \in U} \left\| \frac{k \cdot v_i}{p} \right\|^2 > \log p\right) \leqslant \frac{1}{\log p} \cdot \mathbb{E}\left[\sum_{i \in U} \left\| \frac{k \cdot v_i}{p} \right\|^2\right],$$

by Markov's inequality. Now, if $k \in T_{8\ell}(v)$, then

$$\frac{1}{\log p} \cdot \mathbb{E}\left[\sum_{i \in U} \left\| \frac{k \cdot v_i}{p} \right\|^2\right] = \frac{m}{2n \log p} \sum_{i=1}^{n} \left\| \frac{k \cdot v_i}{p} \right\|^2 \leqslant \frac{8m\ell}{2n \log p} \leqslant \frac{1}{4},$$

since $m \leqslant 2^{12} \log p$ and $\ell = 2^{-16}|v| \leqslant 2^{-16}n$. It follows that

$$\mathbb{P}\left(|T_{8\ell}(v) \setminus F(v_U)| \geqslant \frac{|T_{8\ell}(v)|}{2}\right) \leqslant \frac{2}{|T_{8\ell}(v)|} \cdot \mathbb{E}\big[|T_{8\ell}(v) \setminus F(v_U)|\big] \leqslant \frac{1}{2},$$

by Markov's inequality, as claimed.

Finally, to bound the probability that $F(v_U) \setminus T_t(v)$ is non-empty, we repeat the argument used in the proof of Lemma 4.4.5. To be precise, we define a random variable

$$W(k) := \sum_{i \in U} \left\| \frac{k \cdot v_i}{p} \right\|^2$$

for each $k \in \mathbb{Z}_p$, and observe that, by (4.6),

$$k \in F(v_U) \;\Leftrightarrow\; W(k) \leqslant \log p \qquad \text{and} \qquad k \notin T_t(v) \;\Rightarrow\; \mathbb{E}[W(k)] \geqslant 2^{-8}m.$$

Recalling that $m = \lfloor 2^{12} \log p \rfloor$, it follows by Chernoff's inequality that

$$\mathbb{P}\big(k \in F(v_U)\big) = \mathbb{P}\big(W(k) \leqslant \log p\big) \leqslant \frac{1}{p^2}$$

for every $k \notin T_t(v)$, and hence

$$\mathbb{P}\big(F(v_U) \not\subset T_t(v)\big) \leqslant \mathbb{E}\big[|F(v_U) \setminus T_t(v)|\big] \leqslant \frac{1}{p}.$$

It follows that, with positive probability, the random set U satisfies

$$|U| \leqslant m, \qquad |T_{8\ell}(v)| \leqslant 2 \cdot |F(v_U)| \qquad \text{and} \qquad F(v_U) \subset T_t(v),$$

as required. □

As observed above, it is now straightforward to complete the proof of Theorem 4.1.2.

Proof of Theorem 4.1.2. Let $\mathcal{C}$ be as defined in (4.11), and note that

$$|\mathcal{C}| \leqslant p^m \leqslant \exp\big(2^{12}(\log p)^2\big).$$

For each $v \in \mathbb{Z}_p^n$ with $\rho(v) \geqslant 4/p$ and $|v| \geqslant 2^{18} \log p$, let Y and U be the sets given by Lemmas 4.4.5 and 4.4.6 respectively, and define $B(v) := C\big(F(v_U)\big)$. Now, we have $n/4 \leqslant |Y| \leqslant n/2$, by Lemma 4.4.5, and

$$\big|\big\{i \in [n] : v_i \notin B(v)\big\}\big| \leqslant \frac{n}{4},$$

by Lemma 4.4.1, since $F(v_U) \subset T_t(v)$, where $t = 2^{-7}n$, by Lemma 4.4.6. Finally, we have

$$|B(v)| \leqslant \frac{4p}{|F(v_U)|} \leqslant \frac{2^{15}}{\rho(v_Y)\sqrt{|v|}} \cdot \frac{|T_\ell(v_Y)|}{|F(v_U)|} \leqslant \frac{2^{16}}{\rho(v_Y)\sqrt{|v|}},$$

by Lemmas 4.4.2–4.4.6, since $|T_\ell(v_Y)| \leqslant |T_{8\ell}(v)| \leqslant 2 \cdot |F(v_U)|$. This completes the proof of the inverse Littlewood–Offord theorem. □

4.5 Applying the inverse Littlewood–Offord theorem

In this section we will use our inverse Littlewood–Offord theorem to prove Lemma 4.2.2. Let us fix $n \in \mathbb{N}$ and a prime $2 < p \leqslant \exp\big(2^{-10}\sqrt{n}\big)$ throughout the section. Recall that $\beta \geqslant 4/p$, that

$$q_n(\beta) = \max_{w \in \mathbb{Z}_p^n} \mathbb{P}\big(\exists\, v \in \mathbb{Z}_p^n \setminus \{0\} \,:\, M_n \cdot v = w \text{ and } \rho(v) \geqslant \beta\big),$$

and that our aim is to prove that $q_n(\beta) \leqslant 2^{-n/4}$. We shall do so by using Theorem 4.1.2 to partition the vectors $v \in \mathbb{Z}_p^{n-1} \setminus \{0\}$ into 'containers', and then applying a simple first moment argument inside each container. The simplest container consists of those vectors with small

support, so let us deal with those first. For each $w \in \mathbb{Z}_p^n$, define

$$Q(w) := \big|\{v \in \mathbb{Z}_p^n \setminus \{0\} \,:\, M_n \cdot v = w \text{ and } |v| < 2^8\sqrt{n}\}\big|.$$

Our first lemma bounds the expected size of $Q(w)$.

Lemma 4.5.1. *For every $w \in \mathbb{Z}_p^n$,*

$$\mathbb{E}\big[Q(w)\big] \leqslant 2^{-n/2}.$$

Proof. Fix $w \in \mathbb{Z}_p^n$; the lemma is an easy consequence of the following claim.

Claim: If $v \in \mathbb{Z}_p^n \setminus \{0\}$, then $\mathbb{P}\big(M_n \cdot v = w\big) \leqslant 2^{-n}$.

Proof. Choose $k \in [n]$ such that $v_k \neq 0$, and reveal the entire matrix M_n except for the kth row and the kth column. Observe that if $M_n \cdot v = w$, then

$$m_{ik}v_k = w_i - \sum_{j \neq k} m_{ij}v_j \tag{4.14}$$

for each $i \in [n]$, where m_{ij} are the entries of M_n. Now, for any choice of the entries m_{ij} with $j \neq k$, the event (4.14) has probability at most $1/2$, and these events are independent for different values of $i \neq k$. Finally, having revealed the entire matrix except for m_{kk}, the event (4.14) for $i = k$ has probability at most $1/2$, so $\mathbb{P}\big(M_n \cdot v = w\big) \leqslant 2^{-n}$, as claimed. □

Now, since there are at most $\binom{n}{k}p^k$ vectors $v \in \mathbb{Z}_p^n \setminus \{0\}$ with $|v| < k$, and recalling that $p \leqslant \exp\big(2^{-10}\sqrt{n}\big)$, the claim implies that

$$\mathbb{E}\big[Q(w)\big] \leqslant \binom{n}{2^8\sqrt{n}} p^{2^8\sqrt{n}} \cdot 2^{-n} \leqslant 2^{-n/2}$$

as required. □

From now on, we will therefore restrict our attention to the vectors with large support:

$$\mathcal{V} := \big\{v \in \mathbb{Z}_p^n \,:\, \rho(v) \geqslant \beta,\, |v| \geqslant 2^8\sqrt{n}\big\}.$$

To deal with these vectors, we will define a function

$$f \colon \mathcal{V} \to \mathcal{X} := \Big\{\big(X_i, Y_i, B_i\big)_{i=1}^{\infty} : X_i, Y_i \subset [n] \text{ and } B_i \subset \mathbb{Z}_p \text{ for each } i \in \mathbb{N}\Big\},$$

using Theorem 4.1.2. More precisely, we will define f using the following algorithm, which takes as its input a vector $v \in \mathcal{V}$, and outputs an element of $\mathcal{X}$.

Algorithm 4.5.2. *Let $v \in \mathcal{V}$. At the kth step, if the process has not yet ended, we will have constructed a sequence $(X_i, Y_i, B_i)_{i=1}^{k-1}$ with $X_i, Y_i \subset [n]$ and $B_i \subset \mathbb{Z}_p$ for each $i \in [k-1]$. In this case, set*

$$Z_k := [n] \setminus \bigcup_{i=1}^{k-1} X_i,$$

and do the following:

1. *If $|v_{Z_k}| \geqslant 2^8\sqrt{n}$ then we apply Theorem 4.1.2, and set $Y_k := Y(v_{Z_k})$, $B_k := B(v_{Z_k})$, and*

$$X_k := \big\{i \in Z_k \setminus Y_k \,:\, v_i \in B_k\big\}. \tag{4.15}$$

   Set $k \to k+1$ and repeat the process.

2. *If $|v_{Z_k}| < 2^8\sqrt{n}$, then we set $k^* = k^*(v) := k-1$ and*

$$X_j = Y_j = B_j = \emptyset$$

   for every $j \geqslant k$. The process terminates, and we set $f(v) := \big(X_i, Y_i, B_i\big)_{i=1}^{\infty}$.

Define $\mathcal{U} := \big\{f(v) : v \in \mathcal{V}\big\}$. Theorem 4.1.2 implies the following upper bound on $|\mathcal{U}|$.

Lemma 4.5.3.

$$|\mathcal{U}| \leqslant \exp\big(2^{-6} n \log n\big).$$

Proof. We claim first that, for each $k \in \mathbb{N}$, either $|v_{Z_k}| < 2^8\sqrt{n}$, or

$$|Z_k| \leqslant \left(\frac{3}{4}\right)^{k-1} n. \tag{4.16}$$

Indeed, by Observation 4.4.4 we have $\rho(v_{Z_k}) \geqslant \rho(v) \geqslant \beta \geqslant 4/p$ for every $v \in \mathcal{V}$, and therefore, if $|v_{Z_k}| \geqslant 2^8\sqrt{n} \geqslant 2^{18}\log p$, it follows from Theorem 4.1.2 that $|Y_k| \leqslant |Z_k|/2$ and

$$\big|Z_k \setminus \big(X_k \cup Y_k\big)\big| \leqslant \big|\big\{i \in Z_k : v_i \notin B_k\big\}\big| \leqslant \frac{|Z_k|}{4}. \tag{4.17}$$

Hence $|X_k| \geqslant |Z_k|/4$, and (4.16) follows. In particular, this implies that $k^*(v) \leqslant 2\log n$.

Now, given $(X_i, Y_i, B_i)_{i=1}^{k-1}$, there are at most $2^{|Z_k|}$ choices for X_k and Y_k (since they are subsets of Z_k), and by (4.2) there are at most

$$\exp\Big(2^{12}\big(\log p\big)^2\Big) \leqslant \exp\big(2^{-8} n\big)$$

choices for B_k. It follows that the total number of choices for $f(v)$ is at most

$$\exp\left(2^{-7} n \log n + \sum_{k=1}^{\infty} \left(\frac{3}{4}\right)^{k-1} n\right) \leqslant \exp\big(2^{-6} n \log n\big),$$

as required. □

We will bound, for each sequence $S \in \mathcal{U}$, the probability that some vector $v \in \mathcal{V}$ with $f(v) = S$ satisfies $M_n \cdot v = w$, and then sum over $S \in \mathcal{U}$. To do so, for each $S \in \mathcal{U}$ and $w \in \mathbb{Z}_p^n$, let us define a random variable

$$Q(S,w) := \big|\big\{v \in \mathcal{V} \,:\, f(v) = S \text{ and } M_n \cdot v = w\big\}\big|.$$

The next lemma bounds the expected size of $Q(S,w)$.

Lemma 4.5.4. *If $S = \big(X_i, Y_i, B_i\big)_{i=1}^{\infty} \in \mathcal{U}$ and $w \in \mathbb{Z}_p^n$, then*

$$\mathbb{E}\big[Q(S,w)\big] \leqslant \left(\frac{2^{56}}{n}\right)^{n/16}. \tag{4.18}$$

Proof. If $f(v) = S$, then we have $v_j \in B_i$ for every $j \in X_i$, and $|v_{Z_{k^*}}| < 2^8\sqrt{n}$. There are therefore at most

$$\binom{n}{2^8\sqrt{n}} \cdot p^{2^8\sqrt{n}} \cdot \prod_{i=1}^{k^*} |B_i|^{|X_i|}$$

vectors $v \in \mathcal{V}$ with $f(v) = S$. We claim that, for each such vector v,

$$\mathbb{P}\big(M_n \cdot v = w\big) \leqslant \prod_{i=1}^{k^*} \max_{w_i \in \mathbb{Z}_p^{|X_i|}} \mathbb{P}\big(M_{X_i \times Y_i} \cdot v_{Y_i} = w_i\big) = \prod_{i=1}^{k^*} \rho(v_{Y_i})^{|X_i|}. \tag{4.19}$$

To prove (4.19), recall from (4.15) that

$$X_i \cap Y_i = \emptyset \qquad \text{and} \qquad X_i \cap X_j = Y_i \cap X_j = \emptyset$$

for every $i \in [k^*]$ and every $1 \leqslant j < i$, since $X_i, Y_i \subset Z_i$. It follows that

$$\mathbb{P}\bigg(M_{X_i \times [n-1]} \cdot v = w_{X_i} \;\Big|\; \bigcap_{j=1}^{i-1} M_{X_j \times [n-1]} \cdot v = w_{X_j}\bigg) \leqslant \max_{w_i \in \mathbb{Z}_p^{|X_i|}} \mathbb{P}\big(M_{X_i \times Y_i} \cdot v_{Y_i} = w_i\big)$$

for every $i \in [k^*]$, and moreover the entries of $M_{X_i \times Y_i}$ are all independent. This proves (4.19), and summing over $v \in \mathcal{V}$ with $f(v) = S$ gives

$$\mathbb{E}\big[Q(S,w)\big] \leqslant \binom{n}{2^8\sqrt{n}} \cdot p^{2^8\sqrt{n}} \cdot \prod_{i=1}^{k^*} \Big(|B_i| \cdot \rho(v_{Y_i})\Big)^{|X_i|}.$$

To deduce (4.18), recall from Theorem 4.1.2 that

$$|B_i| \leqslant \frac{2^{16}}{\rho(v_Y)\sqrt{|v|}} \leqslant \frac{2^{12}}{\rho(v_{Y_i})n^{1/4}}$$

for each $i \in [k^*]$, since $|v_{Z_i}| \geqslant 2^8\sqrt{n}$. Since $p \leqslant \exp\left(2^{-10}\sqrt{n}\right)$, and recalling from (4.17) that we have $|X_1| \geqslant n/4$ (since $|v| \geqslant 2^8\sqrt{n}$ for every $v \in \mathcal{V}$), it follows that

$$\mathbb{E}\big[Q(S,w)\big] \leqslant \binom{n}{2^8\sqrt{n}} \cdot p^{2^8\sqrt{n}} \cdot \left(\frac{2^{12}}{n^{1/4}}\right)^{\sum_i |X_i|} \leqslant \left(\frac{2^{14}}{n^{1/4}}\right)^{n/4} = \left(\frac{2^{56}}{n}\right)^{n/16},$$

as required. □

Completing the proof of Lemma 4.2.2, and hence of Theorem 4.1.1, is now straightforward.

Proof of Lemma 4.2.2. By Lemma 4.5.1, for each $w \in \mathbb{Z}_p^n$ the probability that there exists $v \in \mathbb{Z}_p^n \setminus \{0\}$ such that $|v| < 2^8\sqrt{n}$ and $M_n \cdot v = w$ is at most $2^{-n/2}$, and hence

$$q_n(\beta) \leqslant 2^{-n/2} + \sum_{S \in \mathcal{U}} \max_{w \in \mathbb{Z}_p^n} \mathbb{P}\Big(\exists\, v \in \mathcal{V} \,:\, f(v) = S \text{ and } M_n \cdot v = w\Big).$$

Now, by Lemma 4.5.4, we have

$$\mathbb{P}\Big(\exists\, v \in \mathcal{V} \,:\, f(v) = S \text{ and } M_n \cdot v = w\Big) \leqslant \left(\frac{2^{56}}{n}\right)^{n/16}$$

for every $S \in \mathcal{U}$ and $w \in \mathbb{Z}_p^n$, and hence, by Lemma 4.5.3,

$$q_n(\beta) \leqslant 2^{-n/2} + \exp\left(2^{-6} n \log n\right)\left(\frac{2^{56}}{n}\right)^{n/16} \leqslant 2^{-n/4}$$

if n is sufficiently large. This completes the proof of the lemma. □

As observed in Section 4.2, Lemmas 4.2.1 and 4.2.2 together imply Theorem 4.1.1.

4.6 Proof of Lemma 4.2.1

In this section we finish the proof by proving Lemma 4.2.1, which allowed us to reduce the problem of bounding the probability that $\det(M_n) = 0$ to the problem of bounding $q_n(\beta)$. The proof given below is essentially contained in the paper of Ferber and Jain [14], and several of the key lemmas appeared in the papers of Costello, Tao and Vu [9] and Nguyen [31]. We begin by giving an overview of the proof.

4.6.1 Overview of the proof of Lemma 4.2.1

It will be convenient in this section to work over $\mathbb{F}_p$; in particular, we will consider the entries of M_n as elements of $\mathbb{F}_p$, noting that doing so can only increase the probability that M_n is

singular. Observe also that

$$q_n(\beta) = \max_{w \in \mathbb{F}_p^n} \mathbb{P}\Big(\exists\, v \in \mathbb{F}_p^n \setminus \{0\} : M_n \cdot v = w \text{ and } \rho(v) \geqslant \beta\Big), \tag{4.20}$$

where now M_n is a matrix over $\mathbb{F}_p$.

Let us write $\mathrm{rk}(M)$ for the rank of a matrix M over $\mathbb{F}_p$, and M_{n-1} for the random symmetric matrix obtained by removing the first row and column from M_n. The following lemma, which was proved by Nguyen (see [31, Section 2]), allows us to restrict our attention to matrices M_n such that $\mathrm{rk}(M_n) = n-1$ and $\mathrm{rk}(M_{n-1}) \in \{n-2, n-1\}$.

Lemma 4.6.1. *For every $n \in \mathbb{N}$ and prime $p > 2$,*

$$\mathbb{P}\big(\det(M_n) = 0\big) \leqslant 4n \sum_{m=n}^{2n-2} \mathbb{P}\Big(\big\{\mathrm{rk}(M_m) = m-1\big\} \cap \big\{\mathrm{rk}(M_{m-1}) \in \{m-2, m-1\}\big\}\Big).$$

The proof of Lemma 4.6.1 is given in Section 4.6.2. The next two lemmas deal with the cases $\mathrm{rk}(M_{n-1}) = n-2$ and $\mathrm{rk}(M_{n-1}) = n-1$ respectively; the first is more straightforward.

Lemma 4.6.2. *For every $n \in \mathbb{N}$, prime $p > 2$, and $\beta > 0$,*

$$\mathbb{P}\Big(\big\{\mathrm{rk}(M_n) = n-1\big\} \cap \big\{\mathrm{rk}(M_{n-1}) = n-2\big\}\Big) \leqslant \beta + q_{n-1}(\beta).$$

The proof of Lemma 4.6.2, which follows that given in [14, Section 2.2], is described in Section 4.6.3. Finally, the following lemma deals with the case $\mathrm{rk}(M_{n-1}) = n-1$.

Lemma 4.6.3. *For every $n \in \mathbb{N}$, prime $p > 2$, $\beta > 0$, and integer $1 \leqslant k \leqslant n-2$, we have*

$$\mathbb{P}\Big(\mathrm{rk}(M_n) = \mathrm{rk}(M_{n-1}) = n-1\Big) \leqslant 2 \cdot \big(2^k\beta + 2^{-k}\big)^{1/4} + 3^{k+1} q_{n-1}(\beta).$$

The proof of Lemma 4.6.3, which is similar to that given in [14, Section 2.3], is provided in Section 4.6.4. Combining Lemmas 4.6.1, 4.6.2 and 4.6.3, we obtain Lemma 4.2.1.

Proof of Lemma 4.2.1. Observe first that $q_n(\beta) \geqslant 2^{-n}$ for every $\beta < 1/2$ (to see this, set $v = (1, 0, \ldots, 0)$), so the claimed bound holds trivially if $\beta > n^{-1}$ or $\beta < 2^{-n}$. We may therefore assume that $k := \lfloor \log_4(1/\beta) \rfloor \in [n-2]$, and therefore, by Lemmas 4.6.1, 4.6.2 and 4.6.3, we obtain

$$\begin{aligned}
\mathbb{P}\big(\det(A_n) = 0\big) &\leqslant 4n \sum_{m=n}^{2n-2} \Big(\beta + q_{m-1}(\beta) + 2 \cdot \big(3\beta^{1/2}\big)^{1/4} + \beta^{-1} q_{m-1}(\beta)\Big) \\
&\leqslant 16n \sum_{m=n-1}^{2n-3} \left(\beta^{1/8} + \frac{q_m(\beta)}{\beta}\right).
\end{aligned}$$

as required. □

4.6.2 The proof of Lemma 4.6.1

As noted above, Lemma 4.6.1 is a straightforward consequence of [31, Lemmas 2.1 and 2.3]. The first of these two lemmas is as follows.

Lemma 4.6.4 (Lemma 2.1 of [31]). *For any $0 \leqslant k \leqslant n-1$,*

$$\mathbb{P}\big(\mathrm{rk}(M_n) = k\big) \leqslant 2 \cdot \mathbb{P}\big(\mathrm{rk}(M_{2n-k-1}) = 2n-k-2\big). \tag{4.21}$$

To prove Lemma 4.6.4 we will need the following observation of Odlyzko [34]; since it is usually stated in $\mathbb{R}^n$, we provide the short proof.

Observation 4.6.5. *Let V be a subspace of $\mathbb{F}_p^n$ of dimension at most k. Then*

$$\big|V \cap \{-1,1\}^n\big| \leqslant 2^k.$$

Proof. Form an $n \times k$ matrix over $\mathbb{F}_p$ whose columns are a basis $\{v^{(1)}, \dots, v^{(k)}\}$ of V, and choose k linearly independent rows. We obtain an invertible matrix A, and so for each $b \in \{-1,1\}^k$, there is a unique solution in $\mathbb{F}_p^k$ to the set of equations $Ax = b$. The 2^k vectors $\sum_{i=1}^k x_i v^{(i)}$ (one for each $b \in \{-1,1\}^k$) are the only possible elements of $V \cap \{-1,1\}^n$. □

We can now prove Lemma 4.6.4.

Proof of Lemma 4.6.4. We claim that, for any $0 \leqslant k \leqslant n-1$,

$$\mathbb{P}\big(\mathrm{rk}(M_{n+1}) = k+2 \;\big|\; \mathrm{rk}(M_n) = k\big) \geqslant 1 - 2^{k-n}. \tag{4.22}$$

where we remind the reader that M_n is obtained from M_{n+1} by removing the first row and column. Let W be the subspace spanned by the rows of M_n, and note that, by Observation 4.6.5, if $\mathrm{rk}(M_n) = k$ then W intersects $\{-1,1\}^n$ in at most 2^k vectors.

Let $v \in \mathbb{F}_p^n$ be the vector formed by removing the first element from the first row of M_{n+1}. By the remarks above, it follows that $\mathbb{P}(v \notin W) \geqslant 1 - 2^{k-n}$. We claim that if $v \notin W$ then $\mathrm{rk}(M_{n+1}) = k+2$. To see this, note first that if $v \notin W$ then the rank of the final n columns of M_{n+1} is $k+1$. Now, since M_{n+1} is symmetric, the first column of M_{n+1} is the same as the first row, and if $v \notin W$ then v is not in the span of the columns of M_n. It follows that $\mathrm{rk}(M_{n+1}) = k+2$, as claimed, and (4.22) follows.

It follows immediately from (4.22) that

$$\mathbb{P}\big(\mathrm{rk}(M_{n+t}) = k+2t \;\big|\; \mathrm{rk}(M_{n+t-1}) = k+2t-2)\big) \geqslant 1 - 2^{k+t-n-1}$$

for every $k \geqslant 0$ and $1 \leqslant t \leqslant n-k$. Now, building M_{n+t} from M_n by adding one row and column at a time, it follows that

$$\mathbb{P}\big(\mathrm{rk}(M_{2n-k-1}) = 2n-k-2 \,\big|\, \mathrm{rk}(M_n) = k\big) \geqslant \prod_{i=2}^{n-k}(1-2^{-i}) \geqslant \frac{1}{2},$$

which implies (4.21). □

We can now deduce Lemma 4.6.1 using [31, Lemma 2.3], which is the following observation. Let us write $M_n^{(i)}$ for the (symmetric) matrix obtained from M_n by removing the ith row and the ith column.

Lemma 4.6.6 (Lemma 2.3 of [31])**.** *If* $\mathrm{rk}(M_n) = n-1$*, then* $\max_{i\in[n]} \mathrm{rk}(M_n^{(i)}) \geqslant n-2$.

Proof. Choose $n-1$ rows of M_n whose span has dimension $n-1$, and remove the remaining row, giving an $(n-1)\times n$ matrix of rank $n-1$. Hence, removing any column from this matrix, we obtain a matrix of rank at least $n-2$. □

Proof of Lemma 4.6.1. By Lemma 4.6.4, we have

$$\mathbb{P}\big(\det(M_n) = 0\big) = \sum_{k=1}^{n-1} \mathbb{P}\big(\mathrm{rk}(M_n) = k\big) \leqslant 2\sum_{k=1}^{n-1} \mathbb{P}\big(\mathrm{rk}(M_{2n-k-1}) = 2n-k-2\big). \tag{4.23}$$

We therefore need to bound $\mathbb{P}\big(\mathrm{rk}(M_m) = m-1\big)$ for each $n \leqslant m \leqslant 2n-2$. By Lemma 4.6.6, and by symmetry, we have

$$\begin{aligned}\mathbb{P}\big(\mathrm{rk}(M_m) = m-1\big) &\leqslant \sum_{i=1}^{m} \mathbb{P}\Big(\big\{\mathrm{rk}(M_m) = m-1\big\} \cap \big\{\mathrm{rk}(M_m^{(i)}) \geqslant m-2\big\}\Big) \\ &\leqslant m \cdot \mathbb{P}\Big(\big\{\mathrm{rk}(M_m) = m-1\big\} \cap \big\{\mathrm{rk}(M_{m-1}) \in \{m-2, m-1\}\big\}\Big).\end{aligned}$$

Combining this with (4.23) gives the statement of the lemma. □

4.6.3 The case $\mathrm{rk}(M_{n-1}) = n-2$

In this subsection we will prove Lemma 4.6.2, following the presentation in [14, Section 2.2]. Let us write $\mathrm{adj}(M)$ for the *adjugate* of a matrix M over $\mathbb{F}_p$. We will need the following lemma of Nguyen [31], see [14, Lemma 2.5].

Lemma 4.6.7. *If* $\mathrm{rk}(M_{n-1}) = n-2$*, then there exists a non-trivial column* $a \in \mathbb{F}_p^{n-1}$ *of* $\mathrm{adj}(M_{n-1})$ *such that*

(a) $M_{n-1} \cdot a = 0$*, and*

(*b*) *if* $\det(M_n) = 0$, *then* $\sum_{i=2}^{n} a_i x_i = 0$,

where $a = (a_2, \dots, a_n)$, *and* $(x_1, \dots, x_n)$ *is the first row of* M_n.

Proof. Recall (see, e.g., [25, page 22]) that if $\mathrm{rk}(M_{n-1}) = n - 2$, then

$$M_{n-1} \cdot \mathrm{adj}(M_{n-1}) = 0 \qquad \text{and} \qquad \mathrm{rk}\big(\mathrm{adj}(M_{n-1})\big) = 1.$$

It follows that there exists a non-trivial column vector a of $\mathrm{adj}(M_{n-1})$, and $M_{n-1} \cdot a = 0$.

To show that property (*b*) holds, recall that, since M_n is symmetric,

$$\det(M_n) = x_1 \det(M_{n-1}) - \sum_{2 \leqslant i,j \leqslant n} c_{ij} x_i x_j,$$

where c_{ij} are the entries of $\mathrm{adj}(M_{n-1})$. Since $\mathrm{adj}(M_{n-1})$ is a symmetric matrix of rank 1, its entries can be written in the form $c_{ij} = \lambda a_i a_j$ for some $\lambda \in \mathbb{F}_p \setminus \{0\}$. Hence

$$0 = \sum_{2 \leqslant i,j \leqslant n} a_i a_j x_i x_j = \Bigg(\sum_{2 \leqslant i \leqslant n} a_i x_i \Bigg)^2, \tag{4.24}$$

since $\det(M_{n-1}) = \det(M_n) = 0$, as required. □

We now use Lemma 4.6.7 to deduce Lemma 4.6.2, cf. [14, Section 2.2].

Proof of Lemma 4.6.2. By Lemma 4.6.7, it follows that in order to bound the probability that $\mathrm{rk}(M_n) = n - 1$ and $\mathrm{rk}(M_{n-1}) = n - 2$, it suffices to bound the probability that there exists a vector $a \in \mathbb{F}_p^{n-1} \setminus \{0\}$ (unique up to a constant factor) with $M_{n-1} \cdot a = 0$ and $a \cdot x = 0$, where $x \in \{-1, 1\}^{n-1}$ is a random vector chosen uniformly and independent of M_{n-1}.

We will partition this event into 'structured' and 'unstructured' cases, using the event

$$\mathcal{U}_\beta := \big\{ \rho(v) \leqslant \beta \text{ for every vector } v \in \mathbb{F}_p^{n-1} \setminus \{0\} \text{ with } M_{n-1} \cdot v = 0 \big\}.$$

Observe first that, for any $M_{n-1} \in \mathcal{U}_\beta$, and any $a \in \mathbb{F}_p^{n-1} \setminus \{0\}$ with $M_{n-1} \cdot a = 0$, we have

$$\mathbb{P}\big(a \cdot x = 0 \,\big|\, M_{n-1}\big) \leqslant \beta,$$

and hence

$$\mathbb{P}\Big(\big\{ \mathrm{rk}(M_n) = n - 1 \big\} \cap \big\{ \mathrm{rk}(M_{n-1}) = n - 2 \big\} \cap \mathcal{U}_\beta \Big) \leqslant \beta.$$

On the other hand, by the definition (4.20) of $q_n(\beta)$, we have

$$\mathbb{P}\big(\mathcal{U}_\beta^c\big) = \mathbb{P}\big(\exists\, v \in \mathbb{F}_p^{n-1} \setminus \{0\} : M_{n-1} \cdot v = 0 \text{ and } \rho(v) > \beta \big) \leqslant q_{n-1}(\beta).$$

It follows that

$$\mathbb{P}\Big(\big\{\mathrm{rk}(M_n) = n-1\big\} \cap \big\{\mathrm{rk}(M_{n-1}) = n-2\big\}\Big) \leqslant \beta + q_{n-1}(\beta),$$

as required. □

4.6.4 The case $\mathrm{rk}(M_{n-1}) = n-1$

It only remains to prove Lemma 4.6.3. The strategy is similar to that used in the previous subsection (in particular, we will split our event into a 'structured' case and an 'unstructured' case), but now it is trickier to relate our event to $q_n(\beta)$, as we do not have the simple factorisation of the determinant used in Lemma 4.6.7. Instead, we will apply the following 'decoupling' lemma of Costello, Tao and Vu [9].

Lemma 4.6.8 (Lemma 4.7 of [9])**.** *Let X and Y be independent random variables, and let $\mathcal{E}(X,Y)$ be an event that depends on X and Y. Then*

$$\mathbb{P}\big(\mathcal{E}(X,Y)\big) \leqslant \Big(\mathbb{P}\big(\mathcal{E}(X,Y) \cap \mathcal{E}(X',Y) \cap \mathcal{E}(X,Y') \cap \mathcal{E}(X',Y')\big)\Big)^{1/4},$$

where X' and Y' are independent copies of X and Y.

It was remarked in [9] that Lemma 4.6.8 is equivalent to the classical fact (which was essentially proved by Erdős [11] in 1938) that a bipartite graph with parts of size m and n and cmn edges contains at least $c^4m^2n^2$ (possibly degenerate) copies of C_4. Indeed, to deduce Lemma 4.6.8 from this theorem, simply define a bipartite graph, each of whose vertices represents an element of the range of X or Y, and whose edges encode the event $\mathcal{E}$.

In order to state the key technical lemma that we will use to prove Lemma 4.6.3, we need a little notation. Given a vector $v \in \mathbb{F}_p^m$ and a set $J \subset [m]$, let $v_J \in \mathbb{F}_p^{|J|}$ denote the restriction of v to the coordinates of J, and let v_J^* be the vector in $\mathbb{F}_p^m$ whose ith coordinate is $v_i \cdot \mathbb{1}[i \in J]$. Moreover, let $u, u' \in \{-1,1\}^{n-1}$ be chosen uniformly and independently at random, and define $w \in \{-2,0,2\}^{n-1}$ by setting $w_i := u_i - u'_i$ for each $i \in [n-1]$.

The following lemma was essentially proved in [9, Section 4.6] (see also [14, Section 2.3]).

Lemma 4.6.9. *For any non-trivial partition $I \cup J = [n-1]$, we have*

$$\mathbb{P}\big(\mathrm{rk}(M_n) = \mathrm{rk}(M_{n-1}) = n-1\big) \leqslant 2 \cdot \mathbb{E}\Big[\max_{a \in \mathbb{F}_p} \mathbb{P}\big(z_I \cdot w_I = a \,\big|\, M_{n-1}\big)^{1/4} \mathbb{1}\big[\mathrm{rk}(M_{n-1}) = n-1\big]\Big],$$

where $z := M_{n-1}^{-1} \cdot w_J^*$, and the expectation is over the choice of M_{n-1}.

Proof. Let $X := (u_i)_{i\in I}$ and $Y := (u_i)_{i\in J}$ be random variables, and note that u is determined by X and Y. Now define, for each choice of M_{n-1}, an event

$$\mathcal{E}(X,Y) := \big\{\exists\, v \in \mathbb{F}_p^{n-1} : M_{n-1}\cdot v = u \text{ and } u\cdot v \in \{-1,1\}\big\}$$

depending on X and Y. We claim that if $\mathrm{rk}(M_{n-1}) = n-1$, and the first row of M_n is $(x_1, u_1, \dots, u_{n-1})$ for some $x_1 \in \{-1,1\}$, then

$$\big\{\mathrm{rk}(M_n) = n-1\big\} \Rightarrow \big\{u \in \mathcal{E}(X,Y)\big\}.$$

Indeed, since $\det(M_n) = 0 \neq \det(M_{n-1})$ there exists a vector $v \in \mathbb{F}_p^n$ such that $M_n\cdot v = 0$ and $v_1 = -1$. Letting $v' = (v_2,\dots,v_n)$, we see that $M_{n-1}\cdot v' = u$ and $u\cdot v' \in \{-1,1\}$.

Now, for each choice of M_{n-1}, define[5]

$$\mathcal{E}_1 := \mathcal{E}(X,Y)\cap\mathcal{E}(X',Y)\cap\mathcal{E}(X,Y')\cap\mathcal{E}(X',Y').$$

By Lemma 4.6.8, we have

$$\mathbb{P}\big(\mathcal{E}(X,Y)\,\big|\,M_{n-1}\big) \leqslant \mathbb{P}\big(\mathcal{E}_1\,\big|\,M_{n-1}\big)^{1/4},$$

and hence

$$\begin{aligned}\mathbb{P}\Big(\mathrm{rk}(M_n) = \mathrm{rk}(M_{n-1}) = n-1\Big) &\leqslant \mathbb{E}\Big[\mathbb{P}\big(\mathcal{E}(X,Y)\,\big|\,M_{n-1}\big)\mathbb{1}\big[\mathrm{rk}(M_{n-1}) = n-1\big]\Big]\\ &\leqslant \mathbb{E}\Big[\mathbb{P}\big(\mathcal{E}_1\,\big|\,M_{n-1}\big)^{1/4}\mathbb{1}\big[\mathrm{rk}(M_{n-1}) = n-1\big]\Big],\end{aligned}$$

where the expectation is over the choice of M_{n-1}.

To complete the proof of the lemma, it will therefore suffice to show that

$$\mathbb{P}\big(\mathcal{E}_1\,\big|\,M_{n-1}\big) \leqslant 16\cdot\max_{a\in\mathbb{F}_p}\mathbb{P}\big(z_I\cdot w_I = a\,\big|\,M_{n-1}\big) \tag{4.25}$$

for all M_{n-1} with $\mathrm{rk}(M_{n-1}) = n-1$. To prove (4.25), let us fix M_{n-1} (arbitrarily among those with $\mathrm{rk}(M_{n-1}) = n-1$) and set $A := M_{n-1}^{-1}$ and $D := \{-1,1\}$. We claim that if $u \in \mathcal{E}(X,Y)$, then $u^TAu \in D$. To see this, simply observe that

$$u^TAu = u^TA\cdot M_{n-1}v = u^Tv \in \{-1,1\} = D.$$

Recalling that $u = u(X,Y)$, define $f(X,Y) := u^TAu$, and observe that if $\mathcal{E}_1$ holds, then

$$f(X,Y) - f(X',Y) - f(X,Y') + f(X',Y') \in 2D - 2D, \tag{4.26}$$

by the observation above. We claim that the left-hand side of (4.26) is equal to $2z_I\cdot w_I$. To see this, note that

$$f(X,Y) = u^TAu = \sum_{1\leqslant i,j\leqslant n-1} A_{ij}u_iu_j,$$

[5]Here we set $X' := (u_i')_{i\in I}$ and $Y' := (u_i')_{i\in J}$, so X' and Y' are independent copies of X and Y.

and (abusing notation) let us write $f(X,Y)_{ij} := A_{ij}u_iu_j$. Now, observe that if $i,j \in I$, then $f(X,Y)_{ij} = f(X,Y')_{ij}$ and $f(X',Y)_{ij} = f(X',Y')_{ij}$, and therefore

$$\sum_{i,j\in I} f(X,Y)_{ij} - f(X',Y)_{ij} - f(X,Y')_{ij} + f(X',Y)_{ij} = 0.$$

Similarly, if $i,j \in J$ then $f(X,Y)_{ij} = f(X',Y)_{ij}$ and $f(X,Y')_{ij} = f(X',Y')_{ij}$, and hence

$$f(X,Y) - f(X',Y) - f(X,Y') + f(X',Y') = 2\sum_{i\in I}\sum_{j\in J} A_{ij}(u_i - u'_i)(u_j - u'_j).$$

Recalling that $w = u - u'$ and $z_i := \sum_{j\in J} A_{ij}w_j$, it follows that

$$z_I \cdot w_I = \sum_{i\in I}(u_i - u'_i)\sum_{j\in J} A_{ij}(u_j - u'_j),$$

so the left-hand side of (4.26) is equal to $2z_I \cdot w_I$, as claimed. Since $|D| = 2$, it follows that

$$\mathbb{P}\big(\mathcal{E}_1 \,\big|\, M_{n-1}\big) \leqslant 16 \cdot \max_{a\in\mathbb{F}_p} \mathbb{P}\big(z_I \cdot w_I = a \,\big|\, M_{n-1}\big),$$

as claimed. As noted above, this completes the proof of the lemma. □

In the proof of Lemma 4.6.3 we will need the following variant of $\rho(v)$. For any $n \in \mathbb{N}$ and $v \in \mathbb{F}_p^n$, define

$$\rho_{1/2}(v) := \max_{a\in\mathbb{F}_p} \mathbb{P}\big(u_1v_1 + \cdots + u_nv_n = a\big),$$

where $u_1, \ldots, u_n$ are iid random variables taking the value 0 with probability $1/2$, and the values ± 1 each with probability $1/4$. We will need the following simple inequalities.

Lemma 4.6.10 (Lemma 2.8 and 2.9 of [14])**.** *For any $v \in \mathbb{F}_p^n$, and any partition $I \cup J = [n]$,*

$$\rho_{1/2}(v) \leqslant \rho(v) \qquad \text{and} \qquad \rho(v) \leqslant \rho(v_I) \leqslant 2^{|J|}\rho(v).$$

Proof. Observe first that

$$\rho(v) \leqslant \sum_{w\in\{-1,1\}^{|J|}} \mathbb{P}\big(u_J = w\big) \cdot \max_{a\in\mathbb{F}_p} \mathbb{P}\big(u_I \cdot v_I = a - w \cdot v_J \,\big|\, u_J = w\big) \leqslant \rho(v_I).$$

Since $\rho_{1/2}(v) = \rho(v \oplus v)$, it follows that $\rho_{1/2}(v) \leqslant \rho(v)$. Finally, if $a \in \mathbb{F}_p$ maximises $\mathbb{P}\big(u_I \cdot v_I = a\big)$, then

$$\rho(v_I) = 2^{|J|} \cdot \mathbb{P}\big(u_I \cdot v_I = a\big) \prod_{j\in J} \mathbb{P}\big(u_j = 1\big) \leqslant 2^{|J|} \cdot \mathbb{P}\bigg(u \cdot v = a + \sum_{j\in J} v_j\bigg) \leqslant 2^{|J|} \cdot \rho(v),$$

as claimed. □

We are now ready to prove our final lemma, cf. [14, Section 2.3].

Proof of Lemma 4.6.3. Recall that $1 \leqslant k \leqslant n-2$, and let $J \subset [n-1]$ with $|J| = k$. By Lemma 4.6.9, it will suffice to show that

$$\mathbb{E}\Big[\max_{a\in\mathbb{F}_p} \mathbb{P}\big(z_I \cdot w_I = a \,\big|\, M_{n-1}\big)^{1/4} \mathbb{1}\big[\mathrm{rk}(M_{n-1}) = n-1\big]\Big] \leqslant \big(2^{|J|}\beta + 2^{-|J|}\big)^{1/4} + 3^{|J|} q_{n-1}(\beta),$$

where $I = [n] \setminus J$ and $z = M_{n-1}^{-1} \cdot w_J^*$ is defined whenever $\mathrm{rk}(M_{n-1}) = n-1$. Recall that $w \in \{-2,0,2\}^{n-1}$, and observe that therefore $M_{n-1} \cdot z = w_J^* \in W(J)$, where

$$W(J) := \big\{v \in \{-2,0,2\}^{n-1} : v_j = 0 \text{ for all } j \notin J\big\}.$$

We will use the following event to partition into cases:

$$\mathcal{U}_\beta^{(J)} := \Big\{\rho(v) \leqslant \beta \text{ for every vector } v \in \mathbb{F}_p^{n-1} \setminus \{0\} \text{ such that } M_{n-1} \cdot v \in W(J)\Big\}.$$

We will bound the expectation above using the following three claims.

Claim 1: $\mathbb{P}\big(M_{n-1} \notin \mathcal{U}_\beta^{(J)}\big) \leqslant 3^{|J|} q_{n-1}(\beta)$.

Proof of Claim 1. If $\mathcal{U}_\beta^{(J)}$ does not hold for M_{n-1}, then there exists a vector $v \in \mathbb{F}_p^{n-1} \setminus \{0\}$ such that $M_{n-1} \cdot v \in W(J)$ and $\rho(v) > \beta$. For each individual vector $w \in W(J)$, the probability that this holds with $M_{n-1} \cdot v = w$ is at most $q_{n-1}(\beta)$, by (4.20). Hence, summing over $w \in W(J)$, and noting that $|W(J)| = 3^{|J|}$, the claim follows. □

Claim 2: If $\mathrm{rk}(M_{n-1}) = n-1$, then $\mathbb{P}\big(z = 0 \,\big|\, M_{n-1}\big) \leqslant 2^{-|J|}$.

Proof of Claim 2. If $z = 0$ then $w_J^* = M_{n-1} \cdot z = 0$. Since $w_i = 0$ occurs with probability $1/2$ for each $i \in J$, and these events are independent, the claim follows immediately. □

Claim 3: If $M_{n-1} \in \mathcal{U}_\beta^{(J)}$ and $\mathrm{rk}(M_{n-1}) = n-1$, then

$$\max_{a\in\mathbb{F}_p} \mathbb{P}\Big(\big\{z_I \cdot w_I = a\big\} \cap \big\{z \neq 0\big\} \;\big|\; M_{n-1}\Big) \leqslant 2^{|J|}\beta.$$

Proof of Claim 3. Recall that w_J and M_{n-1} together determine z, and that the entries of w_I are independent of w_J, and observe that $\rho_{1/2}(z_I) = \max_{a\in\mathbb{F}_p} \mathbb{P}\big(z_I \cdot w_I = a\big)$. Therefore

$$\mathbb{P}\Big(\big\{z_I \cdot w_I = a\big\} \cap \big\{z \neq 0\big\} \,\big|\, M_{n-1}\Big) \leqslant \mathbb{E}\Big[\rho_{1/2}(z_I)\mathbb{1}\big[z \neq 0\big] \;\big|\; M_{n-1}\Big]$$

for every $a \in \mathbb{F}_p$, where the expectation is over the choice of w_J. Now, by Lemma 4.6.10,

$$\rho_{1/2}(z_I) \leqslant \rho(z_I) \leqslant 2^{|J|}\rho(z).$$

Since $M_{n-1} \in \mathcal{U}_\beta^{(J)}$ and $M_{n-1} \cdot z = w_J^* \in W(J)$, if $z \neq 0$ then $\rho(z) \leqslant \beta$. It follows that

$$\mathbb{E}\Big[\rho_{1/2}(z_I)\mathbb{1}\big[z \neq 0\big] \;\Big|\; M_{n-1}\Big] \leqslant 2^{|J|}\beta,$$

as claimed. □

By Claims 1, 2 and 3, it follows that

$$\mathbb{E}\Big[\max_{a\in\mathbb{F}_p} \mathbb{P}\big(z_I \cdot w_I = a \,\big|\, M_{n-1}\big)^{1/4}\mathbb{1}\big[\mathrm{rk}(M_{n-1}) = n-1\big]\Big] \leqslant \big(2^{|J|}\beta + 2^{-|J|}\big)^{1/4} + 3^{|J|}q_{n-1}(\beta),$$

and, as noted above, this completes the proof of Lemma 4.6.3. □

As shown in Section 4.6.1, this completes the proof of Lemma 4.2.1.

Appendix A

The proof of Theorem 3.9.3

In this section we will deal with some minor technical issues that arise when $\lambda k \approx 2n$, and hence complete the proof of Theorem 3.9.1. First, we need the following variant of Lemma 3.5.4.

Lemma A.0.1. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $n \leqslant \lambda k \leqslant 2n$. Then*

$$|\mathcal{F} \setminus \Lambda^*| \leqslant \big(n - \lambda k/2\big) \cdot |\mathcal{I}|.$$

Proof. We repeat the proof of Lemma 3.5.4, except we now set $d := 1$. To be precise, let $A \in \mathcal{F} \setminus \Lambda^*$, choose $a \in \mathbb{N}$ minimal such that the sets

$$\big\{x \in A : x \leqslant a\big\} \qquad \text{and} \qquad \big\{x \in A : x > a + \lambda k/2\big\}$$

are both non-empty and together contain at most δk elements, define $\varphi(A) := A - a$, and observe that $\varphi(A) \in \mathcal{I}$ (cf. the proof of Lemma 3.5.4). Now, for each set $S \in \mathcal{I}$, there are at most $n - \lambda k/2$ choices for a such that $a + S \subset [n]$, and therefore

$$|\varphi^{-1}(S)| \leqslant n - \lambda k/2$$

for every $S \in \mathcal{I}$, as required. □

We also need the following variant of Lemma 3.9.2.

Lemma A.0.2. *Let $\lambda \geqslant 3$ and $n, k \in \mathbb{N}$, with $n \leqslant \lambda k \leqslant 2n$. Then*

$$\big|\big\{A \subset [n] : |A| = k,\, |A + A| \leqslant \lambda k\big\}\big| \geqslant \frac{1}{\lambda^2} \cdot \big(n - \lambda k/2 + 1\big)\binom{\lambda k/2}{k}.$$

Proof. We consider, for each arithmetic progression P of length $\lambda k/2$ in $[n]$, all subsets $A \subset P$ of size k containing both endpoints of P. All of these sets are distinct, and all satisfy $|A+A| \leqslant \lambda k$.

There are $n - \lambda k/2 + 1$ choices for the arithmetic progression, and therefore

$$|\Lambda| \geqslant \big(n - \lambda k/2 + 1\big)\binom{\lambda k/2 - 2}{k-2} \geqslant \frac{1}{\lambda^2} \cdot \big(n - \lambda k/2 + 1\big)\binom{\lambda k/2}{k},$$

as claimed. □

We can now deduce Theorem 3.9.1 in the case $\lambda k \geqslant n$.

Proof of Theorem 3.9.1. Observe that the theorem is trivial if $\lambda k/2 \geqslant n$ (since in this case $P = [n]$ satisfies the conditions), so let us assume that $n \leqslant \lambda k < 2n$. Replacing Lemma 3.5.4 by Lemma A.0.1 in the proof of Lemma 3.5.1, and recalling that $\varepsilon > e^{-\delta^2 k}$, we obtain

$$|\Lambda \setminus \Lambda^*| \leqslant \big(n - \lambda k/2\big) \cdot |\mathcal{I}| + \frac{\varepsilon}{2\lambda^2}\binom{\lambda k/2}{k}.$$

Now, by Lemmas 3.6.1 and 3.7.1, we have

$$|\mathcal{I}| = |\mathcal{S}| + |\mathcal{D}| \leqslant \frac{\varepsilon}{2\lambda^2}\binom{\lambda k/2}{k}.$$

Finally, by Lemma A.0.2, it follows that

$$|\Lambda \setminus \Lambda^*| \leqslant \frac{\varepsilon}{\lambda^2} \cdot \big(n - \lambda k/2 + 1\big)\binom{\lambda k/2}{k} \leqslant \varepsilon|\Lambda|,$$

as required. □

Bibliography

[1] N. Alon, J. Balogh, R. Morris, and W. Samotij. A refinement of the Cameron–Erdős conjecture. *Proceedings of the London Mathematical Society*, 108(1):44–72, 2013.

[2] N. Alon, J. Balogh, R. Morris, and W. Samotij. Counting sum-free sets in abelian groups. *Israel Journal of mathematics*, 199(1):309–344, 2014.

[3] J. Balogh, H. Liu, and M. Sharifzadeh. The number of subsets of integers with no k-term arithmetic progression. *International Mathematics Research Notices*, 2017(20):6168–6186, 2016.

[4] J. Balogh, Hong Liu, M. Sharifzadeh, and A. Treglown. Sharp bound on the number of maximal sum-free subsets of integers. *Journal of the European Mathematical Society*, 20(8):1885–1911, 2018.

[5] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28(3):669–709, 2015.

[6] J. Balogh, R. Morris, and W. Samotij. The method of hypergraph containers. In *Proceedings of the International Congress of Mathematicians*, 2018.

[7] J. Balogh, R. Morris, W. Samotij, and L. Warnke. The typical structure of sparse K_{r+1}-free graphs. *Transactions of the American Mathematical Society*, 368(9):6439–6485, 2016.

[8] J. Bourgain, V. H. Vu and P. M. Wood, On the singularity probability of discrete random matrices, *J. Funct. Anal.*, **258** (2010), 559–603.

[9] K. P. Costello, T. Tao and V. Vu, Random symmetric matrices are almost surely nonsingular, *Duke Math. J.*, **135** (2006), 395–413.

[10] D. Dellamonica, Jr., Y. Kohayakawa, S. J. Lee, V. Rödl, and W. Samotij. On the number of B_h-sets. submitted.

[11] P. Erdős, On sequences of integers no one of which divides the product of two others and on some related problems, *Mitt. Forsch.-Inst. Math. Mech. Univ. Tomsk*, **2** (1938), 74–82.

[12] P. Erdős, On a lemma of Littlewood and Offord, *Bull. Amer. Math. Soc.*, **51** (1945), 898–902.

[13] A. Ferber and V. Jain, K. Luh, and W. Samotij, On the counting problem in inverse Littlewood–Offord theory, arXiv:1904.10425.

[14] A. Ferber and V. Jain, Singularity of random symmetric matrices – a combinatorial approach to improved bounds, arXiv:1809.04718.

[15] P. Frankl and Z. Füredi, Solution of the Littlewood–Offord problem in high dimensions, *Ann. Math.*, **128** (1988), 259–270.

[16] G. A. Freiman. The addition of finite sets I. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, (6):202–213, 1959.

[17] G. A. Freĭman. Nachala strukturnoĭ teorii slozheniya mnozhestv (Russian).

[18] G. A. Freĭman. Foundations of a structural theory of set addition. *Translations of Mathematical Monographs*, 37, 1973.

[19] B. Green. Approximate algebraic structure. *Proc. Int. Cong. Math.*, 1:341–367.

[20] B. Green. The Cameron–Erdős conjecture. *Bulletin of the London Mathematical Society*, 36(6):769–778, 2004.

[21] B. Green. Counting sets with small sumset, and the clique number of random Cayley graphs. *Combinatorica*, 25(3):307–326, 2005.

[22] B. Green and R. Morris. Counting sets with small sumset and applications. *Combinatorica*, 36(2):129–159, 2016.

[23] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, *Period. Math. Hungar.*, **8** (1977), 197–211.

[24] Y. O. Hamidoune and O. Serra. A note on Pollard's theorem. *arXiv preprint arXiv:0804.2593*, 2008.

[25] R. A. Horn and C. R. Johnson, Matrix Analysis (second edition), Cambridge University Press, 2013.

[26] J. Kahn, J. Komlós and E. Szemerédi, On the probability that a random ± 1 matrix is singular, *J. Amer. Math. Soc.*, **8** (1995), 223–240.

[27] J. Komlós, On the determinant of $(0, 1)$ matrices, *Studia Sci. Math. Hungar.*, **2** (1967), 7–22.

[28] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation III, Rec. Math. (Mat. Sbornik) N.S., **12** (1943), 277–286.

[29] P. Mazur. A structure theorem for sets of small popular doubling. *Acta Arithmetica*, 171:221–239, 2015.

[30] R. Morris, W. Samotij, and D. Saxton. An asymmetric container lemma and the structure of graphs with no induced 4-cycle. *submitted, arXiv:1806.03706*, 2018.

[31] H. H. Nguyen, Inverse Littlewood–Offord problems and the singularity of random symmetric matrices, *Duke Math. J.*, **161** (2012), 545–586.

[32] H. H. Nguyen and V. H. Vu, Optimal inverse Littlewood–Offord theorems, *Adv. Math.*, **226** (2011), 5298–5319.

[33] H. H. Nguyen and V. H. Vu, Small ball probability, inverse theorems, and applications, In: Erdős Centennial, pages 409–463. Springer, 2013.

[34] A. M. Odlyzko, On subspaces spanned by random selections of ± 1 vectors, *J. Combin. Theory Ser. A*, **47** (1988), 124–133.

[35] H. Plünnecke. *Eigenschaften und Abschätzungen von Wirkungsfunktionen.* Number 22. Gesellschaft für Mathematik und Datenverarbeitung, 1961.

[36] H. Plünnecke. Eine zahlentheoretische anwendung der graphentheorie. *Journal für die reine und angewandte Mathematik*, 243:171–183, 1970.

[37] J. M. Pollard. A generalisation of the theorem of Cauchy and Davenport. *Journal of the London Mathematical Society*, 2(3):460–462, 1974.

[38] M. Rudelson and R. Vershynin, The Littlewood–Offord problem and invertibility of random matrices, *Adv. Math.*, **218** (2008), 600–633.

[39] M. Rudelson and R. Vershynin, Smallest singular value of a random rectangular matrix, *Comm. Pure Appl. Math.*, **62** (2009), 1707–1739.

[40] M. Rudelson and R. Vershynin, Non-asymptotic theory of random matrices: extreme singular values, *Proc. Int. Cong. Math.*, Hyderabad, 2010, Vol. 3, 1576–1602.

[41] I. Z. Ruzsa. On the cardinality of A+A and A-A. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, volume 2, pages 933–938, 1978.

[42] I. Z. Ruzsa. An application of graph theory to additive number theory. *Scientia, Ser. A*, 3(97-109):9, 1989.

[43] I. Z. Ruzsa. Generalized arithmetical progressions and sumsets. *Acta Mathematica Hungarica*, 65(4):379–388, 1994.

[44] I. Z. Ruzsa. An analog of freiman's theorem in groups. *Astérisque*, 258(199):323–326, 1999.

[45] T. Sanders. The structure theory of set addition revisited. *Bulletin of the American Mathematical Society*, 50(1):93–127, 2013.

[46] A. A. Sapozhenko. The Cameron-Erdős conjecture. *Dokl. Akad. Nauk*, 393(6):749–752, 2003.

[47] D. Saxton and A. Thomason. Hypergraph containers. *Invent. Math.*, 201(3):925–992, 2015.

[48] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[49] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, *J. Amer. Math. Soc.*, **20** (2007), 603–628.

[50] T. Tao and V. Vu, Inverse Littlewood–Offord theorems and the condition number of random discrete matrices, *Ann. Math.*, **169** (2009), 595–632.

[51] T. Tao and V. Vu, From the Littlewood–Offord problem to the circular law: universality of the spectral distribution of random matrices, *Bull. Amer. Math. Soc.*, **46** (2009), 377–396.

[52] K. Tikhomirov, Singularity of random Bernoulli matrices, arXiv:1812.09016.

[53] R. Vershynin, Invertibility of symmetric random matrices, *Random Structures Algorithms*, **44** (2014), 135–182.

[54] V. Vu, Combinatorial problems in random matrix theory, *Proc. Int. Cong. Math.*, Seoul, 2014, Vol. 4, 489–508.