On the singularity probability of discrete random matrices

Jean Bourgain, Van H. Vu, and Philip Matchett Wood

April 26, 2009

Abstract

Let n be a large integer and M_n be an n by n complex matrix whose entries are independent (but not necessarily identically distributed) discrete random variables. The main goal of this paper is to prove a general upper bound for the probability that M_n is singular.

For a constant $0 < p < 1$ and a constant positive integer r, we will define a property *p-bounded of exponent r*. Our main result shows that if the entries of M_n satisfy this property, then the probability that M_n is singular is at most $\left(p^{1/r} + o(1)\right)^n$. All of the results in this paper hold for any characteristic zero integral domain replacing the complex numbers.

In the special case where the entries of M_n are "fair coin flips" (taking the values $+1, -1$ each with probability $1/2$), our general bound implies that the probability that M_n is singular is at most $\left(\frac{1}{\sqrt{2}} + o(1)\right)^n$, improving on the previous best upper bound of $\left(\frac{3}{4} + o(1)\right)^n$, proved by Tao and Vu [11].

In the special case where the entries of M_n are "lazy coin flips" (taking values $+1, -1$ each with probability $1/4$ and value 0 with probability $1/2$), our general bound implies that the probability that M_n is singular is at most $\left(\frac{1}{2} + o(1)\right)^n$, which is asymptotically sharp.

Our method is a refinement of those from [4] and [11]. In particular, we make a critical use of the Structure Theorem from [11], which was obtained using tools from additive combinatorics.

1 Introduction

Let n be a large integer and M_n be an n by n random matrix whose entries are independent (but not necessarily identically distributed) discrete random variables taking values in the complex numbers. The problem of estimating the probability that M_n is singular is a basic problem in the theory of random matrices and combinatorics. The goal of this paper is to give a bound that applies to a large variety of distributions. The general statement (Theorem 2.2) is a bit technical, so we will first discuss a few corollaries concerning special cases.

The most famous special case is when the entries of M_n are independent identically distributed (i.i.d.) Bernoulli random variables (taking values ± 1 with probability $1/2$). The following conjecture has been open for quite some time:

Conjecture 1.1. *For $M_{\pm 1,n}$ an n by n matrix with each entry an i.i.d. Bernoulli random variable taking the values $+1$ and -1 each with probability $1/2$,*

$$\Pr(M_{\pm 1,n} \text{ is singular}) = \left(\frac{1}{2} + o(1)\right)^n.$$

It is easy to verify that the singularity probability is at least $(1/2)^n$ by considering the probability that there are two equal rows (or columns).

Even in the case of i.i.d. Bernoulli random variables, proving that the singularity probability is $o(1)$ is not trivial. It was first done by Komlós in 1967 [5] (see also [6]; [9] generalizes Komlós's bound to other integer distributions). The first exponential bound was proven by Kahn, Komlós, and Szemerédi [4], who showed that $\Pr(M_{\pm 1,n} \text{ is singular}) \leq .999^n$. This upper bound was improved upon by Tao and Vu in [10] to $.958^n$. A more significant improvement was obtained by the same authors in [11]:

$$\Pr(M_{\pm 1,n} \text{ is singular}) \leq \left(\frac{3}{4} + o(1)\right)^n. \tag{1}$$

This improvement was made possible through the discovery of a new theorem [11, Theorem 5.2] (which was called the Structure Theorem in [11]), which gives a complete characterization of a set with certain additive properties. The Structure Theorem (to be more precise, a variant of it) will play a critical role in the current paper as well.

Our general result has the following corollary in the Bernoulli case:

$$\Pr(M_{\pm 1,n} \text{ is singular}) \leq \left(\frac{1}{\sqrt{2}} + o(1)\right)^n, \tag{2}$$

which gives a slight improvement over Inequality (1) (since $1/\sqrt{2} \approx 0.7071 < .75$).

Let us now discuss a more general class of random matrices. Consider the random variable $\gamma^{(\mu)}$ defined by

$$\gamma^{(\mu)} := \begin{cases} +1 & \text{with probability } \mu/2 \\ 0 & \text{with probability } 1-\mu \\ -1 & \text{with probability } \mu/2, \end{cases} \tag{3}$$

and let $M^{(\mu)}_{\pm 1,n}$ be an n by n matrix with each entry an independent copy of $\gamma^{(\mu)}$. The random variable $\gamma^{(\mu)}$ plays an important role in [4, 10, 11], and the matrices $M^{(\mu)}_{\pm 1,n}$ are of interest in their own right. In fact, giving zero a large weight is a natural thing to do when one would like to (randomly) sparsify a matrix, a common operation used in randomized algorithms (the values of ± 1, as the reader will see, are not so critical). Our general result implies the following upper bounds:

$$\Pr(M^{(\mu)}_{\pm 1,n} \text{ is singular}) \leq (1 - \mu + o(1))^n \qquad \text{for } 0 \leq \mu \leq \frac{1}{2} \tag{4}$$

$$\Pr(M^{(\mu)}_{\pm 1,n} \text{ is singular}) \leq \left(\frac{2\mu+1}{4} + o(1)\right)^n \qquad \text{for } \frac{1}{2} \leq \mu \leq 1 \tag{5}$$

$$\Pr(M^{(\mu)}_{\pm 1,n} \text{ is singular}) \leq \left(\sqrt{1 - 2\mu + \frac{3}{2}\mu^2} + o(1)\right)^n \qquad \text{for } 0 \leq \mu \leq 1. \tag{6}$$

Note that Inequality (5) implies Inequality (1) and that Inequality (6) implies Inequality (2) (in both cases setting $\mu = 1$).

Figure 1 summarizes the upper bounds from Inequalities (4), (5), and (6) and also includes the following lower bounds:

$$(1-\mu+o(1))^n \leq \Pr(M_{\pm 1,n}^{(\mu)} \text{ is singular}) \qquad \text{for } 0 \leq \mu \leq 1 \tag{7}$$

$$\left(1-2\mu+\frac{3}{2}\mu^2+o(1)\right)^n \leq \Pr(M_{\pm 1,n}^{(\mu)} \text{ is singular}) \qquad \text{for } 0 \leq \mu \leq 1. \tag{8}$$

These lower bounds can be derived by computing the probability that one row is all zeros (Inequality (7)) or that there is a dependency between two rows (Inequality (8)). Note that in the case where $\mu \leq 1/2$, the upper bound in Inequality (4) asymptotically equals the lower bound in Inequality (7), and thus our result is the best possible in this case. We also used a Maple program to derive the formulas for lower bounds resulting from a dependency between three, four, or five rows; however, these lower bounds were inferior to those in Inequality (7) and Inequality (8).

We will now present another corollary of the main theorem that has a somewhat different flavor. In this corollary, we treat partially random matrices, which may have many deterministic rows. Our method allows us to obtain exponential bounds so long as there are still at most $c \ln n$ random rows, where $c > 0$ is a particular constant.

Corollary 1.2. *Let p be a real constant between 0 and 1, let c be any positive constant less than $1/\ln(1/p)$, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq O(1)$. Let $N_{\mathfrak{f},n}$ be an n by n complex matrix in which $\mathfrak{f} \leq c \ln n$ rows contain fixed, non-random elements of S and where the other rows contain entries that are independent random variables taking values in S. If the fixed rows are linearly independent and if for every random entry α, we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_{\mathfrak{f},n} \textit{ is singular}) \leq (\sqrt{p}+o(1))^n .$$

Notice that the case $\mathfrak{f} = 0$ and $p = 1/2$ also implies Inequality (2).

Remark 1.3 (Other exponential bounds). The focus of this paper is optimizing the base of the exponent in bounds on the singularity probability for discrete random matrices. One main tool in this optimization is the use of a structure theorem similar to [11, Theorem 5.2] (see Theorem 6.1 below); however, using such a theorem requires additional assumptions to be placed on the values that can appear as entries, and in particular, this is why we assume in Corollary 1.2 that the set S has cardinality $|S| \leq O(1)$ and that $\mathfrak{f} \leq c \ln n$. If one is interested in an exponential bound where there are no conditions on $\mathfrak{f}$ or on the set S (at the expense of having an unspecified constant for the base of the exponential), one can follow the analysis in [10], which does not make use of a structure theorem, along with ideas in this paper to get a result of the following form:

Theorem 1.4. *For every $\epsilon > 0$ there exists $\delta > 0$ such that the following holds. Let $N_{\mathfrak{f},n}$ be an n by n complex matrix in which $\mathfrak{f}$ rows contain fixed, non-random entries and where the other rows contain entries that are independent discrete random variables. If the fixed rows have co-rank k and if for every random entry α, we have $\max_x \Pr(\alpha = x) \leq 1-\epsilon$, then for all sufficiently large n*

$$\Pr(N_{\mathfrak{f},n} \textit{ has co-rank } > k) \leq (1-\delta)^{n-\mathfrak{f}}.$$

Note that Theorem 1.4 holds for any $\mathfrak{f}$ and k, and so in particular, an exponential bound on the singularity probability is achieved whenever $k = 0$ and $\mathfrak{f} \leq cn$, where $c < 1$ is a constant. Also note that the theorem allows the random entries to have discrete distributions taking infinitely many

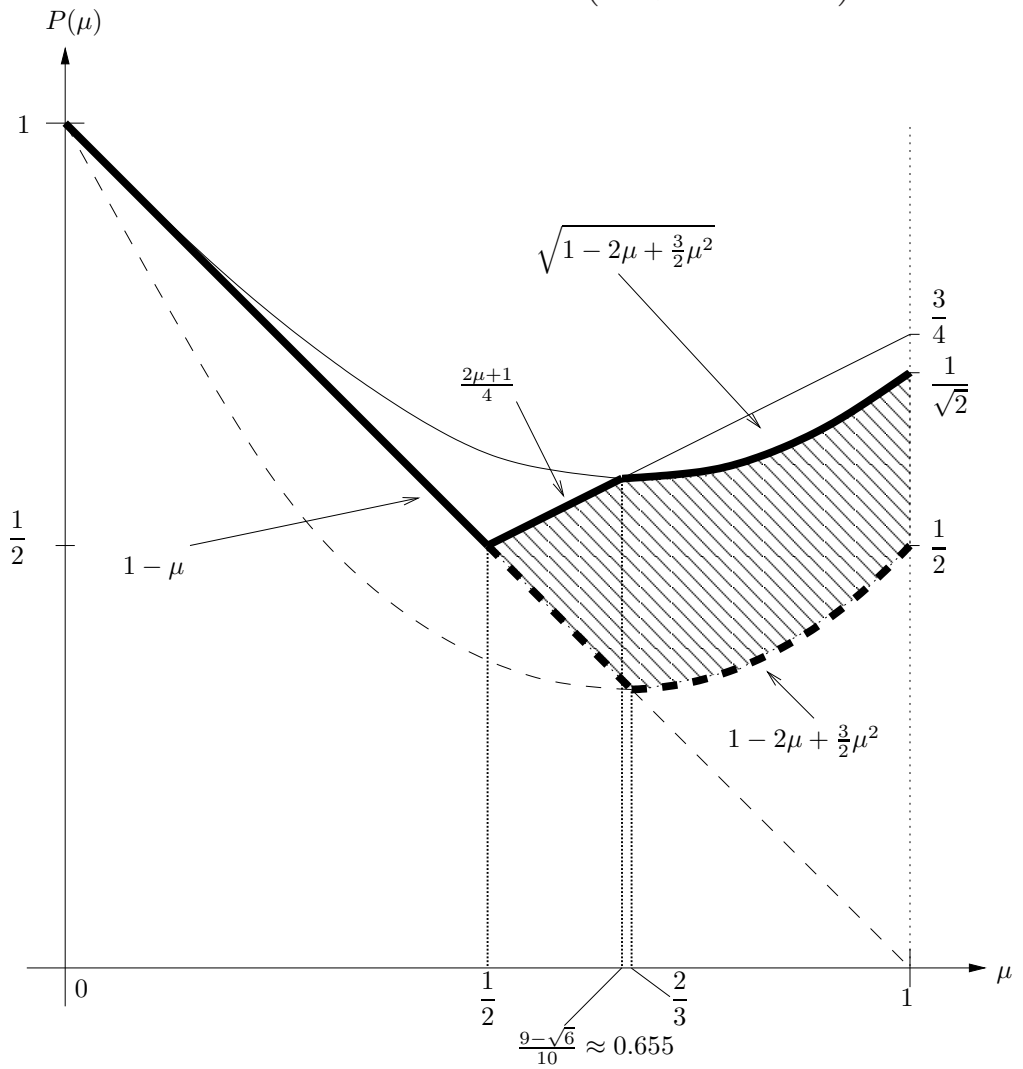

Figure 1: Let $P(\mu) := \lim_{n\to\infty} \Pr\left(M_{\pm 1,n}^{(\mu)} \text{ is singular}\right)^{1/n}$, where $M_{\pm 1,n}^{(\mu)}$ is the n by n matrix with independent random entries taking the value 0 with probability $1-\mu$ and the values $+1$ and -1 each with probability $\mu/2$. The solid lines denote the upper bounds on $P(\mu)$ given by Inequalities (4), (5), and (6), and the dashed lines denote the lower bounds given by Inequalities (7) and (8). The upper and lower bounds coincide for $0 \le \mu \le \frac{1}{2}$, and the shaded area shows the difference between the best known upper and lower bounds for $\frac{1}{2} \le \mu \le 1$. The straight line segments from the point $(0, 1)$ to $(1/2, 1/2)$ and from the point $(1/2, 1/2)$ to $(1, 3/4)$ represent the best upper bounds we have derived using the ideas in [11], and the curve $1 - 2\mu + \frac{3}{2}\mu^2$ for $0 \le \mu \le 1$ represents a sometimes-better upper bound we have derived by adding a new idea. Note that the upper bounds given here also apply to the singularity probability of a random matrix with independent entries having arbitrary symmetric distributions in a set S of complex numbers, so long as each entry is 0 with probability $1 - \mu$ and the cardinality of S is $|S| \le O(1)$ (see Corollary 3.1).

values. Corollary 3.6 proves a version of Theorem 1.4 with a much better exponential bound, given some additional conditions.

The structure of the rest of the paper is as follows. In Section 2 we define p-bounded of exponent r and state the main theorem of this paper. In Section 3, we discuss some corollaries of Theorem 2.2. In particular, we will:

(A) prove Inequalities (4), (5), and (6);

(B) prove general bounds on the singularity probability for discrete random matrices with entries that have symmetric distributions and with entries that have asymmetric distributions;

(C) Prove a version of Corollary 1.2 (namely, Corollary 3.5) that holds for up to $o(n)$ fixed rows, assuming that the entries in the fixed rows take integer values between $-C$ and C for any positive constant C; and

(D) prove that the probability that random matrices with integer entries have a rational eigenvalue is exponentially small.

In Section 4, we discuss Lemma 4.1, a result that is proved in [13] using standard tools from algebraic number theory and algebraic geometry. Lemma 4.1 reduces the question of bounding the singularity probability of a random matrix with entries in $\mathbb{C}$ to a question of bounding the singularity probability of a random matrix with entries in $\mathbb{Z}/Q\mathbb{Z}$ for some large prime Q (in fact, it is possible to replace $\mathbb{C}$ with any characteristic zero integral domain). The proof of Theorem 2.2 is outlined in Section 5, where we also prove some of the easier lemmas needed for the theorem. In Section 6, we state a structure theorem (Theorem 6.1) that completes the proof of our Theorem 2.2 and that is very similar to [11, Theorem 5.2] (which is the Structure Theorem in [11]). We discuss the proof of Theorem 6.1, which uses discrete Fourier analysis and tools from additive combinatorics, in Sections 7 and 8. Finally, in Section 9 we show that the entire argument proving Theorem 2.2 can be generalized to random complex matrices with $\mathfrak{f}$ rows of the matrix containing fixed, non-random entries, so long as $\mathfrak{f} \leq c \ln n$ for a particular constant $c > 0$ (this leads to Corollary 1.2).

2 The general theorem

To prove the results in Inequalities (1) and (2) (and also the results in [4] and [10]), one basic idea is to replace entries of a random matrix with independent copies of the random variable $\gamma^{(\mu)}$ or $2\gamma^{(\mu)}$ (see Equation (3)). One key idea in proving the more general results of the current paper is replacing the entries of a random matrix with more complicated symmetric discrete random variables.

A generalized arithmetic progression of rank $\mathfrak{r}$ is a set of the form $\{v_0 + m_1 v_1 + \cdots + m_{\mathfrak{r}} v_{\mathfrak{r}} : |m_i| \leq M_i/2\}$, where the v_i are elements of a $\mathbb{Z}$-module and the m_i and $M_i > 0$ are integers. Note that whenever the term "symmetric" is used in this paper, it will apply to the distribution of a random variable or to a generalized arithmetic progression; in particular, the term will never apply to matrices. Also, throughout this paper we will use the notation

$$e(x) := \exp(2\pi i x).$$

The following definition lies at the heart of our analysis.

Definition 2.1 (p-bounded of exponent r)**.** Let p be a positive constant such that $0 < p < 1$ and let r be a positive integer constant. A random variable α taking values in the integers (or, respectively, the integers modulo some large prime Q) is *p-bounded of exponent r* if

(i) $\max_x \Pr(\alpha = x) \leq p$, and

if there exists a constant q where $0 < q \leq p$ and a $\mathbb{Z}$-valued (or, respectively, a $\mathbb{Z}/Q\mathbb{Z}$-valued) symmetric random variable $\beta^{(\mu)}$ taking the value 0 with probability $1 - \mu = p$ such that the following two conditions hold:

(ii) $q \leq \min_x \Pr(\beta^{(\mu)} = x)$ and $\max_x \Pr(\beta^{(\mu)} = x) \leq p$, and

(iii) the following inequality holds for every $t \in \mathbb{R}$:

$$|\mathbb{E}(e(\alpha t))|^r \leq \mathbb{E}\left(e(\beta^{(\mu)} t)\right)$$

Here, if the values of α and $\beta^{(\mu)}$ are in $\mathbb{Z}/Q\mathbb{Z}$, we view those values as integers in the range $(-Q/2, Q/2)$ (note that each element in $\mathbb{Z}/Q\mathbb{Z}$ has a unique such integer representation).

We will define p-bounded of exponent r for collections of random variables below, but first we note that the conditions above are easy to verify in practice. In particular, if we have a symmetric random variable

$$\beta^{(\mu)} = \begin{cases} b_\ell & \text{with probability } \mu p_\ell/2 \\ \vdots & \quad\vdots \\ b_1 & \text{with probability } \mu p_1/2 \\ 0 & \text{with probability } 1-\mu \\ -b_1 & \text{with probability } \mu p_1/2 \\ \vdots & \quad\vdots \\ -b_\ell & \text{with probability } \mu p_\ell/2, \end{cases} \tag{9}$$

where $b_s \in \mathbb{Z}$ for all s (or, respectively, $b_s \in \mathbb{Z}/Q\mathbb{Z}$ for all s), then condition (iii) becomes

$$|\mathbb{E}(e(\alpha t))|^r \leq \mathbb{E}\left(e(\beta^{(\mu)} t)\right) = 1 - \mu + \mu \sum_{s=1}^{\ell} p_s \cos 2\pi b_s t, \tag{10}$$

where the equality on the right-hand side is a simple expected value computation.

We say that a collection of random variables $\{\alpha_{jk}\}_{j,k=1}^n$ is p-bounded of exponent r if each α_{jk} is p-bounded of exponent r with the same constants p, q, and r; and, importantly, the same value of $\mu = 1 - p$. We also make the critical assumption that the set of all values that can be taken by the $\beta_{jk}^{(\mu)}$ has cardinality $O(1)$ (a relaxation of this assumption is discussed in Remark 8.5). However, the definition of $\beta_{jk}^{(\mu)}$ is otherwise allowed to vary with j and k. Also, we will use S to denote the set of all possible values taken by the random variables α_{jk}, and we will assume that the cardinality of S is at most $|S| \leq n^{o(n)}$.

If α takes non-integer values in $\mathbb{C}$, we need to map those values to a finite field of prime order so that we may use Definition 2.1, and for this task we will apply Lemma 4.1, which was proved

in [13]. We say that α is p-bounded of exponent r if and only if for each prime Q in an infinite sequence of primes produced by Lemma 4.1, we have $\phi_Q(\alpha)$ is p-bounded of exponent r, where ϕ_Q is the ring homomorphism described in Lemma 4.1 that maps S, the finite set of all possible values taken by the α_{jk}, into $\mathbb{Z}/Q\mathbb{Z}$ in such a way that for any matrix $N_n := (s_{jk})$ with entries in S, the determinant of N_n is zero if and only if the determinant of $\phi_Q(N_n) := (\phi_Q(s_{jk}))$ is zero.

Theorem 2.2. *Let p be a positive constant such that $0 < p < 1$, let r be a positive integer constant, and let S be a generalized arithmetic progression in the complex numbers with rank $O(1)$ (independent of n) and with cardinality at most $|S| \leq n^{o(n)}$. Let N_n be an n by n matrix with entries α_{jk}, each of which is an independent random variable taking values in S. If the collection of random variables $\{\alpha_{jk}\}_{1 \leq j,k \leq n}$ is p-bounded of exponent r, then*

$$\Pr(N_n \text{ is singular}) \leq (p^{1/r} + o(1))^n.$$

In the motivating examples of Section 1 (excluding Corollary 1.2), we discussed the case where the entries of the matrix are i.i.d.; however, in general the distributions of the entries are allowed to differ (and even depend on n), so long as the entries all take values in the same structured set S described above. The condition that S has additive structure seems to be an artifact of the proof (in particular, at certain points in the proof of Theorem 6.1, we need the set $\left\{\sum_{j=1}^n x_j : x_j \in S \text{ for all } j\right\}$ to have cardinality at most $n^{o(n)}$). The easiest way to guarantee that S has the required structure is to assume that the set of values taken by all the α_{jk} has cardinality at most $O(1)$, and this is the approach we take for the corollaries in Section 3, since it also makes it easy to demonstrate that the collection of entries is p-bounded of exponent r.

Remark 2.3 (Strict positivity in Inequality (10)). Note that the constants μ, p_s, b_s must be such that the right-hand side of Equation (10) is non-negative. It turns out for the proof of Theorem 2.2 that we will need slightly more. At one point in the proof, we will apply Lemma 7.3, for which we we must assume that there exists a very small constant $\epsilon_{-1} > 0$ such that $\mathbb{E}(e(\beta_{jk}^{(\mu)} t)) > \epsilon_{-1}$ for all t and for all $\beta_{jk}^{(\mu)}$ used in the definition of p-bounded of exponent r. Of course, if the expectations are not strictly larger than ϵ_{-1}, we can simply reduce μ by $\epsilon_{-1} > 0$. Then, since we are assuming $1 - \mu = p$, we clearly have that all the α_{jk} are $(p + \epsilon_{-1})$-bounded of exponent r (by using $\beta_{jk}^{(\mu-\epsilon_{-1})}$ instead of $\beta_{jk}^{(\mu)}$) and we have that $\mathbb{E}(e(\beta_{jk}^{(\mu-\epsilon_{-1})} t)) > \epsilon_{-1} > 0$. Since Theorem 2.2 would thus yield a bound of $\left((p + \epsilon_{-1})^{1/r} + o(1)\right)^n$ for every $\epsilon_{-1} > 0$, we can conclude a bound of $\left(p^{1/r} + o(1)\right)^n$ by letting ϵ_{-1} tend to 0. Thus, without loss of generality, we will assume that $\mathbb{E}(e(\beta_{jk}^{(\mu)} t)) > \epsilon_{-1}$ for all t and for all $\beta_{jk}^{(\mu)}$ used in the definition of p-bounded of exponent r.

3 Some corollaries of Theorem 2.2

In this section, we will state a number of corollaries of Theorem 2.2, starting with short proofs of Inequalities (4), (5), and (6). The two most interesting results in this section will be more general: first (in Section 3.2), we will show an exponential bound on the singularity probability for a matrix with independent entries each a symmetric random variable taking values in $S \subset \mathbb{C}$, where $|S| \leq O(1)$ and assuming that each entry takes the value 0 with probability $1 - \mu$; and second (in Section 3.3), we will describe a similar (and sometimes better) bound when the condition that the random variables have symmetric distributions is replaced with the assumption that no entry takes

a value with probability greater than p. In the first case, the bound will depend only the value of μ, and in the second case, the bound will depend only on the value of p. In Section 3.4, we will show an exponential bound on the singularity probability for an n by n matrix with $\mathfrak{f} = o(n)$ fixed rows containing small integer values and with the remaining rows containing independent random variables taking values in $S \subset \mathbb{C}$, where $|S| \leq O(1)$ (this is similar to Corollary 1.2, which is proved in Section 9). Finally, in Section 3.5, we will prove an exponential upper bound on the probability that a random integer matrix has a rational eigenvalue.

In each corollary, we will use the definition of p-bounded of exponent 1 and of exponent 2. The definition of p-bounded of exponent 2 is particularly useful, since then the absolute value on the left-hand side of Inequality (10) is automatically dealt with; however, when μ is small (for example whenever $\mu \leq 1/2$), one can get better bounds by using p-bounded of exponent 1. We have not yet found an example where the best possible bound from Theorem 2.2 is found by using p-bounded of an exponent higher than 2.

3.1 Proving Inequalities (4), (5), and (6)

To prove Inequality (4), we note for $0 \leq \mu \leq \frac{1}{2}$ that (using the definition in Equation (3) of $\gamma^{(\mu)}$)

$$\left|\mathbb{E}(e(\gamma^{(\mu)}t))\right| = 1 - \mu + \mu\cos(2\pi t),$$

and thus $\gamma^{(\mu)}$ is $(1-\mu)$-bounded of exponent 1 (i.e., take $\beta^{(\mu)} := \gamma^{(\mu)}$), and so Inequality (4) follows from Theorem 2.2.

To prove Inequality (5), we note for $\frac{1}{2} \leq \mu \leq 1$ that

$$\left|\mathbb{E}(e(\gamma^{(\mu)}t))\right| = |1 - \mu + \mu\cos(2\pi t)| \leq \left(\frac{2\mu+1}{4}\right) + (1-\mu)\cos(2\pi t) + \left(\frac{2\mu-1}{4}\right)\cos(4\pi t)$$

(the inequality above may be checked by squaring both sides and expanding as polynomials in $\cos(2\pi t)$). Thus, we can take

$$\beta^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{2\mu-1}{8} \\ -2 & \text{with probability } \frac{2\mu-1}{8} \\ +1 & \text{with probability } \frac{1-\mu}{2} \\ -1 & \text{with probability } \frac{1-\mu}{2} \\ 0 & \text{with probability } \frac{2\mu+1}{4} \end{cases}$$

to see that $\gamma^{(\mu)}$ is $\left(\frac{2\mu+1}{4}\right)$-bounded of exponent 1, and so Inequality (5) follows from Theorem 2.2.

To prove Inequality (6), we note for $0 \leq \mu \leq 1$ that

$$\left|\mathbb{E}(e(\gamma^{(\mu)}t))\right|^2 = |1 - \mu + \mu\cos(2\pi t)|^2 = 1 - 2\mu + \frac{3}{2}\mu^2 + 2(1-\mu)\mu\cos(2\pi t) + \left(\frac{\mu^2}{2}\right)\cos(4\pi t).$$

Thus, we can take

$$\beta^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{\mu^2}{4} \\ -2 & \text{with probability } \frac{\mu^2}{4} \\ +1 & \text{with probability } (1-\mu)\mu \\ -1 & \text{with probability } (1-\mu)\mu \\ 0 & \text{with probability } 1 - 2\mu + \frac{3}{2}\mu^2 \end{cases}$$

to see that $\gamma^{(\mu)}$ is $\left(1 - 2\mu + \frac{3}{2}\mu^2\right)$-bounded of exponent 2, and so Inequality (6) follows from Theorem 2.2.

3.2 Matrices with entries having symmetric distributions

In this subsection, we will prove a singularity bound for an n by n matrix $N_n^{(\mu)}$ for which each entry is a symmetric discrete random variable taking the value 0 with probability $1 - \mu$.

Corollary 3.1. *Let S be a set of complex numbers with cardinality $|S| \leq O(1)$. If $N_n^{(\mu)}$ is an n by n matrix in which each entry is an independent symmetric complex random variable taking values in S and taking the value 0 with probability $1 - \mu$, then*

$$\Pr(N_n^{(\mu)} \text{ is singular}) \leq \begin{cases} (1 - \mu + o(1))^n & \text{for } 0 \leq \mu \leq \frac{1}{2} \\ \left(\frac{2\mu+1}{4} + o(1)\right)^n & \text{for } \frac{1}{2} \leq \mu \leq 1 \\ \left(\sqrt{1 - 2\mu + \frac{3}{2}\mu^2} + o(1)\right)^n & \text{for } 0 \leq \mu \leq 1. \end{cases}$$

In particular, the same upper bounds as in Inequalities (4), (5), *and* (6) *(which are shown in Figure 1) apply to the singularity probability for $N_n^{(\mu)}$.*

Proof. Let α_{ij} be an entry of $N_n^{(\mu)}$. Since α_{ij} is symmetric and takes the value 0 with probability $1-\mu$, we may write $\alpha_{ij} = \gamma_{ij}^{(\mu)}\eta_{ij}$, where $\gamma_{ij}^{(\mu)}$ is an independent copy of $\gamma^{(\mu)}$ as defined in Equation (3) and η_{ij} is a random variable that shares no values with $-\eta_{ij}$. This description of α_{ij} was inspired by [1], and it allows us to condition on η_{ij} and then use the remaining randomness in $\gamma_{ij}^{(\mu)}$ to get a bound on the singularity probability. In particular,

$$\Pr(N_n^{(\mu)} \text{ is singular}) = \sum_{(c_{ij})} \Pr(N_n^{(\mu)} \text{ is singular} | \{\eta_{ij} = c_{ij}\}) \Pr(\{\eta_{ij} = c_{ij}\}),$$

where the sum runs over all (n^2)-tuples $(c_{ij})_{1 \leq i,j \leq n}$ of possible values taken by random variables η_{ij}. Since $\sum_{(c_{ij})} \Pr(\{\eta_{ij} = c_{ij}\}) = 1$, we can complete the proof by proving an exponential bound on $\Pr(N_n^{(\mu)} \text{ is singular} | \{\eta_{ij} = c_{ij}\})$, and we will use Theorem 2.2 for this task.

Consider the random matrix $N_n^{(\mu)}\Big|_{\{\eta_{ij}=c_{ij}\}}$, where the i,j entry is the random variable $c_{ij}\gamma_{ij}^{(\mu)}$ for some constant c_{ij}. Note that the entries of $N_n^{(\mu)}\Big|_{\{\eta_{ij}=c_{ij}\}}$ take values in S, a set with cardinality

$O(1)$, and let ϕ_Q be the map from Lemma 4.1, which lets us pass to the case where $N_n^{(\mu)}\Big|_{\{\eta_{ij}=c_{ij}\}}$ has entries in $\mathbb{Z}/Q\mathbb{Z}$. Defining $\theta_{ij} := 2\pi\phi_Q(c_{ij})$, we compute

$$\begin{aligned}\left|\mathbb{E}e(\phi_Q(c_{ij}\gamma_{ij}^{(\mu)})t)\right| &= |1-\mu+\mu\cos(\theta_{ij}t)| \\ &\leq \begin{cases} 1-\mu+\mu\cos(\theta_{ij}t) & \text{for } 0\leq\mu\leq\frac{1}{2}, \\ \frac{2\mu+1}{4}+(1-\mu)\cos(\theta_{ij}t)+\left(\frac{2\mu-1}{4}\right)\cos(2\theta_{ij}t) & \text{for } \frac{1}{2}\leq\mu\leq 1, \text{ and} \\ \left(1-2\mu+\frac{3}{2}\mu^2+2(1-\mu)\mu\cos(\theta_{ij}t)+\frac{\mu^2}{2}\cos(2\theta_{ij}t)\right)^{1/2} & \text{for } 0\leq\mu\leq 1. \end{cases}\end{aligned}$$

We have thus shown that the entries of $N_n^{(\mu)}\Big|_{\{\eta_{ij}=c_{ij}\}}$ are

$$\begin{aligned}&(1-\mu)\text{-bounded of exponent 1 for } 0\leq\mu\leq\frac{1}{2}, \\ &\left(\frac{2\mu+1}{4}\right)\text{-bounded of exponent 1 for } \frac{1}{2}\leq\mu\leq 1, \text{ and} \\ &\left(1-2\mu+\frac{3}{2}\mu^2\right)\text{-bounded of exponent 2 for } 0\leq\mu\leq 1.\end{aligned}$$

Applying Theorem 2.2 completes the proof. □

Corollary 3.1 is tight for $0\leq\mu\leq\frac{1}{2}$, since the probability of a row of all zeroes occurring is $(1-\mu+o(1))^n$; however, for any specific case, Theorem 2.2 can usually prove better upper bounds than those given by Corollary 3.1.

For example, consider the case of a matrix $M_{\{\pm2,\pm1\},n}^{(\mu)}$ with each entry an independent copy of the symmetric random variable

$$\alpha^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{\mu}{4} \\ -2 & \text{with probability } \frac{\mu}{4} \\ +1 & \text{with probability } \frac{\mu}{4} \\ -1 & \text{with probability } \frac{\mu}{4} \\ 0 & \text{with probability } 1-\mu \end{cases}$$

Corollary 3.2. *For $M_{\{\pm2,\pm1\},n}^{(\mu)}$ as defined above, we have*

$$\Pr(M_{\{\pm2,\pm1\},n}^{(\mu)} \text{ is singular}) \leq \begin{cases} (1-\mu+o(1))^n & \text{for } 0\leq\mu\leq\frac{16}{25} \\ \left(\sqrt{1-2\mu+\frac{5}{4}\mu^2}+o(1)\right)^n & \text{for } 0\leq\mu\leq 1. \end{cases}$$

Proof. By the definition of $\alpha^{(\mu)}$ we have

$$\left|\mathbb{E}e(\alpha^{(\mu)}t)\right| = 1-\mu+\frac{\mu}{2}\cos(2\pi t)+\frac{\mu}{2}\cos(4\pi t), \qquad \text{for } 0\leq\mu\leq\frac{16}{25}$$

(i.e., the right-hand side of the equation above is non-negative for such μ), which proves the first bound.

Also, we have

$$\left|\mathbb{E}e(\alpha^{(\mu)}t)\right|^2 = 1 - 2\mu + \frac{5}{4}\mu^2 + \left(\mu - \frac{3}{4}\mu^2\right)\cos(2\pi t) + \left(\mu - \frac{7}{8}\mu^2\right)\cos(4\pi t) + \frac{\mu^2}{4}\cos(6\pi t) + \frac{\mu^2}{8}\cos(8\pi t)$$

for $0 \le \mu \le 1$, which proves the second bound. □

We also have the following lower bounds for the singularity probability of $M^{(\mu)}_{\{\pm 2,\pm 1\},n}$:

$$(1 - \mu + o(1))^n \qquad \text{(from one row of all zeroes)} \tag{11}$$

$$\left(1 - 2\mu + 5\mu^2/4 + o(1)\right)^n \qquad \text{(from a two-row dependency)} \tag{12}$$

The results of Corollary 3.2 and the corresponding lower bounds are shown in Figure 2, and one should note that the upper bounds are substantially better than those guaranteed by Corollary 3.1.

3.3 Random matrices with entries having arbitrary distributions

A useful feature of the definition of p-bounded of exponent 2 is that it lets one bound the singularity probability of matrices with independent discrete random variables that are asymmetric.

Corollary 3.3. *Let p be a constant such that $0 < p \le 1$ and let $S \subset \mathbb{C}$ be a set with cardinality $|S| \le O(1)$. If N_n is an n by n matrix with independent random entries taking values in S such that for any entry α, we have $\max_x \Pr(\alpha = x) \le p$, then*

$$\Pr(N_n \text{ is singular}) \le (\sqrt{p} + o(1))^n.$$

We will need the following slightly more general corollary in Section 3.4. For a set A and an integer m, we will use the notation $mA := \{\sum_{j=1}^m a_j : a_j \in A\}$ and $A^m := \{\prod_{j=1}^m a_j : a_j \in A\}$.

Corollary 3.4. *Let p be a constant such that $0 < p \le 1$, let $S \subset \mathbb{C}$ be a set with cardinality $|S| \le O(1)$, and let X_n be an n by n matrix with fixed, non-random entries in $n^{o(n)}(S \cup \{-1, 0, 1\})^{O(1)}$. If N_n is an n by n matrix with independent random entries taking values in S such that for any entry α, we have $\max_x \Pr(\alpha = x) \le p$, then*

$$\Pr(X_n + N_n \text{ is singular}) \le (\sqrt{p} + o(1))^n.$$

Note that that Corollary 3.4 implies Corollary 3.3 by taking X_n to be the matrix of all zeroes.

Proof of Corollary 3.4. Let α_{ij} be an entry in N_n. Our goal is to describe α_{ij} in a two-step random process, condition on one of the steps, and then use the randomness in the other step to bound the singularity probability. The conditioning approach is the same as that used in the symmetric case (Corollary 3.1) and was inspired by [1]. The conditioning argument is useful since some entries of the random matrix may take some values with very small probability (i.e. probability less than any constant); recall that while the entries of the random matrix always take values in a fixed set S of cardinality $O(1)$, the distributions of those random variables within S are allowed to vary with n.

Asymptotic Upper and Lower Bounds for $\Pr\left(M^{(\mu)}_{\{\pm 2,\pm 1\},n} \textbf{ is singular}\right)^{1/n}$ **for** $0 \le \mu \le 1$

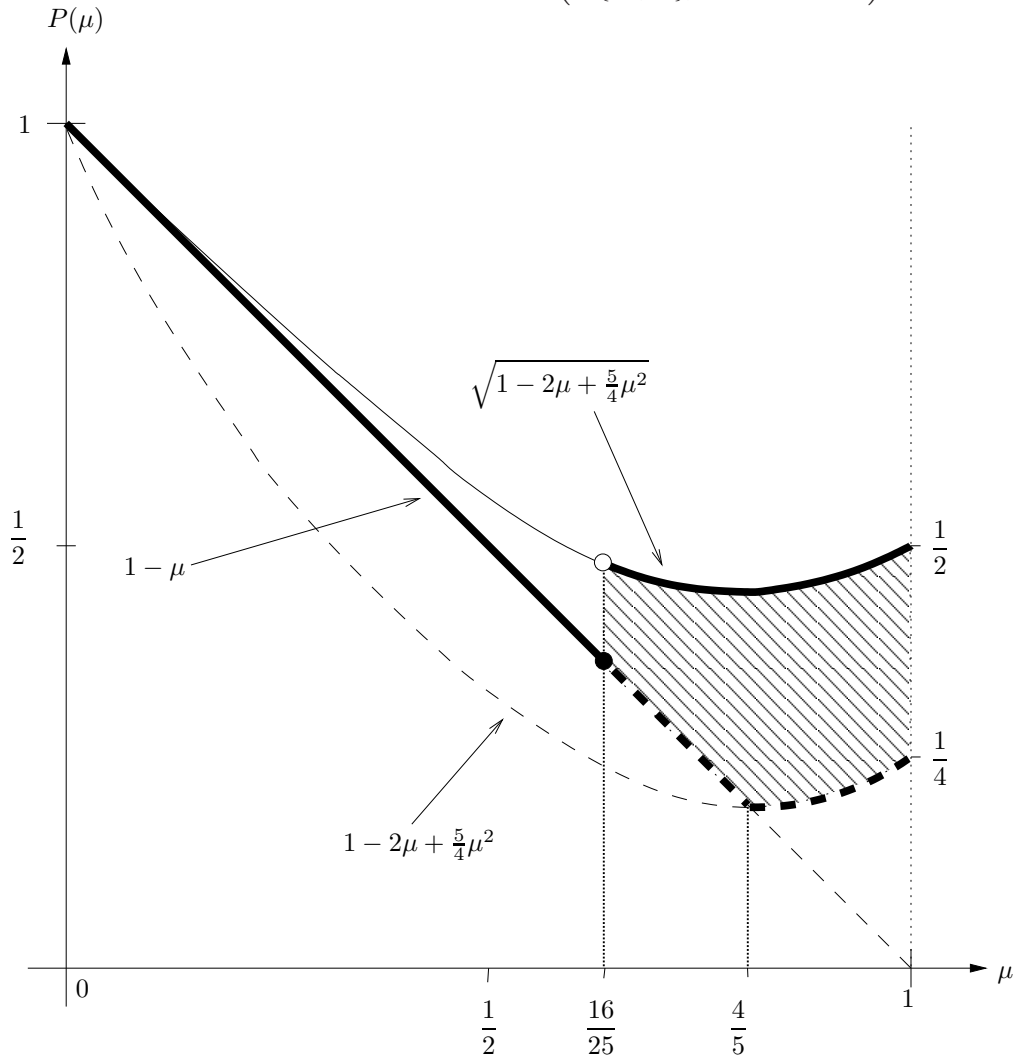

Figure 2: Let $P(\mu) := \lim_{n\to\infty} \Pr\left(M^{(\mu)}_{\{\pm 2,\pm 1\},n} \text{ is singular}\right)^{1/n}$, where $M^{(\mu)}_{\{\pm 2,\pm 1\},n}$ is the n by n matrix with independent random entries taking the value 0 with probability $1-\mu$ and the values $+2, -2, +1, -1$ each with probability $\mu/4$. This figure summarizes the upper bounds on $P(\mu)$ from Corollary 3.2 and the lower bounds from Displays (11) and (12). The best upper bounds (shown in thick solid lines) match the best lower bounds (thick dashed lines) for $0 \le \mu \le \frac{16}{25}$; and it is not hard to improve the upper bound a small amount by finding a bound (of exponent 1) to bridge the discontinuity. One should note that even as stated above, the upper bounds are substantially better than those given by Corollary 3.1 (which are shown in Figure 1). The shaded area represents the gap between the upper and lower bounds.

(Note that making use of Remark 8.5 would provide an alternate way of dealing with entries that take some values with very small probability.)

Say that α_{ij} takes the values $v_1, \dots, v_t$ with probabilities $\varrho_1, \dots, \varrho_t$, respectively, where $\varrho_1 \geq \varrho_2 \geq \cdots \geq \varrho_t$. Define new random variables η_{ijk} such that for some i_0 and i_1, the values taken by η_{ijk} are $v_{i_0}, v_{i_0+1}, \dots, v_{i_0+i_1}$ with corresponding probabilities $\varrho_{i_0}/p_k, \varrho_{i_0+1}/p_k, \dots, \varrho_{i_0+i_1}/p_k$, where $p_k := \sum_{i=1}^{i_1} \varrho_{i_0+i}$. Thus, we can write

$$\alpha_{ij} = \begin{cases} \eta_{ij1} & \text{with probability } p_1 \\ \eta_{ij2} & \text{with probability } p_2 \\ \vdots & \vdots \\ \eta_{ij\ell} & \text{with probability } p_\ell. \end{cases} \tag{13}$$

Furthermore, the η_{ijk} can be constructed so that $p_k \leq p$ for every k, so that $p/2 \leq p_k$ for $1 \leq k \leq \ell - 1$, and so that no two η_{ijk} with different k's ever take the same value.

There are two cases to consider for the technical reason that p_ℓ is not necessarily bounded below by a constant. Let $\epsilon > 0$ be a very small constant, so for example $p/2 > \epsilon$. Case 1 is when $\epsilon \leq p_\ell$, and in this case each p_k is bounded below by ϵ and above by p. We will consider Case 1 first and then discuss the small changes needed to deal with Case 2.

As in the proof of Corollary 3.1, we will condition on the values taken by the η_{ijk} in order to prove a bound on the singularity probability. We have that

$$\Pr(X_n + N_n \text{ is singular}) = \sum_{(c_{ijk})} \Pr(X_n + N_n \text{ is singular} | \{\eta_{ijk} = c_{ijk}\}) \Pr(\{\eta_{ijk} = c_{ijk}\}),$$

where the sum runs over all possible values (c_{ijk}) that the η_{ijk} can take. Thus, it is sufficient to prove a bound on the singularity probability for the random matrix $X_n + N_n\Big|_{\{\eta_{ijk}=c_{ijk}\}}$ which has random entries

$$x_{ij} + \widetilde{\alpha}_{ij} = \begin{cases} x_{ij} + c_{ij1} & \text{with probability } p_1 \\ x_{ij} + c_{ij2} & \text{with probability } p_2 \\ \vdots & \vdots \\ x_{ij} + c_{ij\ell} & \text{with probability } p_\ell, \end{cases}$$

where x_{ij} and the c_{ijk} are constants.

Note the entries of $X_n + N_n\Big|_{\{\eta_{ijk}=c_{ijk}\}}$ take values in $n^{o(n)} \left(S \cup \{-1, 0, 1\}\right)^{O(1)}$, a generalized arithmetic progression with rank $O(1)$ and cardinality at most $n^{o(n)}$, and let ϕ_Q be the map from Lemma 4.1, which lets us pass to the case where $X_n + N_n\Big|_{\{\eta_{ijk}=c_{ijk}\}}$ has entries in $\mathbb{Z}/Q\mathbb{Z}$. Defining $\theta_{ijk} := 2\pi\phi_Q(c_{ijk})$ and letting $\widetilde{\alpha}'_{ij}$ be an i.i.d. copy of $\widetilde{\alpha}_{ij}$, we compute

$$\begin{aligned} |\mathbb{E}e(\phi_Q(x_{ij} + \widetilde{\alpha}_{ij})t)|^2 &= \mathbb{E}e\left(\phi_Q(x_{ij} + \widetilde{\alpha}_{ij} - x_{ij} - \widetilde{\alpha}'_{ij})t\right) = \mathbb{E}e\left(\phi_Q(\widetilde{\alpha}_{ij} - \widetilde{\alpha}'_{ij})t\right) \\ &= \sum_{k=1}^{\ell} p_k^2 + 2 \sum_{1 \leq k_1 < k_2 \leq \ell} p_{k_1} p_{k_2} \cos((\theta_{ijk_1} - \theta_{ijk_2})t). \end{aligned}$$

Thus, $x_{ij} + \widetilde{\alpha}_{ij}$ is $\left(\sum_{k=1}^{\ell} p_k^2\right)$-bounded of exponent 2 (using the constant $q = \epsilon^2$ in Definition 2.1, so q does not depend on n). Given that $0 < p_k \leq p$ for every k, it is not hard to show that $\sum_{k=1}^{\ell} p_k^2 \leq p < p + \epsilon$, and so from Definition 2.1, we see that the collection $\{x_{ij} + \widetilde{\alpha}_{ij} : \widetilde{\alpha}_{ij}$ has corresponding probability $p_\ell \geq \epsilon\}$ is $(p+\epsilon)$-bounded of exponent 2. We are thus finished with Case 1.

Case 2 is when the decomposition of α_{ij} given in Equation (13) has $p_\ell < \epsilon$. In this case we need to modify Equation (13) slightly, deleting $\eta_{ij\ell}$ and replacing $\eta_{ij(\ell-1)}$ with a new variable $\eta'_{ij(\ell-1)}$ that takes all the values previously taken by $\eta_{ij\ell}$ and by $\eta_{ij(\ell-1)}$ with the appropriate probabilities. Thus, in Case 2, we have that $p/2 \leq p_k < p + \epsilon$ for all $1 \leq k \leq \ell - 1$, showing that each p_k is bounded below by a constant and is bounded above by $p + \epsilon$ (here we are using $p_{\ell-1}$ to denote the probability that α_{ij} draws a value from the random variable $\eta'_{ij(\ell-1)}$).

For Case 2, we use exactly the same reasoning as in Case 1 above to show that such entries of $X_n + N_n\Big|_{\{\eta_{ijk}=c_{ijk}\}}$ are $\left(\sum_{k=1}^{\ell-1} p_k^2\right)$-bounded of exponent 2 (using the constant $q = \epsilon^2 < p^2/4$ in Definition 2.1, so q does not depend on n). Noting that $\sum_{k=1}^{\ell-1} p_k^2 < p + \epsilon$ and using Definition 2.1, we see that the collection $\{x_{ij} + \widetilde{\alpha}_{ij} : \widetilde{\alpha}_{ij}$ has corresponding probability $p_\ell < \epsilon\}$ is $(p+\epsilon)$-bounded of exponent 2.

Combining Case 1 and Case 2, we have that the collection $\{x_{ij} + \widetilde{\alpha}_{ij}\}$ is $(p+\epsilon)$-bounded of exponent 2, and so by and by Theorem 2.2 we have that $\Pr(X_n + N_n\Big|_{\{\eta_{ijk}=c_{ijk}\}}$ is singular$) \leq (\sqrt{p+\epsilon} + o(1))^n$.

The constant $\epsilon > 0$ was chosen arbitrarily, and so letting ϵ tend to zero, we get that

$$\Pr(X_n + N_n \text{ is singular} | \{\eta_{ijk} = c_{ijk}\}) \leq (\sqrt{p} + o(1))^n .$$

□

3.4 Partially random matrices

In this subsection, we prove a bound on the singularity probability for partly random matrices where many rows are deterministic.

Corollary 3.5. *Let p be a real constant between 0 and 1, let K be a large positive constant, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq K$. Let $N_{\mathfrak{f},n}$ be an n by n matrix in which $\mathfrak{f}$ rows contain fixed, non-random integers between $-K$ and K and where the other rows contain entries that are independent random variables taking values in S. If $\mathfrak{f} \leq o(n)$, if the $\mathfrak{f}$ fixed rows are linearly independent, and if for every random entry α, we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_{\mathfrak{f},n} \text{ is singular}) \leq (\sqrt{p} + o(1))^{n-\mathfrak{f}} .$$

Corollary 3.5 applies to partly random matrices with $\mathfrak{f} = o(n)$ fixed, non-random rows containing integers bounded by a constant and with random entries taking at most $O(1)$ values in the complex numbers. Corollary 1.2, on the other hand, holds with the fixed entries also allowed to take values in the complex numbers and gives a sligtly better bound, but additionally requires $\mathfrak{f} \leq O(\ln n)$ (which is far smaller in general than $o(n)$). Proving Corollary 1.2 requires mirroring the entire argument used to prove the main theorem (Theorem 2.2) in the case where $\mathfrak{f}$ rows contain fixed, non-random entires, and we discuss this argument in Section 9. Proving Corollary 3.5, however, can be done directly from Theorem 2.2, as we will show below. First, we will state a generalization of Corollary 3.5.

Corollary 3.6. *Let p be a real constant between 0 and 1, let K be a large positive constant, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq K$. Let $N_{\mathfrak{f},n}$ be an n by n matrix in which $\mathfrak{f}$ rows contain fixed, non-random integers between $-K$ and K and where the other rows contain entries that are independent random variables taking values in S. If $\mathfrak{f} \leq o(n)$, if the fixed rows have co-rank k, and if for every random entry α, we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_{\mathfrak{f},n} \text{ has co-rank } > k) \leq (\sqrt{p} + o(1))^{n-\mathfrak{f}}.$$

To obtain Corollary 3.6 from Corollary 3.5, find a collection $\mathcal{C}$ of $\mathfrak{f} - k$ linearly independent rows among the deterministic rows. Replace the rest of the deterministic rows with a collection $\mathcal{C}'$ of rows containing integer values between $-K$ and K such that $\mathcal{C}'$ is linearly independent from $\mathcal{C}$. Finally, apply Corollary 3.5 to the new partially random matrix whose deterministic rows are from $\mathcal{C} \cup \mathcal{C}'$, thus proving Corollary 3.6.

Proof of Corollary 3.5. By reordering the rows and columns, we may write

$$N_{\mathfrak{f},n} = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right),$$

where A is an $\mathfrak{f}$ by $\mathfrak{f}$ non-random invertible matrix, B is an $\mathfrak{f}$ by $n - \mathfrak{f}$ non-random matrix, C is an $n - \mathfrak{f}$ by $\mathfrak{f}$ random matrix, and D is an $n - \mathfrak{f}$ by $n - \mathfrak{f}$ random matrix. Note that $N_{\mathfrak{f},n}$ is singular if and only if there exists a vector $\mathbf{v}$ such that $N_{\mathfrak{f},n}\mathbf{v} = 0$. Let $\mathbf{v}_1$ be the first $\mathfrak{f}$ coordinates of $\mathbf{v}$ and let $\mathbf{v}_2$ be the remaining $n - \mathfrak{f}$ coordinates. Then $N_{\mathfrak{f},n}\mathbf{v} = 0$ if and only if

$$\begin{cases} A\mathbf{v}_1 + B\mathbf{v}_2 = 0, \text{ and} \\ C\mathbf{v}_1 + D\mathbf{v}_2 = 0. \end{cases}$$

Since A is invertible, these two equations are satisfied if and only if $(-CA^{-1}B + D)\mathbf{v}_2 = 0$, that is, if and only if the random matrix $-CA^{-1}B + D$ is singular.

We want to show that every entry that can appear in $-CA^{-1}B$ is an element of $n^{o(n)}\left(S \cup \{-1, 0, 1\}\right)^{O(1)}$. By the cofactor formula for A^{-1}, we know that the i, j entry of A^{-1} is $(-1)^{i+j}(\det A_{ij})/\det A$, where A_{ij} is the $\mathfrak{f} - 1$ by $\mathfrak{f} - 1$ matrix formed by deleting the i-th row and j-th column of A. Thus, $A^{-1} = \frac{1}{\det A}\widetilde{A}$, where the i, j entry of $\widetilde{A}$ is $(-1)^{i+j} \det A_{ij}$. By the volume formula for the determinant, we know that $|\det A|$ is at most the product of the lengths of the row vectors of A; and thus $|\det A| \leq n^{o(n)}$ (here we need that A has integer entries between $-K$ and K, where K is a constant, and that $\mathfrak{f} \leq o(n)$). Similarly, we have $|\det A_{ij}| \leq n^{o(n)}$. Every entry of $\widetilde{A}$ is thus in $n^{o(n)}\{-1, 0, 1\}$, every entry of C is in S, and every entry of B is in $O(1)\{-1, 0, 1\}$; thus, every entry of $-C\widetilde{A}B$ is an element of $n^{o(n)}(S \cup \{-1, 0, 1\})$.

Conditioning on the values taken by all the entries in C, we have

$$\begin{aligned} \Pr(N_{\mathfrak{f},n} \text{ is singular}) &= \Pr(-CA^{-1}B + D \text{ is singular}) \\ &= \sum_{(c_{ij})} \Pr(-CA^{-1}B + D \text{ is singular} | C = (c_{ij})) \Pr(C = (c_{ij})), \end{aligned} \tag{14}$$

where the sum runs over all possible matrices (c_{ij}) that C can produce. Considering the entries in $C = (c_{ij})$ to be fixed (note that A and B are fixed by assumption), we now need to bound

$$\Pr(-(c_{ij})A^{-1}B + D \text{ is singular}) = \Pr(-(c_{ij})\widetilde{A}B + (\det A)D \text{ is singular}).$$

Note that every entry of $-(c_{ij})\widetilde{A}B$ is an element of $n^{o(n)}\left(S \cup \{-1,0,1\}\right)^{O(1)}$ and that the random matrix $(\det A)D$ has entries that take values in the fixed set $\{(\det A)s : s \in S\}$ having cardinality $O(1)$. Thus, by Corollary 3.4, we have that

$$\Pr(-(c_{ij})\widetilde{A}B + (\det A)D \text{ is singular}) \leq (\sqrt{p} + o(1))^{n-\mathfrak{f}}.$$

Plugging this bound back into Equation (14) completes the proof. □

3.5 Integer matrices and rational eigenvalues

Let η_k be the random variable taking the values $-k, -k+1, \ldots, k-1, k$ each with equal probability, and let M_n be the n by n matrix where each entry is an independent copy of η_k. In [7], Martin and Wong show that for any $\epsilon > 0$,

$$\Pr(M_n \text{ has a rational eigenvalue}) \leq \frac{c(n,\epsilon)}{k^{1-\epsilon}},$$

where $c(n,\epsilon)$ is a constant depending on n and ϵ. (One goal in [7] is to study this bound as k goes to ∞ while n is fixed, which is why $c(n,\epsilon)$ is allowed to depend on n.)

Below, we prove a similar result for random integer matrices with entries between $-k$ and k (with k fixed), where we allow each entry to have a different (independent) distribution and we also allow the distributions to be very general.

Corollary 3.7. *Fix a positive integer k, and let $M_{k,n}$ be a random integer matrix with independent entries, each of which takes values in the set $\{-k, -k+1, \ldots, k-1, k\}$. Let c be a constant such that for every entry α, we have $\max_{-k \leq x \leq k} \Pr(\alpha = x) \leq c/k$. Then*

$$\Pr(M_{k,n} \textit{ has a rational eigenvalue}) \leq \left(\frac{c}{k} + o(1)\right)^{n/2},$$

where the $o(1)$ term goes to zero as n goes to ∞.

For example, in the case where each independent entry has the uniform distribution on $\{-k, -k+1, \ldots, k-1, k\}$ (as in [7]), one can set $c = 1/2$ in the corollary above.

Proof. The proof given below follows the same outline as the main theorem of [7], with Corollary 1.2 replacing an appeal to [7, Lemma 1].

The characteristic polynomial for $M_{k,n}$ is monic with integer coefficients, and thus the only possible rational eigenvalues are integers (by the rational roots theorem). Every eigenvalue of $M_{k,n}$ has absolute value at most nk (see [7, Lemma 4]); thus, the only possible integer eigenvalues are between $-nk$ and nk.

The matrix $M_{k,n}$ has λ as an eigenvalue if and only if $M_{k,n} - \lambda I$ is singular (where I is the n by n identity matrix). By Corollary 1.2 (with $\mathfrak{f} = 0$), we have

$$\Pr(M_{k,n} - \lambda I \text{ is singular}) \leq \left(\sqrt{\frac{c}{k}} + o(1)\right)^{n}.$$

Using the union bound, we have

$$\begin{aligned}\Pr(M_{k,n} \text{ has a rational eigenvalue}) &= \Pr(M_{k,n} - \lambda I \text{ is singular, for some } \lambda \in \{-nk, \dots, nk\}) \\ &\leq \sum_{\lambda=-nk}^{nk} \Pr(M_{k,n} - \lambda I \text{ is singular}) \\ &\leq (2nk+1)\left(\sqrt{\frac{c}{k}} + o(1)\right)^n \\ &\leq \left(\frac{c}{k} + o(1)\right)^{n/2}.\end{aligned}$$

□

4 Random matrices with complex entries: A reduction technique

The original work on discrete random matrices in [5, 4, 10, 11] is concerned with matrices having integer entries, which can also be viewed as matrices with entries in $\mathbb{Z}/Q\mathbb{Z}$ where Q is a very large prime. In this section we show that one can pass from a (random) matrix with entries in $\mathbb{C}$ to one with entries in $\mathbb{Z}/Q\mathbb{Z}$ where Q is an arbitrarily large prime number, all without affecting the probability that the determinant is zero, thanks to the following lemma.

Lemma 4.1 ([13])**.** *Let S be a finite subset of $\mathbb{C}$. There exist infinitely many primes Q such that there is a ring homomorphism $\phi_Q : \mathbb{Z}[S] \to \mathbb{Z}/Q\mathbb{Z}$ satisfying the following two properties:*

(i) the map ϕ_Q is injective on S, and

(ii) for any n by n matrix $(s_{ij})_{1\leq i,j\leq n}$ with entries $s_{ij} \in S$, we have

$$\det\Big((s_{ij})_{1\leq i,j\leq n}\Big) = 0 \quad \textit{if and only if} \quad \det\Big((\phi_Q(s_{ij}))_{1\leq i,j\leq n}\Big) = 0.$$

In order to apply this lemma, let us point out that the proof of Theorem 2.2, which is discussed in Sections 5 through 8, works exclusively in $\mathbb{Z}/Q\mathbb{Z}$; though at various points, it is necessary to assume Q is extremely large with respect to n and various constants. For this paper, S will be the set of all possible values taken by the random variables α_{jk}. Recall that by assumption, $|S| \leq n^{o(n)}$, so in particular, S is finite.

Remark 4.2 (On the size of Q). When we apply Lemma 4.1, we will take $Q > \exp(\exp(Cn))$ for some constant C in order for Freiman-type theorems such as [11, Theorem 6.3] (which is restated in Theorem 8.1 below) to apply, and we will also choose Q large enough so that the integral approximation in Inequality (44) holds and so that Q is large with respect to various constants. One should note that while Q can be taken arbitrarily large with respect to n, we cannot choose Q so that it is arbitrarily large with respect to $\phi_Q(s)$ for all $s \in S$, where S is the set of all values that could appear in the given random matrix. For example, if $\sqrt{2} \in S$, then the smallest positive integer representative for $\phi_Q(\sqrt{2})$ must be larger than $\sqrt{Q}$ (since $(\phi_Q(\sqrt{2}))^2 = 2$ in $\mathbb{Z}/Q\mathbb{Z}$). Finally, if we were in a situation where $S \subset \mathbb{Q}$, then we could avoid using Lemma 4.1 altogether by clearing denominators to pass to $\mathbb{Z}$ and then take $Q \approx \exp(\exp(Cn))$, as is done in [11].

Lemma 4.1 is a corollary of the main theorem of [13] and its proof is given in detail in [13, Section 6]. The paper [13] also contains further applications of the method used to prove Lemma 4.1, for example proving a sum-product result for the complex numbers and proving a Szemerédi-Trotter-type result for the complex numbers, where the applications follow from the analogous results for $\mathbb{Z}/Q$ where Q is a prime (see [3]). The results in [13], including Lemma 4.1, all go through with the complex numbers being replaced by any characteristic zero integral domain. Thus, the results stated in Sections 1, 2, and 3 above for the complex numbers $\mathbb{C}$ also all go through with $\mathbb{C}$ replaced by any characteristic zero integral domain. For example, Corollary 3.3 becomes

Corollary 4.3. *Let p be a constant such that $0 < p \leq 1$ and let D be a characteristic zero integral domain. Let $S \subset D$ have cardinality $|S| \leq O(1)$. If N_n is an n by n matrix with independent random entries, each taking values in S, such that for every entry α, we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_n \text{ is singular}) \leq (\sqrt{p} + o(1))^n.$$

5 Proof of the main theorem (Theorem 2.2)

The proof of Theorem 2.2 very closely follows the proof of [11, Theorem 1.2]. Our goal is to highlight the changes that need to be made to generalize the proof in [11] so that it proves Theorem 2.2. A reader interested in the details of the proof of Theorem 2.2 should read this paper alongside of [11]. Throughout the proof, we will assume that n is sufficiently large, and we will allow constants hidden in the $o(\cdot)$ and $O(\cdot)$ notation to depend on the constants $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, p, q, r, c_{\text{MedDim}}, c_{\text{LgDim}}, c_{\text{LO}}$, and c_m. The constants $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2$ should be considered very small, and, in fact, we will let them tend to zero to prove the full strength of Theorem 2.2. The constants $p, q, r, c_{\text{MedDim}}, c_{\text{LgDim}}, c_{\text{LO}}$, and c_m can be thought of as absolute, except possibly for depending on each other.

5.1 Definitions and preliminaries

Given an n by n matrix N_n with entries α_{ij}, we assume that the collection of independent random variables $\{\alpha_{ij}\}_{1 \leq i,j \leq n}$ is p-bounded of exponent r for some fixed constants p, q, and r (here, q is the constant from Definition 2.1 which is independent of n). We also assume that each α_{ij} takes at most $n^{o(n)}$ distinct values. Using Lemma 4.1, we may assume without loss of generality that each α_{ij} takes values in $\mathbb{Z}/Q\mathbb{Z}$ for some very large prime Q. The entirety of the proof will take place over the field $\mathbb{Z}/Q\mathbb{Z}$, and so terminology such as "linearly independent", "span", "dimension", "rank" and so forth will always be with respect to the field $\mathbb{Z}/Q\mathbb{Z}$.

Let $X_i := (\alpha_{i,1}, \ldots, \alpha_{i,n})$ denote the i-th row of N_n. We note that N_n has determinant zero if and only if there is a linear dependency among the rows X_i. It has been shown (see [10, Lemma 5.1] and also [4]) that the dominant contribution to the singularity probability comes from the X_i spanning a hyperplane (of dimension $n-1$). In particular,

$$\Pr\left(N_n \text{ is singular}\right) = p^{-o(n)} \sum_{\substack{V \text{ a non-trivial} \\ \text{hyperplane in } (\mathbb{Z}/Q\mathbb{Z})^n}} \Pr(A_V), \tag{15}$$

where A_V denotes the event that $X_1, \ldots, X_n$ span V, and *non-trivial* means that V contains the origin, V is spanned by vectors in S^n (where S is the set of all possible values that can occur in N_n), and $\Pr(X_i \in V) > 0$ for all i.

As in [11], we will divide the non-trivial hyperplanes into n^2 classes, since it is then sufficient to show that the sum of $\Pr(A_V)$ over all V in a particular class is at most $(p^{1/r} + o(1))^n$.

Definition 5.1 (combinatorial dimension)**.** Let $\mathcal{D} := \left\{\frac{a}{n} : 0 \le a \le n^2, a \in \mathbb{Z}\right\}$. For any $d_\pm \in \mathcal{D}$ such that $d_\pm \ge \frac{1}{n}$, we define the *combinatorial Grassmannian* $\mathrm{Gr}(d_\pm)$ to be the set of all non-trivial hyperplanes V in $(\mathbb{Z}/Q\mathbb{Z})^n$ such that

$$p^{n-d_\pm+1/n} < \max_{1\le i\le n} \Pr(X_i \in V) \le p^{n-d_\pm}. \tag{16}$$

For $d_\pm = 0$, we define $\mathrm{Gr}(0)$ to be the set of all non-trivial hyperplanes such that

$$\max_{1\le i\le n} \Pr(X_i \in V) \le p^n.$$

We will refer to $d_\pm$ as the *combinatorial dimension* of V.

Note that $\mathrm{Gr}(d_\pm) = \emptyset$ for $d_\pm \ge n - 1 + 1/n$ (by Lemma B.1). We will consider hyperplanes V with combinatorial dimension in three main regions: $d_\pm$ small, $d_\pm$ medium-sized, and $d_\pm$ large. The two lemmas and the proposition below suffice to prove Theorem 2.2.

Lemma 5.2 (Small combinatorial dimension, [4], [10], [11])**.** *For any $\delta > 0$ we have*

$$\sum_{d_\pm\in\mathcal{D} \text{ s.t. } p^{n-d_\pm}\le\delta^n} \quad \sum_{V\in\mathrm{Gr}(d_\pm)} \Pr(A_V) \le n\delta^n.$$

In proving Theorem 2.2, we will take $\delta = (p + c_{\mathrm{MedDim}}\epsilon_0)^{1/r}$ to take care of all small $d_\pm$ not covered by Proposition 5.4 below.

Proof. The reasoning here is the same as in [11, Lemma 2.3], making use of fact that $\Pr(X_i \in V) \le \max_{1\le i\le n} \Pr(X_i \in V) \le p^{n-d_\pm} \le \delta^n$. In particular,

$$\Pr(A_V) \le \sum_{i=1}^{n} \Pr\left(\{X_j\}_{1\le j\le n} \setminus \{X_i\} \text{ spans } V\right) \Pr(X_i \in V),$$

which completes the proof since the summing the right-hand side over all V is at most $n \max_i \Pr(X_i \in V)$ (note that an instance of the vectors $\{X_j\}_{1\le j\le n} \setminus \{X_i\}$ can span at most one hyperplane). □

Lemma 5.3 (Large combinatorial dimension, [4],[10],[11])**.** *We have*

$$\sum_{d_\pm\in\mathcal{D} \text{ s.t. } \frac{c_{\mathrm{LgDim}}}{\sqrt{n}}\le p^{n-d_\pm}} \quad \sum_{V\in\mathrm{Gr}(d_\pm)} \Pr(A_V) \le (p + o(1))^n$$

Here we choose the constant c_{LgDim} so that $c_{\mathrm{LgDim}} \ge c_{\mathrm{LO}} p^{-1/n}\sqrt{\frac{2r}{q}}$, where c_{LO} is the constant from the Littlewood-Offord inequality (see Lemma A.1 in Appendix A) and q is the constant from Definition 2.1.

Proof. Our proof is essentially the same as [11, Lemma 2.4]. Fix $V \in \mathrm{Gr}(d_\pm)$, where $\frac{c_{\mathrm{LgDim}}}{\sqrt{n}} \leq p^{n-d_\pm}$. Let $i_{\max}$ be an index such that $\Pr(X_{i_{\max}} \in V) = \max_{1 \leq i \leq n} \Pr(X_i \in V)$. By assumption,

$$\Pr(X_{i_{\max}} \in V) \geq p^{n-d_\pm+1/n} \geq \frac{c_{\mathrm{LgDim}}}{\sqrt{n}} p^{1/n} \geq c_{\mathrm{LO}} \sqrt{\frac{2r}{qn}}.$$

Noting that $X_{i_{\max}} \in V$ if and only if $X_{i_{\max}}$ is orthogonal to the normal vector for V, we have by Lemma A.1 that

$$\Pr(X_{i_{\max}} \in V) \leq c_{\mathrm{LO}} \sqrt{\frac{r}{qk}},$$

where k is the number of nonzero coordinates in the normal vector to V. Combining the two inequalities above shows that $k \leq n/2$.

Thus, we have

$$\sum_{d_\pm \in \mathcal{D} \text{ s.t. } \frac{c_{\mathrm{LgDim}}}{\sqrt{n}} \leq p^{n-d_\pm}} \quad \sum_{V \in \mathrm{Gr}(d_\pm)} \Pr(A_V) \leq \Pr\left(\left\{\begin{array}{l}\text{there exists a vector } \mathbf{v} \text{ with at} \\ \text{most } n/2 \text{ nonzero coordinates} \\ \text{such that } N_n \cdot \mathbf{v} = 0\end{array}\right\}\right)$$

$$\leq (p + o(1))^n \qquad \text{(by Lemma A.2)}$$

(Lemma A.2 is a natural generalization of [4, Section 3.1]; see also [6], [10, Lemma 5.1], and [2, Lemma 14.10].) □

Proposition 5.4 (Medium combinatorial dimension estimate)**.** *Let $0 < \epsilon_0$ be a constant much smaller than 1, and let $d_\pm \in \mathcal{D}$ be such that $(p + c_{\mathrm{MedDim}}\epsilon_0)^{n/r} < p^{n-d_\pm} < \frac{c_{\mathrm{LgDim}}}{\sqrt{n}}$. Then*

$$\sum_{V \in \mathrm{Gr}(d_\pm)} \Pr(A_V) \leq o(1)^n.$$

Here we choose the constant c_{MedDim} so that $c_{\mathrm{MedDim}} > \left(\frac{1}{100} + c_m\right)$, where c_m is some absolute constant such that $0 < c_m < 1$ (the $\frac{1}{100}$ here comes from $\underline{\mu}$ as defined in Section 5.2 below; in [11], it happens that the constant c_m is also taken to be $\frac{1}{100}$).

To prove Theorem 2.2, we can simply combine Lemma 5.2 with $\delta = (p + c_{\mathrm{MedDim}}\epsilon_0)^{1/r}$, Lemma 5.3, and Proposition 5.4. Thus, proving Proposition 5.4 will complete the proof of Theorem 2.2. To prove Proposition 5.4, as in [11, Proposition 2.5], we will separate hyperplanes V of medium combinatorial dimension into two classes, which we will call *exceptional* and *unexceptional* (see Definition 5.5). See [11, Section 3] for motivation. The unexceptional case will be proved in the remainder of this section, and the exceptional case will be proved in Sections 6, 7, and 8.

The results in [10] and [4] were derived using the ideas that we will use for the unexceptional medium combinatorial dimension case. The idea of considering the exceptional case separately in [11] (and using tools from additive combinatorics in the exceptional case) is what lead to the improvement of Inequality (1), which gives a bound of asymptotically $\left(\frac{3}{4}\right)^n$, over the $.999^n$ bound in [4].

5.2 Proof of the medium combinatorial dimension

Before defining exceptional and unexceptional hyperplanes, we will need some new notation. By assumption, the collection of random variables $\{\alpha_{ij}\}_{1\le i,j\le n}$ is p-bounded of exponent r with a constant $\mu = 1-p$, with random variables $\beta_{ij}^{(\mu)}$ corresponding to each α_{ij}, and with a constant $0 < q \le p$ (see Definition 2.1). We also define a constant slightly smaller than μ, namely $\underline{\mu} := \mu - \dfrac{\epsilon_0}{100}$. We will let $Y_i := (y_{i,1}, \dots, y_{i,n}) := (\beta_{i,1}^{(\mu)}, \dots, \beta_{i,n}^{(\mu)})$ denote another row vector that corresponds to the row vector X_i ($\beta_{i,j}^{(\mu)}$ comes from the definition of p-bounded of exponent r). Also, we will let

$$Z_{i,k}^* := (\overbrace{0, \dots, 0}^{\substack{k_{\text{start}}-1 \\ \text{zeroes}}}, y_{i,k_{\text{start}}}, \dots, y_{i,k_{\text{end}}}, \overbrace{0, \dots, 0}^{\substack{n-k_{\text{end}} \\ \text{zeroes}}}), \tag{17}$$

where $k_{\text{start}} := \left\lfloor (k-1)\frac{n}{r} \right\rfloor + 1$ and $k_{\text{end}} := \left\lfloor k\frac{n}{r} \right\rfloor$. The vector $Z_{i,k}^*$ can be thought of as the k-th segment of Y_i (out of r roughly equal segments). Note that Y_i and $Z_{i,k}^*$ are both defined using $\underline{\mu} := \mu - \frac{\epsilon_0}{100}$, not μ. Finally, let ϵ_1 be a positive constant that is small with respect to ϵ_0, c_m, and r.

Definition 5.5 (exceptional and unexceptional)**.** Consider a hyperplane V of medium combinatorial dimension (that is, $d_\pm$ satisfies the condition in Proposition 5.4). We say V is *unexceptional* if there exists an i_0 where $1 \le i_0 \le n$ and there exists a k_0 where $1 \le k_0 \le r$ such that

$$\max_{1\le j\le n} \{\Pr(X_j \in V)\} < \epsilon_1 \Pr(Z_{i_0,k_0}^* \in V).$$

We say V is *exceptional* if for every i where $1 \le i \le n$ and for every k where $1 \le k \le r$ we have

$$\epsilon_1 \Pr(Z_{i,k}^* \in V) \le \max_{1\le j\le n} \{\Pr(X_j \in V)\}. \tag{18}$$

In particular, there exists $i_{\max}$ such that $\Pr(X_{i_{\max}} \in V) = \max_{1\le j\le n}\{\Pr(X_j \in V)\}$; and so if V is exceptional, then

$$\epsilon_1 \Pr(Z_{i_{\max},k}^* \in V) \le \Pr(X_{i_{\max}} \in V) \qquad \text{for every } k. \tag{19}$$

We will refer to $X_{i_{\max}}$ as the *exceptional row*.

Inequality (10) following Definition 2.1 can be used to give another relationship between $\Pr(Z_{i_{\max},k}^* \in V)$ and $\Pr(X_{i_{\max}} \in V)$ that, together with Inequality (19), will be of critical importance in Section 7.

Proposition 5.4 follows from the two lemmas below, so long as ϵ_1 is chosen suitably small with respect to ϵ_0, c_m, and r.

Lemma 5.6 (Unexceptional space estimate)**.** *We have*

$$\sum_{V\in \mathrm{Gr}(d_\pm):\ V \text{ is unexceptional}} \Pr(A_V) \le p^{-o(n)} 2^n \epsilon_1^{c_m \epsilon_0 n/r}.$$

Lemma 5.7 (Exceptional space estimate)**.** *We have*

$$\sum_{V\in \mathrm{Gr}(d_\pm):\ V \text{ is exceptional}} \Pr(A_V) \le n^{-\frac{n}{2}+o(n)}.$$

We will prove Lemma 5.6 in Section 5.3, and we will prove Lemma 5.7 in Section 6.

5.3 The unexceptional medium combinatorial dimension case

The general idea for the case of an unexceptional hyperplane V is to replace some of the rows X_i in the matrix N_n with rows that concentrate more sharply on the subspace V. In the case where the exponent $r = 1$, replacing a row X_i with $Y_i := (\beta_{i,1}^{(\underline{\mu})}, \ldots, \beta_{i,n}^{(\underline{\mu})})$ is successful; however, in the exponent $r = 2$ case, for example, replacing the entire row results in a bound that is off by an exponential factor. We solve this problem by replacing X_i with only half of Y_i (with the other half of the entries being zero). This idea easily extends to any integer $r \geq 2$ and is the motivation for defining the vectors $Z_{i,k}^*$ to have all zeros except for roughly n/r coordinates, as is done in Equation (17). The basic utility of Z_{i_0,k_0}^* (from Definition 5.5) is that it concentrates more sharply on the unexceptional subspace V than the vector X_i for any i.

Let Z_{i_0,k_0}^* be the vector from the definition of unexceptional (Definition 5.5) such that $\Pr(X_i \in V) < \epsilon_1 \Pr(Z_{i_0,k_0}^* \in V)$ for every i, and set $Z := Z_{i_0,j_0}^*$. Let m be the closest integer to $\frac{c_m \epsilon_0 n}{r}$, where c_m is a small positive absolute constant (for example, in [11], c_m is taken to be $\frac{1}{100}$). Finally, let $Z_1, \ldots, Z_m$ be copies of Z, independent of each other and of $X_1, \ldots, X_n$.

Lemma 5.8 (see Lemma 4.4 in [11])**.** *Let $B_{V,m}$ be the event that $Z_1, \ldots, Z_m$ are linearly independent and lie in V. Then,*

$$\Pr(B_{V,m}) \geq p^{o(n)} \left(\frac{\max_{1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1} \right)^m$$

Proof. The argument follows the same reasoning as [11, Lemma 4.4], however, the quantity $2^{d_\pm - n}$ in [11] should be replaced by $\max_{1 \leq i \leq n} \Pr(X_i \in V)$. Details are provided in Appendix B. □

To conclude the proof of Lemma 5.6, we follow the "row-swapping" argument at the end of [11, Section 4], with the small change of bounding $\Pr(X_i \in V)$ by $\max\limits_{1 \leq i \leq n} \Pr(X_i \in V)$, which we use in place of the quantity $2^{d_\pm - n}$. Details are provided in Appendix B.

6 Analyzing the exceptional medium combinatorial dimension case

The approach for exceptional V in [11] is very different from that used in the unexceptional case or in the large or small combinatorial dimension cases. Using some powerful tools from additive combinatorics, the general idea is to put exceptional hyperplanes V in correspondence with a particular additive structure called a generalized arithmetic progression, and then to show that the number of the particular generalized arithmetic progressions that arise in this way is exceedingly small. The key to this approach is a structure theorem—namely, [11, Theorem 5.3]. In this section, we state a slightly modified structure theorem (Theorem 6.1), and then we show how to use Theorem 6.1 to prove Lemma 5.7. In the beginning of Section 7, we outline the changes needed to prove the the structure theorem for our current context, and in Sections 7 and 8 we provide details.

Before stating the structure theorem, we need some definitions and notation. A *generalized arithmetic progression* of rank $\mathfrak{r}$ is a set of the form

$$P = \{v_0 + m_1 v_1 + \cdots + m_{\mathfrak{r}} : |m_i| \leq M_i/2\},$$

where the *basis vectors* $v_0, v_1, \ldots, v_{\mathfrak{r}}$ are elements of a $\mathbb{Z}$-module (here, $\mathbb{Z}/Q\mathbb{Z}$) and where the *dimensions* $M_1, \ldots, M_{\mathfrak{r}}$ are positive integers. We say that v_i has *corresponding dimension* M_i. For

a given element $a = v_0 + m_1 v_1 + \cdots + m_{\mathfrak{r}}$ in P, we refer to $m_1, \ldots, m_{\mathfrak{r}}$ as *coefficients* for a. A generalized arithmetic progression P is *symmetric* if $v_0 = 0$, and P is *proper* if for each $a \in P$, the there is a unique $\mathfrak{r}$-tuple $(m_1, \ldots, m_{\mathfrak{r}})$ with $|m_i| < M_i/2$ that gives the coefficients for a. If P is proper and symmetric, we define the P-norm $\|a\|_P$ of an element $a \in P$ to be

$$\|a\|_P := \left(\sum_{i=1}^{\mathfrak{r}} \left(\frac{m_i}{M_i} \right)^2 \right)^{1/2}.$$

We will use the notation mP, where m is a positive integer, to denote the set $\{\sum_{i=1}^m x_i : x_i \in P\}$ and the notation P^m, where m is a positive integer, to denote the set $\{\prod_{i=1}^m x_i : x_i \in P\}$. If P is a generalized arithmetic progression of rank $\mathfrak{r}$, then so is mP, while P^m, on the other hand, is a generalized arithmetic progression of rank at most $\mathfrak{r}^m$. Also note that $|mP| \leq m^{\mathfrak{r}} |P|$ and that $|P^m| \leq |P|^m$.

Let V be an exceptional hyperplane of medium combinatorial dimension in $\mathrm{Gr}(d_\pm)$ and let $X_{i_{\max}} = (\alpha_1, \ldots, \alpha_n)$ be the exceptional row (here we are using α_j as shorthand for $\alpha_{i_{\max},j}$). Let $(\beta_1^{(\mu)}, \ldots, \beta_n^{(\mu)})$ be the row of random variables corresponding to $X_{i_{\max}}$ from the definition of p-bounded of exponent r, and let $b_{j,s}$ with $1 \leq j \leq n$ and $1 \leq s \leq \ell_j$ be the values taken by $\beta_j^{(\mu)}$ (see Equation (9) for the definition of $\beta_j^{(\mu)}$).

Given an exceptional hyperplane V, there exists a representation of the form

$$V = \{(x_1, x_2, \ldots, x_n) \in (\mathbb{Z}/Q\mathbb{Z})^n \ : \ x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0\}$$

for some elements $a_1, a_2, \ldots, a_n \in \mathbb{Z}/Q\mathbb{Z}$. We will call $a_1, a_2, \ldots, a_n$ the *defining coordinates* of V. Finally, let $\tilde{a}_j := b_{j,1} a_j$. We will refer to $(\tilde{a}_1, \ldots, \tilde{a}_n)$ as the *scaled defining coordinates of* V. Note that once $i_{\max}$ is fixed, so are the elements $b_{j,1}$. We should also note that the choice of $b_{j,1}$ among $b_{j,s}$ for $1 \leq s \leq \ell_j$ is arbitrary—since $\beta_j^{(\mu)}$ takes the values $b_{j,s}$ each with probability at least q, any value of s will do; and so we have taken $s = 1$ for convenience.

Let $\mathbb{H}$ denote the *highly rational numbers*, that is, those numbers in $\mathbb{Z}/Q\mathbb{Z}$ of the form $a/b \pmod{Q}$ where a, b are integers such that $|a|, |b| \leq n^{o(n)}$ and $b \neq 0$. The highly rational numbers were defined in [11, Section 8], and we will need a small extension for the current paper, due to the fact that we are using the scaled defining coordinates of V instead of simply the defining coordinates of V. If we were to assume that $b_{j,1}$ was an $O(1)$ integer for all j and that every possible value taken by α_{ij} was an $O(1)$ integer for all i, j, then we could still use the same definition of highly rational as in [11]. However, if there is a $b_{j,1}$ or an entry α_{ij} in the matrix N_n that ever takes an irrational value, then when we pass to $\mathbb{Z}/Q\mathbb{Z}$ using Lemma 4.1 we have to account for values possibly on the order of Q (see Remark 4.2), and the highly rational numbers are not sufficient for this task. We can overcome this difficulty by extending to the highly T-rational numbers, which contain the highly rational numbers along with all the values in a structured set T (described below). We will now give a rigorous definition the highly T-rational numbers.

Let T be a generalized arithmetic progression in $\mathbb{Z}/Q\mathbb{Z}$ with rank $O(1)$ and having cardinality at most $n^{o(n)}$. As in the definition of p-bounded of exponent r (Definition 2.1), we will take S to be the generalized arithmetic progression containing all possible values in $\mathbb{Z}/Q\mathbb{Z}$ taken by the random variables α_{ij} that are the entries of N_n; thus, by assumption $|S| \leq n^{o(n)}$. By the definition of p-bounded of exponent r, we know that all of the random variables $\beta_{ij}^{(\mu)}$ take values in a set with cardinality $O(1)$. Thus, there is a symmetric generalized arithmetic progression T with rank $O(1)$

and cardinality $|T| \leq n^{o(n)}$ such that T contains S, such that T contains the set $\{-1, 0, 1\}$, and such that T contains all the values taken by the $\beta_{ij}^{(\mu)}$. To construct T from S, one can, for example, add each distinct value taken by a $\beta_{ij}^{(\mu)}$ as a new basis vector v' with corresponding dimension $M' := 3$ (say).

A *highly T-rational number* h is any element of $\mathbb{Z}/Q\mathbb{Z}$ of the form a/b, where $a, b \in n^{o(n)}T^{O(1)}$. Note that therefore, the cardinality of the highly T-rational numbers is at most $(n^{do(n)}\,|T|)^{O(1)} = n^{o(n)}$, where $d = O(1)$ is the rank of T (here we used the fact that $|T| \leq n^{o(n)}$).

Theorem 6.1 (Structure Theorem)**.** *There is a constant $C = C(\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, q, r, \mu)$ such that the following holds. Let V be an exceptional hyperplane and let $\tilde{a}_1, \ldots, \tilde{a}_n$ be its scaled defining coordinates (as described above). Then there exist integers*

$$1 \leq \mathfrak{r} \leq C$$

and $M_1, \ldots, M_{\mathfrak{r}} \geq 1$ with the volume bound

$$M_1 \cdots M_{\mathfrak{r}} \leq C \Pr(X_{i_{\max}} \in V)^{-1}$$

and nonzero elements $v_1, \ldots, v_{\mathfrak{r}} \in \mathbb{Z}/Q\mathbb{Z}$ such that the following holds

- *(i) (Scaled defining coordinates lie in a progression) The symmetric generalized arithmetic progression*

$$P := \{m_1 v_1 + \cdots + m_{\mathfrak{r}} v_{\mathfrak{r}} : -M_i/2 < m_i < M_i/2\}$$

  is proper and contains all of the $\tilde{a}_j$.

- *(ii) (Bounded norm) The $\tilde{a}_j$ have small P-norm:*

$$\sum_{j=1}^{n} \|\tilde{a}_j\|_P^2 \leq C.$$

- *(iii) (Rational T-commensurability) The set $\{v_1, \ldots, v_{\mathfrak{r}}\} \cup \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ is contained in the set*

$$\{hv_1 : \ h \text{ is highly } T\text{-rational}\}.$$

Note that unlike [11], part (iii) above does not necessarily place $\{v_1, \ldots, v_{\mathfrak{r}}\} \cup \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ in a simple arithmetic progression.

We will discuss the proof of the structure theorem in Sections 7 and 8. In the remainder of this section, we will discuss how to use the structure theorem to prove Lemma 5.7.

Fix $d_\pm$ of medium combinatorial dimension (see Proposition 5.4). Using independence of the rows, we have

$$\begin{aligned}\sum_{\substack{V \in \mathrm{Gr}(d_\pm): \\ V \text{ is exceptional}}} \Pr(A_V) &\leq \sum_{\substack{V \in \mathrm{Gr}(d_\pm): \\ V \text{ is exceptional}}} \prod_{i=1}^{n} \Pr(X_i \in V) \\ &\leq |\{V \in \mathrm{Gr}(d_\pm) : V \text{ is exceptional}\}| \cdot \left(\max_{1 \leq i \leq n} \Pr(X_i \in V)\right)^n. \qquad (20)\end{aligned}$$

In [11, Section 5], it is shown using Theorem 6.1(i) and (ii) and Gaussian-type methods (and the fact that $\mathfrak{r}$ is bounded by a constant) that

$$|\{V \in \mathrm{Gr}(d_\pm) : V \text{ is exceptional}\}| \leq \frac{n^{o(n)}}{Q-1} \sum_{\substack{\mathfrak{r},\{M_1,\ldots,M_\mathfrak{r}\} \\ \{v_1,\ldots,v_\mathfrak{r}\}}} \left(1 + n^{-1/2} M_1 \cdots M_\mathfrak{r}\right)^n,$$

where the sum runs over all possible values for $\mathfrak{r}$, for the M_i, and for $v_1, \ldots, v_\mathfrak{r}$. By Theorem 6.1, we know that $\mathfrak{r} \leq C = O(1)$ and that $M_i \leq M_1 M_2 \cdots M_\mathfrak{r} \leq C \Pr(X_{i_{\max}} \in V)^{-1} \leq O(1/p^n)$; thus, there are at most $n^{o(n)}$ choices for $\mathfrak{r}$ and the M_i. Furthermore, there are at most $Q-1$ choices for v_1 (since $v_1 \neq 0$), and once the value for v_1 has been fixed, (iii) tells us that there are at most $n^{o(n)}$ choices for $\{v_2, \ldots, v_\mathfrak{r}\}$ (since $\left|n^{o(n)} T^{O(1)}\right| \leq n^{o(n)}$). Thus, the sum runs over at most $n^{o(n)}$ terms. (This is the point in the proof where it is essential that $n^{o(n)} T^{O(1)}$ has cardinality $n^{o(n)}$.)

Plugging the volume bound on $M_1 \cdots M_\mathfrak{r}$ into the previous displayed inequality, we have

$$\begin{aligned} |\{V \in \mathrm{Gr}(d_\pm) : V \text{ is exceptional}\}| &\leq n^{o(n)} \left(1 + n^{-\frac{1}{2}} C \Pr(X_{i_{\max}} \in V)^{-1}\right)^n \\ &= n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-n}, \end{aligned} \tag{21}$$

using the fact that $\Pr(X_{i_{\max}} \in V) \leq \frac{c_{\mathrm{LgDim}}}{\sqrt{n}}$, which is a consequence of $d_\pm$ being of medium combinatorial dimension. Plugging in Inequality (21) into Inequality (20) and summing over all $d_\pm$ of medium combinatorial dimension completes the proof of Lemma 5.7 (recall that by assumption $\max_{1 \leq i \leq n} \Pr(X_i \in V) = \Pr(X_{i_{\max}} \in V)$).

7 Halász-type arguments

The proof of the structure theorem has two main ingredients: tools from additive combinatorics, and Halász-type arguments using discrete Fourier analysis. Our proof of Theorem 6.1 will follow the proof of [11, Theorem 5.2] very closely. We will use results about additive combinatorics from [11, Section 6] directly, and we will discuss below the extent to which the Halász-type arguments of [11, Section 7] need to be modified to work for our current context. The proof of Theorem 6.1 will be given in Section 8 using results from the current section, [11, Section 6], [11, Section 7], and [11, Section 8]. Our Section 8 follows [11, Section 8] closely, with a few modifications to prove rational T-commensurability instead of only rational commensurability.

In this section we discuss modifications to the lemmas in [11, Section 7] that are needed in order to prove Theorem 6.1.

We will use $e_Q(\cdot)$ to denote the primitive character

$$e_Q(x) := \exp(2\pi i x/Q).$$

Let $i_{\max}$ be the index of the exceptional row, so for every $1 \leq k \leq r$ we have

$$\epsilon_1 \Pr(Z^*_{i_{\max},k} \in V) \leq \Pr(X_{i_{\max}} \in V), \tag{22}$$

and recall that by Definition 5.5 we have $\Pr(X_{i_{\max}} \in V) = \max_i \Pr(X_i \in V)$. Let $(\alpha_1, \ldots, \alpha_n) := X_{i_{\max}}$ with the corresponding random variables $(\beta_1^{(\mu)}, \ldots, \beta_n^{(\mu)})$ from the definition of p-bounded of

exponent r (see Definition 2.1 and Equation (9)), and let $(a_1, \dots, a_n)$ be the defining coordinates of V. Then, using the Fourier expansion, we can compute

$$\begin{aligned}
\Pr(X_{i_{\max}} \in V) = \mathbb{E}(\mathbf{1}_{\{X_{i_{\max}} \in V\}}) &= \mathbb{E}\left(\frac{1}{Q}\sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} e_Q(\sum_{j=1}^{n} \alpha_j a_j \xi)\right) \\
&\le \frac{1}{Q}\sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^{n} |\mathbb{E}(e_Q(\alpha_j a_j \xi))| \\
&\le \frac{1}{Q}\sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^{n} \mathbb{E}\left(e_Q(\beta_j^{(\mu)} a_j \xi)\right)^{1/r} \\
&= \frac{1}{Q}\sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^{n} \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r} \qquad (23) \\
&\le \frac{1}{Q}\sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^{n} \left(1 - \underline{\mu} + \underline{\mu} \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r} \\
&\le \prod_{k=1}^{r} \Pr(Z^*_{i_{\max},k} \in V)^{1/r}, \qquad (24)
\end{aligned}$$

where the last line is an application of Hölder's inequality.

Define

$$f(\xi) := \prod_{j=1}^{n} \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \qquad (25)$$

$$f_j(\xi) := \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \quad \text{and} \qquad (26)$$

$$g_k(\xi) := \prod_{(k-1)\frac{n}{r} < j \le k\frac{n}{r}} \left(1 - \underline{\mu} + \underline{\mu} \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \qquad (27)$$

where $\underline{\mu} := \mu - \frac{\epsilon_0}{100}$, as defined in Section 5.2. Note that $f(\xi) = \prod_{j=1}^{n} f_j(\xi)$.

We will need the following analog of [11, Lemma 7.1]:

Lemma 7.1. *For all $\xi \in \mathbb{Z}/Q\mathbb{Z}$, we have*

$$\prod_{j=1}^{n} f_j(\xi)^{r\underline{\mu}/\mu} \le \prod_{k=1}^{r} g_k(\xi)$$

Proof. This inequality may be proven pointwise (for each j after expanding out the definition of g_k) using the convexity of the log function, just as in the proof of [11, Lemma 7.1] (see also [10, Lemma 7.1]. □

Let ϵ_2 be sufficiently small compared to ϵ_1 (we will specify how small in Inequality (33) while proving Lemma 7.2). Following [11], we define the *spectrum* $\Lambda \subset \mathbb{Z}/Q\mathbb{Z}$ of $\{b_{1,1}a_1, \ldots, b_{n,1}a_n\} = \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ (the scaled defining coordinates of V) to be

$$\Lambda := \{\xi \in \mathbb{Z}/Q\mathbb{Z} : f(\xi) \geq \epsilon_2\}. \tag{28}$$

Let $\|x\|_{\mathbb{R}/\mathbb{Z}}$ denote the distance from $x \in \mathbb{R}$ to the nearest integer. Using the elementary inequality $\cos(2\pi x) \leq 1 - \frac{1}{100}\|x\|^2_{\mathbb{R}/\mathbb{Z}}$, we have

$$\begin{aligned} f(\xi) &\leq \exp\left(-\frac{\mu}{100r}\sum_{j=1}^{n}\sum_{s=1}^{\ell_j} p_{j,s}\,\|b_{j,s}a_j\xi/Q\|^2_{\mathbb{R}/\mathbb{Z}}\right) \\ &\leq \exp\left(-\frac{q}{50r}\sum_{j=1}^{n}\|b_{j,1}a_j\xi/Q\|^2_{\mathbb{R}/\mathbb{Z}}\right) \end{aligned} \tag{29}$$

($\mu p_{j,1} \geq 2q$ since $\min_x \Pr(\beta_j^{(\mu)} = x) \geq q$ by Definition 2.1).

Thus, there is a constant $C(\epsilon_2, q, r)$ such that

$$\left(\sum_{j=1}^{n}\|\tilde{a}_j\xi/Q\|^2_{\mathbb{R}/\mathbb{Z}}\right)^{1/2} = \left(\sum_{j=1}^{n}\|b_{j,1}a_j\xi/Q\|^2_{\mathbb{R}/\mathbb{Z}}\right)^{1/2} \leq C(\epsilon_2, q, r), \tag{30}$$

for every $\xi \in \Lambda$. (E.g., the constant $C(\epsilon_2, q, r) := \left(\frac{50r}{q}\ln\left(\frac{1}{\epsilon_2}\right)\right)^{1/2}$ suffices.)

Lemma 7.2. *There exists a constant C depending on $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, q, r$, and μ such that*

$$C^{-1}Q\Pr(X_{i_{\max}} \in V) \leq |\Lambda| \leq CQ\Pr(X_{i_{\max}} \in V). \tag{31}$$

Furthermore, for every integer $k \geq 4$ we have

$$|k\Lambda| \leq \binom{C+k-3}{k-2} CQ\Pr(X_{i_{\max}} \in V). \tag{32}$$

Proof. Our goal is to bound $\sum_{\xi\in\Lambda} f(\xi)$ from above and below, and then pass to bounds on $|\Lambda|$ using the fact that $\epsilon_2 \leq f(\xi) \leq 1$ for all $\xi \in \Lambda$.

Note that

$$\frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}} f(\xi) \geq \Pr(X_{i_{\max}} \in V) \qquad \text{(by Equation (25) and Equation (23)).}$$

Also,

$$\begin{aligned}
\frac{1}{Q}\sum_{\xi\notin\Lambda} f(\xi) = \frac{1}{Q}\sum_{\xi\notin\Lambda}\prod_{j=1}^{n} f_j(\xi) \quad &\leq \quad \epsilon_2^{1-\underline{\mu}/\mu}\frac{1}{Q}\sum_{\xi\notin\Lambda}\prod_{j=1}^{n} f_j(\xi)^{\underline{\mu}/\mu} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu}\frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}\prod_{k=1}^{r} g_k(\xi)^{1/r} && \text{(Lemma 7.1)} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu}\frac{1}{Q}\left(\prod_{k=1}^{r}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}} g_k(\xi)\right)^{1/r} && \text{(Hölder's inequality)} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu}\left(\frac{1}{\epsilon_1}\right)\Pr(X_{i_{\max}}\in V) && \text{(by Inequality (22))} \quad .
\end{aligned}$$

For the lower bound, we have

$$\begin{aligned}
\sum_{\xi\in\Lambda} f(\xi) &= \sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}} f(\xi) - \sum_{\xi\notin\Lambda} f(\xi) \\
&\geq Q\Pr(X_{i_{\max}}\in V) - \frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1} Q\Pr(X_{i_{\max}}\in V) \\
&= Q\Pr(X_{i_{\max}}\in V)\left(1-\frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1}\right).
\end{aligned}$$

We can choose ϵ_2 sufficiently small with respect to ϵ_1 and $1-\underline{\mu}/\mu$ so that, for example,

$$1-\frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1} \geq \frac{1}{2}. \tag{33}$$

For the upper bound, we have

$$\begin{aligned}
\sum_{\xi\in\Lambda} f(\xi) &\leq \sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}} f(\xi) \\
&\leq Q\prod_{k=1}^{r}\Pr(Z^*_{i_{\max},k}\in V)^{1/r} && \text{(Inequality (24))} \\
&\leq Q\frac{1}{\epsilon_1}\Pr(X_{i_{\max}}\in V) && \text{(Inequality (22))} \quad .
\end{aligned}$$

Thus, we have shown that $\sum_{\xi\in\Lambda} f(\xi) = \Theta(Q\Pr(X_{i_{\max}}\in V))$. Since $\epsilon_2 \leq f(\xi) \leq 1$ for all $\xi\in\Lambda$, we have proven Inequality (31).

Making use of [11, Lemma 6.4], we can prove Inequality (32) by showing $|4\Lambda| \leq C|\Lambda|$ for some constant C. Using Lemma 7.3 below (for which we need to assume strict positivity of $\mathbb{E}(e(\beta_j^{(\mu)}t))$—see Remark 2.3), we have that there exists a constant $c := c(\epsilon_{-1},\epsilon_2)$ such that

$$f(\xi) \geq c(\epsilon_{-1},\epsilon_2),$$

for every $\xi \in 4\Lambda$. Thus,

$$\begin{aligned}|4\Lambda| &\leq \frac{1}{c(\epsilon_{-1},\epsilon_2)} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} f(\xi) \\ &\leq \left(\frac{1}{c(\epsilon_{-1},\epsilon_2)}\right) \frac{Q}{\epsilon_1} \Pr(X_{i_{\max}} \in V) = C\,|\Lambda|\,,\end{aligned}$$

for some constant C. This completes the proof of Lemma 7.2. □

We now state and prove a lemma showing that $f(\xi)$ is at least a constant for all $\xi \in 4\Lambda$. In [11], the lemma below is unnecessary because an inequality following from [11, Inequality (30)] (which corresponds to Inequality (30)) and the triangle inequality suffices.

Lemma 7.3. *Let Λ and f be defined as in Equation (28) and Equation (25), respectively. If $\xi \in 4\Lambda$, then*

$$f(\xi) \geq \left(\epsilon_2 \epsilon_{-1}^{\ln(1/\epsilon_2)}\right)^{320000} =: c(\epsilon_{-1},\epsilon_2).$$

Note that $c(\epsilon_{-1},\epsilon_2)$ is a constant.

Proof. Note that Inequality (29) implies that for any $\xi' \in \Lambda$ we have

$$\left(\sum_{j=1}^{n}\sum_{s=1}^{\ell_j} p_{j,s} \left\|b_{j,s}a_j\xi'/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} \leq \left(\frac{100r}{\mu}\ln\left(\frac{1}{\epsilon_2}\right)\right)^{1/2}.$$

Thus, by the triangle inequality, we have for any $\xi \in 4\Lambda$ that

$$\left(\sum_{j=1}^{n}\sum_{s=1}^{\ell_j} p_{j,s} \left\|b_{j,s}a_j\xi/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} \leq 4\left(\frac{100r}{\mu}\ln\left(\frac{1}{\epsilon_2}\right)\right)^{1/2}. \tag{34}$$

Fix $\xi \in 4\Lambda$. Let k_0 be the number of indices j such that

$$100\mu \sum_{s=1}^{\ell_j} p_{j,s} \left\|b_{j,s}a_j\xi/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2 > \frac{1}{2},$$

and without loss of generality, say that these indices are $j = 1, 2, \ldots, k_0$. Squaring Inequality (34), we see that $\frac{k_0}{200\mu} \leq \frac{1600r}{\mu}\ln\left(\frac{1}{\epsilon_2}\right)$, and so we have

$$k_0 \leq 320000r\ln\left(\frac{1}{\epsilon_2}\right),$$

which is a constant. Thus, for the vast majority of the indices j, namely $j = k_0+1, k_0+2, \ldots, n$, we have

$$100\mu \sum_{s=1}^{\ell_j} p_{j,s} \left\|b_{j,s}a_j\xi/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2 \leq \frac{1}{2}. \tag{35}$$

We may now compute that

$$
\begin{aligned}
f(\xi) &:= \prod_{j=1}^{n} \left(1-\mu+\mu\sum_{s=1}^{\ell_j} p_{j,s}\cos(2\pi b_{j,s}a_j\xi/Q)\right)^{1/r} \\
&\geq \epsilon_{-1}^{k_0/r} \prod_{j=k_0+1}^{n} \left(1-\mu+\mu\sum_{s=1}^{\ell_j} p_{j,s}\cos(2\pi b_{j,s}a_j\xi/Q)\right)^{1/r} && \text{(since } f(\xi')\geq\epsilon_{-1} \text{ for any } \xi' \text{ by the assumption of strict positivity—see Remark 2.3))} \\
&\geq \epsilon_{-1}^{k_0/r} \prod_{j=k_0+1}^{n} \left(1-100\mu\sum_{s=1}^{\ell_j} p_{j,s}\left\|b_{j,s}a_j\xi/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/r} && \text{(since } \cos(2\pi x)\geq 1-100\left\|x\right\|_{\mathbb{R}/\mathbb{Z}}^2 \text{ and the factors are all positive by Inequality (35))} \\
&\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r}\sum_{j=k_0+1}^{n}\sum_{s=1}^{\ell_j} p_{j,s}\left\|b_{j,s}a_j\xi/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right) && (1-x\geq e^{-2x} \text{ for } 0\leq x\leq .79) \\
&\geq \epsilon_{-1}^{320000\ln\left(\frac{1}{\epsilon_2}\right)} \exp\left(-320000\ln\left(\frac{1}{\epsilon_2}\right)\right) && \text{(by Inequality (34))} \\
&= \left(\epsilon_2\epsilon_{-1}^{\ln(1/\epsilon_2)}\right)^{320000}.
\end{aligned}
$$

This completes the proof. □

We have shown that the spectrum Λ has small doubling, and the next step is to use this fact to show that a set containing most of the scaled defining coordinates $\tilde{a}_j$ also has small doubling. Towards that end, we will use the Λ-norm from [11], which is defined as follows: for $x \in \mathbb{Z}/Q\mathbb{Z}$, let $\|x\|_\Lambda$ be defined by

$$\|x\|_\Lambda := \left(\frac{1}{|\Lambda|^2}\sum_{\xi,\xi'\in\Lambda}\left\|x(\xi-\xi')/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2}.$$

Note that $0 \leq \|x\|_\Lambda \leq 1$ for all x and that the triangle inequality holds: $\|x+y\|_\Lambda \leq \|x\|_\Lambda + \|y\|_\Lambda$. We also have that

$$
\begin{aligned}
\|x\|_\Lambda &\leq \left(\frac{1}{|\Lambda|^2}\sum_{\xi,\xi'\in\Lambda}\|x\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} + \left(\frac{1}{|\Lambda|^2}\sum_{\xi,\xi'\in\Lambda}\left\|x\xi'/Q\right\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} \\
&= 2\left(\frac{1}{|\Lambda|}\sum_{\xi\in\Lambda}\|x\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2}.
\end{aligned}
$$

Thus, squaring Inequality (30) and summing over all $\xi \in \Lambda$, we have

$$\sum_{j=1}^{n}\|\tilde{a}_j\|_\Lambda^2 \leq 4C(\epsilon_2,q,r) =: C'. \tag{36}$$

We will now show that the set of all x with small Λ-norm, which by Inequality (36) includes most of the $\tilde{a}_j$, has small doubling.

Lemma 7.4. [11, Lemma 7.4] *There is a constant C such that the following holds. Let $A \subseteq \mathbb{Z}/Q\mathbb{Z}$ denote the "Bohr set":*

$$A := \{x \in \mathbb{Z}/Q\mathbb{Z} : \|x\|_\Lambda < \frac{1}{100}\}.$$

Then we have

$$C^{-1} \Pr(X_{i_{\max}} \in V)^{-1} \leq |A| \leq |A + A| \leq C \Pr(X_{i_{\max}} \in V)^{-1}.$$

The proof of Lemma 7.4 is the same as in [11], with the small modification that a_j should be replaced with $\tilde{a}_j := b_{j,1} a_j$ and the quantity $2^{d_\pm - n}$ should be replaced with $\Pr(X_{i_{\max}} \in V)$ (and, of course, the field F in [11] should be replaced with $\mathbb{Z}/Q\mathbb{Z}$). Also, one should note that [11, Inequality (30)], [11, Inequality (31)], and [11, Inequality (32)] correspond to, respectively, Inequalities (30), (31), and (32).

In the next section, we will complete the proof of the structure theorem using the lemma above.

8 Proof of the Structure Theorem (Theorem 6.1)

The key to proving the structure theorem is an application of Freiman's Theorem for finite fields.

Theorem 8.1 (see Lemma 6.3 in [11])**.** *For any constant C there are constants $\mathfrak{r}$ and δ such that the following holds. Let A be a non-empty subset of $\mathbb{Z}/Q\mathbb{Z}$, a finite field of prime order Q, such that $|A + A| \leq C|A|$. Then, if Q is sufficiently large depending on $|A|$, there is a symmetric generalized arithmetic progression P of rank $\mathfrak{r}$ such that $A \subset P$ and $|A|/|P| \geq \delta$.*

Note that by Lemma 4.1 we can assume that Q is sufficiently large with respect to $|A| \leq C \Pr(X_{i_{\max}} \in V)^{-1} \leq C(1/p)^n$ (this follows from V being of medium combinatorial dimension).

The set A from Lemma 7.4 satisfies $|A + A| \leq C^2|A|$, where $C \leq O(1)$, and also contains all but $O(1)$ of the scaled defining coordinates $\tilde{a}_j$, since $\tilde{a}_j \notin A$ implies that $\|\tilde{a}_j\|_\Lambda \geq 1/100$ and Inequality (36) shows that there can be at most $100C' = O(1)$ such $\tilde{a}_j$. By Theorem 8.1, there exists a symmetric generalized arithmetic progression $P = \{m_1 v_1 + \cdots + m_{\mathfrak{r}} v_{\mathfrak{r}} : |m_i| < M_i/2\}$ containing A and satisfying the bounds:

$$\text{rank}(P) = \mathfrak{r} \leq O(1) \text{ and} \tag{37}$$

$$|P| \leq M_1 M_2 \cdots M_{\mathfrak{r}} \leq O(\Pr(X_{i_{\max}} \in V)^{-1}). \tag{38}$$

The symmetric generalized arithmetic progression P is close to what is needed for Theorem 6.1, since it satisfies the required volume and rank bounds. We will show below that P can be altered in ways that preserve Inequalities (37) and (38) (except possibly for changing the implicit constants) so that P satisfies conditions (i), (ii), and (iii) of Theorem 6.1.

To show Theorem 6.1(i), we will first add the remaining scaled defining coordinates $\{\tilde{a}_1, \ldots, \tilde{a}_n\} \setminus P$ (i.e., those $\tilde{a}_j$ such that $\|\tilde{a}_j\|_\Lambda \geq 1/100$) as new basis vectors v'_k with corresponding dimensions M'_k equal to (say) 3. The resulting generalized arithmetic progression, which we will continue to call P by abuse of notation, satisfies both Inequalities (37) and (38), since there are only $O(1)$ of the $\tilde{a}_j$ with $\|\tilde{a}_j\|_\Lambda \geq 1/100$ (by Inequality (36)). Second, we need to ensure that P is proper, for which we will use the following lemma:

Lemma 8.2 (cf. Lemma 9.3 in [11])**.** *There is an absolute constant $C_0 \geq 1$ such that the following holds. Let P be a symmetric progression of rank $\mathfrak{r}$ in a abelian group G, such that every nonzero*

element of G has order at least $\mathfrak{r}^{C_0\mathfrak{r}^3}|P|$. Then there exists a proper symmetric generalized arithmetic progression P' of rank at most $\mathfrak{r}$ containing P such that

$$|P'| \leq \mathfrak{r}^{C_0\mathfrak{r}^3}|P|.$$

Furthermore, if P is not proper and $\mathfrak{r} \geq 2$, then P' can be chosen to have rank an most $\mathfrak{r}-1$

One can conclude Lemma 8.2 from the proof of [11, Lemma 9.3] (the only difference is noting that the rank can be reduced by at least 1 if P is not proper to begin with). Note that we can always choose Q larger than $\mathfrak{r}^{C_0\mathfrak{r}^3}|P| \leq O\left(\frac{1}{p}\right)^n$.

Applying Lemma 8.2 gives us a proper symmetric generalized arithmetic progression, which again we call P by abuse of notation, that contains all the $\tilde{a}_j$ and satisfies both Inequalities (37) and (38).

The next task is to show that P can be further altered so to meet the condition (ii) in Theorem 6.1. Note that there are only $O(1)$ scaled defining coordinates $\tilde{a}_j$ such that $\|\tilde{a}_j\|_\Lambda \geq 1/100$, and so these $\tilde{a}_j$ contribute only a constant to the sum $\sum_{j=1}^n \|\tilde{a}_j\|_P^2$. On the other hand, for any $\tilde{a}_j$ with $\|\tilde{a}_j\|_\Lambda < 1/100$, we have that $k\tilde{a}_j \in A \subset P$ for every positive integer $k < \frac{1}{100\|\tilde{a}_j\|_\Lambda}$. We will exploit this fact, and to do so will need the following notation. Let $\Phi_P : P \to \mathbb{Z}^{\mathfrak{r}}$ be the map sending a point $m_1v_1 + \cdots + m_{\mathfrak{r}}v_{\mathfrak{r}}$ in the proper generalized arithmetic progression P to the unique r-tuple of coefficients $(m_1, \ldots, m_{\mathfrak{r}})$.

If the representation for $\tilde{a}_j$ in P is $\tilde{a}_j = m_1v_1 + \cdots + m_{\mathfrak{r}}v_{\mathfrak{r}}$ and $k\tilde{a}_j$ is in P, we would like to be able to say that the representation for $k\tilde{a}_j$ is $km_1v_1 + \cdots + km_{\mathfrak{r}}v_{\mathfrak{r}}$; i.e., we hope that $\Phi_P(k\tilde{a}_j)$ is equal to $k\Phi_P(\tilde{a}_j)$. If this were true, then we would have $|km_i| \leq M_i$ for $1 \leq i \leq \mathfrak{r}$, which, if k is large, would show that $\|\tilde{a}_j\|_P$ is small. However, at this point we may well have $\Phi_P(k\tilde{a}_j) \neq k\Phi_P(\tilde{a}_j)$. A priori, changing this to equality would require replacing P with kP and then applying Lemma 8.2 to get a proper symmetric generalized arithmetic progression, but since k may be large, this would increase the volume of P too much, violating Inequality (38). Luckily, the lemma below provides a way around this difficulty. We will say that P is *(k_j, x_j)-proper* if $\Phi_P(k_jx_j) = k_j\Phi_P(x_j)$.

Lemma 8.3. *There exists an absolute constant C_1 such that the following holds. Let P be a symmetric proper generalized arithmetic progression with rank $\mathfrak{r}$ containing elements $x_1, \ldots, x_m$, and let $k_1, \ldots, k_m$ be positive integers such that $\ell_j x_j \in P$ for every $1 \leq \ell_j \leq k_j$ and for every j. Then, there exists a proper symmetric generalized arithmetic progression P' of rank at most $\mathfrak{r}$ such that P' contains P,*

$$\begin{aligned} &\left|P'\right| \leq \mathfrak{r}^{C_1\mathfrak{r}^4}\left|P\right|, \text{ and} \\ &\quad P \text{ is } (k_j, x_j)\text{-proper for every } j. \end{aligned}$$

Furthermore, if $r \geq 2$ and if there is some j for which P is not (k_j, x_j)-proper, then P' can be chosen to have rank at most $\mathfrak{r}-1$.

The proof of this lemma relies on an application of Lemma 8.2 to $2P$ (which contains P) along with the fact that if $\|\tilde{a}_j\|_\Lambda < 1/100$ then $k\tilde{a}_j \in P$ for *every* $1 \leq k < \frac{1}{100\|\tilde{a}_j\|_\Lambda}$.

Proof. We proceed by induction on the rank $\mathfrak{r}$. For the base case, let $\mathfrak{r} = 1$ and consider $x_j \in P$ such that $k_jx_j \in P$. Since P has rank 1 in this case, we have that $x_j = \Phi_P(x_j)v_1$ and $k_jx_j = \Phi_P(k_jx_i)v_1$. Combining these two equations we have $k_j\Phi_P(x_j)v_1 = \Phi_P(k_jx_j)v_1$, and dividing by v_1 (note that

we may assume that $v_1 \neq 0$), we see that $k_j \Phi_P(x_j) = \Phi_P(k_j x_j)$. Thus P is (k_j, x_j)-proper for every j.

For $\mathfrak{r} \geq 2$, we may assume that there is some j_0 such that $k_{j_0}\Phi_P(x_{j_0}) \neq \Phi_P(k_{j_0}x_{j_0})$ (i.e., we assume that P is not (k_{j_0}, x_{j_0})-proper). We may assume that P has the form $\{m_1v_1 + \cdots + m_{\mathfrak{r}}v_{\mathfrak{r}} : |m_i| < M_i/2\}$. Let $M := (M_1, \ldots, M_{\mathfrak{r}})$, and let $(-M/2, M/2)$ denote the box $\{(m_1, \ldots, m_{\mathfrak{r}}) : |m_i| < M_i/2\}$.

Let $\underline{k}$ be the largest integer such that $\Phi_P(\underline{k}x_{j_0}) = \underline{k}\Phi_P(x_{j_0})$, so $1 \leq \underline{k} < k_{j_0}$ and $\Phi_P((\underline{k}+1)x_{j_0}) \neq (\underline{k}+1)\Phi_P(x_{j_0})$. Since $\underline{k}x_{j_0} \in P$ and $x_{j_0} \in P$, we know that $\Phi_P(x_{j_0}) \in (-M/2, M/2)$ and $\Phi_P(\underline{k}x_{j_0}) = \underline{k}\Phi_P(x_{j_0}) \in (-M/2, M/2)$; and thus, $(\underline{k}+1)\Phi_P(x_{j_0}) \in (-M, M)$. This shows that $2P$, which has dimensions $2M = (2M_1, \ldots, 2M_{\mathfrak{r}})$, is not proper, since it has two distinct representations for $(\underline{k}+1)x_{j_0}$.

We can now apply Lemma 8.2 to $2P$, thus finding a proper symmetric generalized arithmetic progression P' of rank at most $\mathfrak{r}-1$ containing $2P$ (which contains P) such that

$$\left|P'\right| \leq \mathfrak{r}^{C_0\mathfrak{r}^3} \left|2P\right| \leq \mathfrak{r}^{2C_0\mathfrak{r}^3} \left|P\right|.$$

Since P' has rank at most $\mathfrak{r}-1$, we have by induction that there exists P'' a proper symmetric generalized arithmetic progression of rank at most $\mathfrak{r}-1$ containing P' and such that

$$\left|P''\right| \leq (\mathfrak{r}-1)^{C_1(\mathfrak{r}-1)^4} \left|P'\right| \leq \mathfrak{r}^{C_1(\mathfrak{r}-1)^4}\mathfrak{r}^{2C_0\mathfrak{r}^3} \left|P\right|,$$

and such that P'' is (k_j, x_j)-proper for every j. Choosing $C_1 \geq 2C_0$ (for example) guarantees that $\mathfrak{r}^{C_1(\mathfrak{r}-1)^4}\mathfrak{r}^{2C_0\mathfrak{r}^3} \leq \mathfrak{r}^{C_1\mathfrak{r}^4}$, which completes the induction. □

Applying Lemma 8.3, we can generate a new proper symmetric generalized arithmetic progression, which again we will call P by abuse of notation, such that P contains the $\tilde{a}_j$, satisfies Inequalities (37) and (38), and is $(k_j, \tilde{a}_j)$-proper for every $\tilde{a}_j$ such that $\|\tilde{a}_j\|_\Lambda < 1/100$, where $k_j := \left\lceil \frac{1}{200\|\tilde{a}_j\|_\Lambda} \right\rceil \geq 1$. We will now show that such P satisfies part (ii) of Theorem 6.1. For $\tilde{a}_j$ such that P is $(k_j, \tilde{a}_j)$-proper, we have that $|k_j m_i| \leq M_i$ for each $1 \leq i \leq \mathfrak{r}$, and so

$$\|\tilde{a}_j\|_P = \sum_{i=1}^{\mathfrak{r}} \left(\frac{m_i}{M_i}\right)^2 \leq \sum_{i=1}^{\mathfrak{r}} \left(\frac{1}{k_j}\right)^2 \leq \sum_{i=1}^{\mathfrak{r}} (200\|\tilde{a}_j\|_\Lambda)^2 = 40000\mathfrak{r}\|\tilde{a}_j\|_\Lambda^2.$$

Thus, part (ii) of Theorem 6.1 follows from Inequality (36), since P is $(k_j, \tilde{a}_j)$-proper for all but $O(1)$ of the $\tilde{a}_j$.

The next step is to make further alterations to P so that we can prove part (iii) of Theorem 6.1. The key property that we will use for (iii) is to have the set of vectors $\{\Phi_P(\tilde{a}_j) : 1 \leq j \leq n\}$ span all of $\mathbb{R}^{\mathfrak{r}}$, and we will use a rank reduction argument on P to produce a new proper symmetric generalized arithmetic progression satisfying this full rank property.

Lemma 8.4. [11] *Let P be a proper symmetric generalized arithmetic progression of rank $\mathfrak{r}$ containing a set B such that the set of vectors $\Phi_P(B)$ does not span $\mathbb{R}^{\mathfrak{r}}$. Then there exists a symmetric generalized arithmetic progression P' containing P such that*

$$\begin{aligned} \operatorname{rank}(P') &\leq \mathfrak{r}-1 \text{ and} \\ \left|P'\right| &\leq \left|P\right|. \end{aligned}$$

Note that the resulting P' is not necessarily proper or $(k_j, \tilde{a}_j)$-proper, even if P had these properties.

Proof. We use the same proof here as appears in [11, Section 8]. If $\{\Phi_P(\tilde{a}_j) : 1 \le j \le n\}$ does not have rank $\mathfrak{r}$, then it is contained is a subspace of $\mathbb{R}^{\mathfrak{r}}$ of dimension $\mathfrak{r} - 1$. Thus, there exists an integer vector $(\alpha_1, \ldots, \alpha_{\mathfrak{r}})$ with all the α_i coprime such that $(\alpha_1, \ldots, \alpha_{\mathfrak{r}})$ is orthogonal to every vector in $\{\Phi_P(\tilde{a}_j) : 1 \le j \le n\}$. Thus, for every $w \in \mathbb{Z}/Q\mathbb{Z}$ and and any $\tilde{a}_j = m_1 v_1 + \cdots + m_{\mathfrak{r}} v_{\mathfrak{r}}$, we have that

$$\tilde{a}_j = m_1 v_1 + \cdots + m_{\mathfrak{r}} v_{\mathfrak{r}} = m_1(v_1 - w\alpha_1) + \cdots + m_{\mathfrak{r}}(v_{\mathfrak{r}} - w\alpha_{\mathfrak{r}}).$$

Since not all the α_i are zero, we may assume that $\alpha_{\mathfrak{r}} \neq 0$. Setting $w = v_{\mathfrak{r}}/\alpha_{\mathfrak{r}}$ so that $v_{\mathfrak{r}} - w\alpha_{\mathfrak{r}} = 0$, we see that P is contained in the symmetric generalized arithmetic progression

$$P' := \{m_1' v_1' + \cdots + m_{\mathfrak{r}-1}' v_{\mathfrak{r}-1}' : |m_i'| < M_i/2\}$$

with rank $\mathfrak{r} - 1$, dimensions $M_1, \ldots, M_{\mathfrak{r}-1}$ (which are the same as the corresponding dimensions for P), and basis vectors $v_i' := v_i - \alpha_i v_{\mathfrak{r}}/\alpha_{\mathfrak{r}}$. By construction $|P'| \le |P|$. □

We can now run the following algorithm to create a generalized arithmetic progression with all the desired properties. As the input, we take the generalized arithmetic progression P that we arrived at after applying Lemma 8.3, thus the input P contains all the $\tilde{a}_j$, satisfies Inequalities (37) and (38), and is $(k_j, \tilde{a}_j)$-proper for every $\tilde{a}_j$ such that $\|\tilde{a}_j\|_\Lambda < 1/100$; however, we do not yet know whether $\Phi_P(\{\tilde{a}_j : 1 \le j \le n\})$ spans $\mathbb{R}^{\mathfrak{r}}$.

1. If $\Phi_P(\{\tilde{a}_j : 1 \le j \le n\})$ spans $\mathbb{R}^{\mathfrak{r}}$, then do nothing; otherwise apply Lemma 8.4.
2. If P is proper, then do nothing; otherwise apply Lemma 8.2.
3. If for every $\tilde{a}_j$ with $\|\tilde{a}_j\|_\Lambda < 1/100$ we have that P is $(k_j, \tilde{a}_j)$-proper, then do nothing; otherwise apply Lemma 8.3.
4. If P satisfies the three properties given in steps 1, 2, and 3, halt; otherwise, return to step 1.

Each application of a lemma in the algorithm may disrupt some property that other two lemmas preserve; however, we also know that each step in the algorithm either does not change P or reduces the rank of P by at least 1. Since the original input P has rank $O(1)$, the algorithm must terminate in $O(1)$ steps, giving us a generalized arithmetic progression of rank $\mathfrak{r}$ that satisfies Inequalities (37) and (38), satisfies conditions (i) and (ii) of Theorem 6.1, and satisfies the condition that $\Phi_P(\{\tilde{a}_j : 1 \le j \le n\}$ spans all of $\mathbb{R}^{\mathfrak{r}}$.

Thus, all that is left to prove is part (iii), the claim of rational T-commensurability. Though we will not need it in the current section, one should recall that Theorem 6.1 is only useful when $\left|n^{o(n)}T^{O(1)}\right| = n^{o(n)}$, where T is the symmetric generalized arithmetic progression containing $\{-1, 0, 1\}$ and all possible values taken by the $\beta_{ij}^{(\mu)}$ and the α_{ij} (see Section 6).

We say that a set W *economically T-spans* a set U if each $u \in U$ can be represented as a highly T-rational linear combination of elements in W, where each coefficient may be expressed as a/b where $a, b \in n^{o(n)}T^{O(1)}$ and where the implicit constants in the $o(\cdot)$ and $O(\cdot)$ notation are uniform over U.

Comparing our definitions with those from [11, Section 8], we note that "highly rational" means the same thing as "highly $\{-1, 0, 1\}$-rational", and "economically spans" means the same thing as

"economically $\{-1,0,1\}$-spans". Thus, it is clear that any highly rational number is also highly T-rational for any T containing $\{-1,0,1\}$, and also the statement "W economically spans U" implies "W economically T-spans U" for any set T containing $\{-1,0,1\}$. The remainder of this section paraphrases (with some notational changes) the latter portion of [11, Section 8].

We know that $\Phi_P(\{\tilde{a}_j : 1 \leq j \leq n\}$ spans $\mathbb{R}^{\mathfrak{r}}$. Thus, there exists a subset $U \subset \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ of cardinality $\mathfrak{r}$ such that $\Phi_P(U)$ spans $\mathbb{R}^{\mathfrak{r}}$. Renumbering if necessary, we can write $U = \{\tilde{a}_1, \ldots, \tilde{a}_{\mathfrak{r}}\}$. It will be important later on that U has cardinality $O(1)$.

The set $\{v_1, \ldots, v_{\mathfrak{r}}\}$ of basis vectors for P economically $\{-1,0,1\}$-spans $\{\tilde{a}_1, \ldots, \tilde{a}_n\}$ by the definition of P (note that $M_i \leq O(\Pr(X_{i_{\max}} \in V)^{-1}) \leq O(p^{-n}) = n^{o(n)}$), and so by Cramer's rule, the vectors $\Phi_P(U)$ economically $\{-1,0,1\}$-span the standard basis vectors $\{e_1, \ldots, e_{\mathfrak{r}}\}$ for $\mathbb{R}^{\mathfrak{r}}$. Applying Φ_P^{-1} (recall that Φ_P is a bijection since P is proper) shows that U economically $\{-1,0,1\}$-spans $\{v_1, \ldots, v_{\mathfrak{r}}\}$.

Following this paragraph, we will show that there exists a single vector v_{i_0} where $1 \leq i_0 \leq \mathfrak{r}$ such that v_{i_0} economically T-spans U, which will show by transitivity that v_{i_0} economically T-spans $\{\tilde{a}_1, \ldots, \tilde{a}_n\}$ (since U economically T-spans $\{v_1, \ldots, v_{\mathfrak{r}}\}$ which economically T-spans $\{\tilde{a}_1, \ldots, \tilde{a}_n\}$; the relation "economically T-spans" is transitive here since the sets U and $\{v_1, \ldots, v_{\mathfrak{r}}\}$ have cardinality $O(1)$).

Let s be the smallest integer such that there exists a subset of cardinality s of $\{v_1, \ldots, v_{\mathfrak{r}}\}$ (by renumbering, say the set is $\{v_1, \ldots, v_s\}$) so that for some nonzero $d \in n^{o(n)}T^{O(1)}$ and some $c_{ij} \in n^{o(n)}T^{O(1)}$ we have

$$d\tilde{a}_i = \sum_{j=1}^{s} c_{ij} v_j \text{ for every } 1 \leq i \leq n. \tag{39}$$

Note that d does not depend on i, and so this statement is slightly stronger than having $\{v_1, \ldots, v_s\}$ economically T-span $\{\tilde{a}_1, \ldots, \tilde{a}_n\}$. Also, note that Equation (39) holds (for example) with $s = \mathfrak{r}$ by the definition of P and since T contains $\{-1,0,1\}$.

We now consider two cases:

- The $n \times s$ matrix $C = (c_{ij})$ has rank 1 in $\mathbb{Z}/Q\mathbb{Z}$. In this case, $\tilde{a}_{i_1}/\tilde{a}_{i_2}$ is highly T-rational for all i_1, i_2 (Since all the c_{ij} are highly T-rational). We know that U economically T-spans $\{v_1, \ldots, v_{\mathfrak{r}}\}$, and so the numbers v_{i_1}/v_{i_2} are also highly T-rational (note that it is critical here that U has cardinality $O(1)$). This means that v_1 (for example) economically T-spans $\{v_1, \ldots, v_{\mathfrak{r}}\}$, and so by transitivity v_1 economically T-spans U.

- The matrix C has rank at least 2. Recall that $(a_1, \ldots, a_n)$ is the normal vector for V and that V is spanned by $(n-1)$ linearly independent vectors with entries in S (recall that S contains all possible values taken by the α_{ij}). We can scale the j-th coordinate of each of these vectors by $b_{j,1}^{-1}$ to get a set of $n-1$ linearly independent vectors each of which is orthogonal to $\tilde{a} := (\tilde{a}_1, \ldots, \tilde{a}_n)$. Among these $(n-1)$ linearly independent vectors that are orthogonal to $(\tilde{a}_1, \ldots, \tilde{a}_n)$, we can find at least one, say $w = (b_{1,1}^{-1}w_1, \ldots, b_{n,1}^{-1}w_n)$ that is not orthogonal to every column of C (since C has column rank at least 2). Let $B := \{b_{j,1} : 1 \leq j \leq n\}$, and let $\tilde{w} := w\prod_{b \in B} b = (\tilde{w}_1, \ldots, \tilde{w}_n)$. Thus $\tilde{w}$ is orthogonal to $\tilde{a}$ and every coordinate $\tilde{w}_i$ of $\tilde{w}$ is an element of $T^{O(1)}$ (since T contains S and B and $|B| = O(1)$ by the definition of p-bounded of exponent r).

Remark 8.5. Note that the line above is the only place in the proof where we use the assumption from the definition of p-bounded of exponent r that the $\beta_{ij}^{(\mu)}$ take values in a set with cardinality $O(1)$. As is evidenced here, the following weaker assumption suffices instead: say that for each $1 \leq i \leq n$ there exists a set B_i such that $|B_i| = O(1)$ and such that $\beta_{i1}^{(\mu)}, \beta_{i2}^{(\mu)}, \ldots, \beta_{in}^{(\mu)}$ each take a nonzero value in B_i with probability at least q. In fact, this weaker assumption also replaces the assumption in the definition of p-bounded of exponent r that $q \leq \min_x \Pr(\beta_{ij}^{(\mu)} = x)$ for every i, j: It suffices for each $\beta_{ij}^{(\mu)}$ to take one value in B_i with probability at least q, instead of taking every value with probability at least q.

We may now compute:

$$0 = d\tilde{a} \cdot \tilde{w} = \sum_{i=1}^{n} d\tilde{a}_i \tilde{w}_i = \sum_{i=1}^{n} \sum_{j=1}^{s} c_{ij} v_j \tilde{w}_i = \sum_{j=1}^{s} \left(\sum_{i=1}^{n} c_{ij} \tilde{w}_i \right) v_j.$$

Since $\tilde{w}$ is not orthogonal to every column of $C = (c_{ij})$, we can assume (reordering if necessary), that the coefficient for v_s above is nonzero, and thus we have

$$v_s = \frac{-1}{\sum_{\ell=1}^{n} c_{\ell s} \tilde{w}_\ell} \sum_{j=1}^{s-1} \left(\sum_{\ell=1}^{n} c_{\ell j} \tilde{w}_\ell \right) v_j.$$

Plugging this last equation into Equation (39), we arrive at

$$d \left(\sum_{\ell=1}^{n} c_{\ell s} \tilde{w}_\ell \right) \tilde{a}_i = \sum_{j=1}^{s-1} \left(c_{ij} \sum_{\ell=1}^{n} c_{\ell s} \tilde{w}_\ell - c_{is} \sum_{\ell=1}^{n} c_{\ell j} \tilde{w}_\ell \right) v_j.$$

Since the coefficient for $\tilde{a}_i$ on the left is an element of $n^{o(n)} T^{O(1)}$ and the coefficient for each v_j on the right is an element of $n^{o(n)} T^{O(1)}$, we have contradicted the minimality of s.

Thus, we have completed the proof of the structure theorem (Theorem 6.1). □

9 A generalization: ϝ rows have fixed, non-random values

In this section, we will give a generalization of Theorem 2.2 to the case where the random matrix N_n has $\mathfrak{f} \leq O(\ln n)$ rows that are assumed to be linearly independent and contain fixed, non-random entries. The proof of the generalized result is very similar to the proof of Theorem 2.2, and we will sketch the main differences in the two proofs below.

Definition 9.1 (a random matrix $N_{\mathfrak{f},n}$ with entries in S)**.** Let $\mathfrak{f}$ be an integer between 1 and n, let S be a subset of a ring, and let $N_{\mathfrak{f},n}$ be an n by n matrix defined as follows. For $1 \leq i \leq \mathfrak{f}$ and $1 \leq j \leq n$, let the entries s_{ij} of $N_{\mathfrak{f},n}$ be fixed (non-random) elements of S such that the rows $(s_{i,1}, \ldots, s_{i,n})$ for $1 \leq i \leq \mathfrak{f}$ are linearly independent. For $\mathfrak{f} + 1 \leq i \leq n$ and $1 \leq j \leq n$, let the

entries α_{ij} of $N_{\mathfrak{f},n}$ be discrete finite random variables taking values in S. Thus,

$$N_{\mathfrak{f},n} := \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,n} \\ \vdots & \cdots & \cdots & \vdots \\ s_{\mathfrak{f},1} & \cdots & \cdots & s_{\mathfrak{f},n} \\ \alpha_{\mathfrak{f}+1,1} & \alpha_{\mathfrak{f}+1,2} & \cdots & \alpha_{\mathfrak{f}+1,n} \\ \alpha_{\mathfrak{f}+2,1} & \alpha_{\mathfrak{f}+2,2} & \cdots & \alpha_{\mathfrak{f}+2,n} \\ \alpha_{\mathfrak{f}+3,1} & \alpha_{\mathfrak{f}+3,2} & \cdots & \alpha_{\mathfrak{f}+3,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} \text{Fixed rows; assumed to be linearly independent} \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \\ a \\ a \end{matrix}}\right\} \text{Random rows} \end{matrix}$$

Theorem 9.2. *Let p be a positive constant such that $0 < p < 1$, let r be a positive integer constant, and let S be a generalized arithmetic progression in the complex numbers with rank $O(1)$ (independent of n) and with cardinality at most $|S| \leq n^{o(n)}$. Consider the matrix $N_{\mathfrak{f},n}$ with entries in S (see Definition 9.1 above), where $\mathfrak{f} \leq \left(\frac{r}{2\ln(1/p)} - o(1)\right) \ln n$. If the collection of random variables $\{\alpha_{jk}\}_{\mathfrak{f}+1 \leq j \leq n, 1 \leq k \leq n}$ is p-bounded of exponent r, then*

$$\Pr(N_{\mathfrak{f},n} \text{ is singular}) \leq \max\left\{(p^{1/r} + o(1))^n, (p + o(1))^{n-\mathfrak{f}}\right\}.$$

Note that the bound on the singularity probability of $N_{\mathfrak{f},n}$ for $r \geq 2$ is the same as in Theorem 2.2 (since for $r \geq 2$, we have $n/r \ll n - c\ln n = n - \mathfrak{f}$). This is a reflection of the fact that only the large dimension case uses the randomness in all the rows simultaneously, and in that case the exponential bound does not depend on r. Generally speaking, the best known lower bounds on the singularity probability of a discrete random matrix come from a dependency among at most two random rows, and since $N_{\mathfrak{f},n}$ certainly has more than two random rows, the upper bounds given in Theorem 9.2 seem reasonable.

Theorem 9.2 leads to Corollary 1.2 by following a conditioning argument very similar to that given in Section 3.3.

9.1 Outline of the proof of Theorem 9.2

The proof of Theorem 9.2 follows the same lines of reasoning as that of Theorem 2.2. In this subsection, we will state the main lemmas with the necessary modifications, and we will mention a few important considerations when making the modifications.

Note that Equation (15), which reduces the question of singularity to one of the rows spanning non-trivial hyperplane of dimension $n-1$ holds in the current context, using the same definition of A_V and "non-trivial hyperplane" (both are defined after Equation (15) in Section 5.1).

Definition 9.3 (combinatorial dimension with $\mathfrak{f}$ fixed rows)**.** Let $\mathcal{D} := \left\{\frac{a}{n} : 0 \leq a \leq n^2, a \in \mathbb{Z}\right\}$. For any $d_\pm \in \mathcal{D}$, we define the *combinatorial Grassmannian* $\mathrm{Gr}_{\mathfrak{f}}(d_\pm)$ to be the set of all non-trivial hyperplanes V in $(\mathbb{Z}/Q\mathbb{Z})^n$ such that

$$p^{n-d_\pm+1/n} < \max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V) \leq p^{n-d_\pm}.$$

For $d_\pm = 0$, we define $\mathrm{Gr}_{\mathfrak{f}}(0)$ to be the set of all non-trivial hyperplanes such that

$$\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V) \leq p^n.$$

We will refer to $d_\pm$ as the *combinatorial dimension* of V.

Lemma 9.4 (Small combinatorial dimension, with $\mathfrak{f}$ fixed rows)**.** *For any $\delta > 0$ we have*

$$\sum_{d_\pm \in \mathcal{D} \text{ s.t. } T^{d_\pm} q^n \le \delta^n} \; \sum_{V \in \mathrm{Gr}_\mathfrak{f}(d_\pm)} \Pr(A_V) \le (n-\mathfrak{f})\delta^n.$$

Proof. The proof is the same as that for Lemma 5.2; also see [4], [10], [11]. □

Lemma 9.5 (Large combinatorial dimension, with $\mathfrak{f}$ fixed rows)**.** *We have*

$$\sum_{d_\pm \in \mathcal{D} \text{ s.t. } \frac{c_{\mathrm{LgDim}}}{n^{1/2}} \le T^{d_\pm} q^n} \; \sum_{V \in \mathrm{Gr}_\mathfrak{f}(d_\pm)} \Pr(A_V) \le (p + o(1))^{n-\mathfrak{f}}$$

Here, c_{LgDim} is the same as in Lemma 5.3.

Proof. The proof is the same as that for Lemma 5.3, except now we appeal to Lemma A.2 with $\mathfrak{f} > 0$. Note that we must assume $\mathfrak{f} \le n/2$ in order to apply Lemma A.2. See also [4],[10],[11]. □

Proposition 9.6 (Medium combinatorial dimension estimate, with $\mathfrak{f}$ fixed rows)**.** *Let $0 < \epsilon_0$ be a constant much smaller than 1, and let $d_\pm \in \mathcal{D}$ be such that $(p + c_{\mathrm{MedDim},\mathfrak{f}}\epsilon_0)^{n/r} < T^{d_\pm} q^n < \frac{c_{\mathrm{LgDim}}}{\sqrt{n}}$. If $\mathfrak{f} \le \left(\frac{r}{2\ln(1/p)} - o(1)\right)\ln n$, then*

$$\sum_{V \in \mathrm{Gr}_\mathfrak{f}(d_\pm)} \Pr(A_V) \le (p + o(1))^{n/r}.$$

Here we choose the constant $c_{\mathrm{MedDim},\mathfrak{f}}$ so that $c_{\mathrm{MedDim},\mathfrak{f}} > \left(c_m + c_\mathfrak{f} + \frac{1}{100}\right)$, where c_m and $c_\mathfrak{f}$ are positive absolute constants (in particular, we need $c_\mathfrak{f}$ such that $\mathfrak{f} \le \frac{c_\mathfrak{f}\epsilon_0 n}{r}$, which is true for any positive constant $c_\mathfrak{f}$ since $\mathfrak{f} \le O(\ln n)$). As before, we will prove this proposition by separating V with medium combinatorial dimension into two cases: exceptional and unexceptional, which are defined below using the definition of $Z^*_{i,k}$ from Equation (17) (this definition is the same as in Definition 5.5 with the small change that i and j are required to be between $\mathfrak{f}+1$ and n instead of between 1 and n).

Definition 9.7. Consider a hyperplane V of medium combinatorial dimension (that is, $d_\pm$ satisfies the condition in Proposition 9.6). We say V is *unexceptional* if there exists an i_0 where $\mathfrak{f}+1 \le i_0 \le n$ and there exists a k_0 where $1 \le k_0 \le r$ such that

$$\max_{\mathfrak{f}+1 \le j \le n} \{\Pr(X_j \in V)\} < \epsilon_1 \Pr(Z^*_{i_0,k_0} \in V).$$

We say V is *exceptional* if for every i where $\mathfrak{f}+1 \le i \le n$ and for every k where $1 \le k \le r$ we have

$$\epsilon_1 \Pr(Z^*_{i,k} \in V) \le \max_{\mathfrak{f}+1 \le j \le n} \{\Pr(X_j \in V)\}. \tag{40}$$

In particular, there exists $i_{\max}$ such that $\Pr(X_{i_{\max}} \in V) = \max_{\mathfrak{f}+1 \le j \le n}\{\Pr(X_j \in V)\}$; and so if V is exceptional, then

$$\epsilon_1 \Pr(Z^*_{i_{\max},k} \in V) \le \Pr(X_{i_{\max}} \in V) \qquad \text{for every } k. \tag{41}$$

We will refer to $X_{i_{\max}}$ as the *exceptional row*.

Lemma 9.8 (Unexceptional space estimate, with $\mathfrak{f}$ fixed rows)**.** *If $\mathfrak{f} \leq \frac{c_{\mathfrak{f}}\epsilon_0 n}{r}$ for some positive constant $c_{\mathfrak{f}}$, then we have*

$$\sum_{V \in \mathrm{Gr}_{\mathfrak{f}}(d_\pm): V \text{ is unexceptional}} \Pr(A_V) \leq p^{-o(n)} 2^n \epsilon_1^{c_m \epsilon_0 n/r}.$$

Notice that the bound is the same as in Lemma 5.6, except that we replaced c_{MedDim} with $c_{\mathrm{MedDim},\mathfrak{f}}$ when defining "unexceptional".

Proof. The proof follows in the same way as that for Lemma 5.6; however, when replacing rows X_i of $N_{\mathfrak{f},n}$ with rows Z_i that concentrate more sharply on V, we must take care to only replace random rows of $N_{\mathfrak{f},n}$ (i.e., rows $X_1, \ldots, X_{\mathfrak{f}}$ must not be replaced by Z_i). See Appendix B for details. □

In the exceptional case, The same structure theorem (Theorem 6.1) holds, leading to the following lemma.

Lemma 9.9 (Exceptional space estimate, with $\mathfrak{f}$ fixed rows)**.** *If $\mathfrak{f} \leq \left(\frac{r}{2\ln(1/p)} - o(1)\right) \ln n$, then*

$$\sum_{V \in \mathrm{Gr}(d_\pm): V \text{ is exceptional}} \Pr(A_V) \leq p^{n/r} \tag{42}$$

Note that this upper bound is dramatically worse than the analogous upper bound in Lemma 5.7 of $n^{-\frac{n}{2}+o(n)}$.

Proof. As in Lemma 5.7, the main step in the proof is applying the structure theorem (Theorem 6.1). In the current context, Inequality (20) holds with $n-\mathfrak{f}$ as the exponent instead of n (since there are only $n-\mathfrak{f}$ random rows). If we combine this modified version of Inequality (20) with Inequality (21), then we have the bound

$$\begin{aligned}\sum_{\substack{V \in \mathrm{Gr}_{\mathfrak{f}}(d_\pm): \\ V \text{ is exceptional}}} \Pr(A_V) &\leq n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-n} \Pr(X_{i_{\max}} \in V)^{n-\mathfrak{f}} \\ &= n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-\mathfrak{f}},\end{aligned}$$

where by assumption $X_{i_{\max}}$ is the random row such that $\Pr(X_{i_{\max}} \in V) = \max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)$. In order for this upper bound to achieve the desired bound in Inequality (42), it is sufficient to have

$$n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-\mathfrak{f}} \leq p^{n/r}. \tag{43}$$

Using the assumption that $\Pr(X_{i_{\max}} \in V) \geq (p + c_{\mathrm{MedDim},\mathfrak{f}}\epsilon_0)^{n/r} > p^{n/r}$ (since V is of medium combinatorial dimension), we see that Inequality (43) holds whenever

$$\mathfrak{f} \leq \left(\frac{r}{2\ln(1/p)} - o(1)\right) \ln n,$$

which completes the proof. □

A Two background results

A.1 A version of the Littlewood-Offord result in $\mathbb{Z}/Q\mathbb{Z}$

If $S \subset \mathbb{Q}$, then we can clear denominators and prove (as in [11, Lemma 2.4]) the large combinatorial dimension estimate in $\mathbb{R}$ instead of working in $\mathbb{Z}/Q\mathbb{Z}$, in which case we can also use the Littlewood-Offord result over $\mathbb{R}$ (see [12, Corollary 7.13]), instead of the version over $\mathbb{Z}/Q\mathbb{Z}$ given here in Lemma A.1. When working in $\mathbb{R}$, the integral approximation of Inequality (44) can be replaced by a limit going to infinity, and we do not need any extra assumptions on Q. In particular, we may take $Q \approx \exp(\exp(Cn))$ (see Remark 4.2).

For Q sufficiently large with respect to q, r, and n, it is clear that we have

$$\frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}(1-2q+2q\cos(2\pi\xi/Q))^{k/r} \leq \int_0^1 (1-2q+2q\cos(2\pi t))^{k/r}\,dt + \frac{1}{n}, \tag{44}$$

for all $1 \leq k \leq n$.

Lemma A.1. *Let Q be sufficiently large to satisfy Inequality (44), and let $v_1, \dots, v_n \in \mathbb{Z}/Q\mathbb{Z}$ be such that $v_1, \dots, v_k$ are nonzero. Let $\{\alpha_j\}_{j=1}^n$ be a collection of random variables that are p-bounded of exponent r, and let $X_{\mathbf{v}} := \alpha_1 v_1 + \cdots + \alpha_n v_n$. Then, for every $x \in \mathbb{Z}/Q\mathbb{Z}$ we have*

$$\Pr(X_{\mathbf{v}} = x) \leq \frac{c_{\mathrm{LO}}\sqrt{r}}{\sqrt{qk}} = O\left(\frac{1}{\sqrt{k}}\right),$$

where c_{LO} is an absolute constant.

Proof. Our proof is closely modeled on the proof of [12, Corollary 7.13]. Let $\beta_j^{(\mu)}$ be the symmetric random variables from the definition of p-bounded of exponent r corresponding to α_j (see Equation (9)). Then, we can compute

$$\begin{aligned}
\Pr(X_{\mathbf{v}} = x) &\leq \frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}\prod_{j=1}^{k}|\mathbb{E}(e_Q(\alpha_j a_j \xi))| && \text{(note that } a_j = 0 \text{ for } j > k\text{)}\\
&\leq \prod_{j=1}^{k}\left(\frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}|\mathbb{E}(e_Q(\alpha_j a_j \xi))|^k\right)^{1/k} && \text{(Hölder's inequality)}\\
&\leq \frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}|\mathbb{E}(e_Q(\alpha_{j_0} a_{j_0} \xi))|^k && \text{(where } j_0 \text{ corresponds to the largest factor in the previous line)}\\
&\leq \frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}\left(1-\mu+\mu\sum_{s=1}^{\ell_{j_0}} p_{j_0,s}\cos(2\pi b_{j_0,s} v_{j_0}\xi/Q)\right)^{k/r} && \text{(since } \alpha_{j_0} \text{ is } p\text{-bounded of exponent } r\text{)}\\
&\leq \frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}(1-2q+2q\cos(2\pi b_{j_0,1} v_{j_0}\xi/Q))^{k/r} && \text{(since } \mu p_{j_0,1} \geq 2q\text{)}\\
&= \frac{1}{Q}\sum_{\xi\in\mathbb{Z}/Q\mathbb{Z}}(1-2q+2q\cos(2\pi\xi/Q))^{k/r} && \text{(by reordering the sum) .}
\end{aligned}$$

Combining the above inequalities with Inequality (44) and following the proof of [12, Corollary 7.13] to bound the integral, we have

$$\begin{aligned}\Pr(X_{\mathbf{v}} = x) &\le \int_0^1 \left(1 - 2q + 2q\cos(2\pi t)\right)^{k/r}\, dt + \frac{1}{n} \\ &= \frac{c_{\mathrm{LO}}\sqrt{r}}{\sqrt{qk}} = O\left(\frac{1}{\sqrt{k}}\right),\end{aligned}$$

where c_{LO} is an absolute constant. □

A.2 A generalization of a lemma due to Komlós [6]

This lemma is a generalization of the result in [6] (see also [2, Lemma 14.10], [4, Section 3.1], and [10, Lemma 5.3]).

Lemma A.2. *Fix n, and let p be a positive constants such that $0 < p < 1$ and let r be a positive integer constant. Consider the matrix $N_{\mathfrak{f},n}$ taking values in $\mathbb{Z}/Q\mathbb{Z}$, where $\mathfrak{f} \le n/2$ and Q is large enough to satisfy Inequality (44). If the collection of random entries in $N_{\mathfrak{f},n}$ is p-bounded of exponent r, then*

$$\Pr\Big(\textit{there exists } \mathbf{v} \in \Omega_1 \textit{ such that } N_{\mathfrak{f},n} \cdot \mathbf{v} = 0\Big) \le (p + o(1))^{n-\mathfrak{f}},$$

where

$$\Omega_1 := \Big\{(v_1, \ldots, v_n) \in \mathbb{Z}/Q\mathbb{Z} : \textit{ at most } (n-\mathfrak{f})\left(1 - \frac{c}{\ln n}\right) + 1 \textit{ of the } v_i \textit{ are nonzero}\Big\} \setminus \{\underline{0}\},$$

where the constant c can be taken to be $c \ge 2\ln(100/p)$, and where $\underline{0}$ denotes the zero vector.

Proof. Let $E_k = \{$there exists $v \in \Omega_1$ with at most k nonzero coordinates such that $N_{\mathfrak{f},n} \cdot v = 0\}$. Clearly,

$$\Pr\Big(\text{there exists } v \in \Omega_1 \text{ such that } N_{\mathfrak{f},n} \cdot v = 0\Big) \le \sum_{1 \le k \le (n-\mathfrak{f})\left(1 - \frac{c}{\ln n}\right) + 1} \Pr(E_k \setminus E_{k-1}).$$

Let S be the set of all possible values that could appear as entries in $N_{\mathfrak{f},n}$, and let $N_{\mathfrak{f},n}|_{j_1,\ldots,j_k}$ be the n by k matrix consisting of columns $j_1, \ldots, j_k$ of $N_{\mathfrak{f},n}$. Following [6, Lemma 2] (see also [2, Lemma 14.10] and [10, Lemma 5.3]) we can write

$$\Pr(E_k \setminus E_{k-1}) \le \sum_{\substack{1 \le j_1 < \cdots \\ \cdots < j_k \le n}} \sum_{\substack{1 \le i_1 < \cdots \\ \cdots < i_{k-1} \le n}} \sum_{\substack{H \text{ a } (k-1)\text{-dimensional} \\ \text{hyperplane} \\ \text{spanned by } S^k}} \Pr(\mathrm{RwSpn}_{i_1,\ldots,i_{k-1},H}) \Pr(\mathrm{RwIn}_{i_1,\ldots,i_{k-1},H}),$$

where

$$\begin{aligned}\mathrm{RwSpn}_{i_1,\ldots,i_{k-1},H} &:= \Big\{\text{rows } i_1, \ldots, i_{k-1} \text{ of } N_{\mathfrak{f},n}|_{j_1,\ldots,j_k} \text{ span } H\Big\}, \text{ and} \\ \mathrm{RwIn}_{i_1,\ldots,i_{k-1},H} &:= \Big\{\text{all rows of } N_{\mathfrak{f},n}|_{j_1,\ldots,j_k} \text{ except } i_1, \ldots, i_{k-1} \text{ are in } H\Big\}.\end{aligned}$$

Let $U(k,p,q)$ be a uniform upper bound for $\Pr(\text{row } i \text{ is in } H)$, where $\mathfrak{f}+1 \le i \le n$ and q is the constant from Definition 2.1 (here, we mean uniform with respect to the index sets $\{j_1,\dots,j_k\}$ and $\{i_1,\dots,i_k\}$). Then we have

$$\Pr(E_k \setminus E_{k-1}) \le U(k,p,q)^{n-k-\mathfrak{f}+1}\binom{n}{k}\binom{n}{k-1},$$

since $k-1$ fixed rows of $N_{\mathfrak{f},n}|_{j_1,\dots,j_k}$ can span at most 1 hyperplane H of dimension $k-1$.

For $k \le \dfrac{2^8 c_{\mathrm{LO}}^2 r}{p^2 q}$ (a constant), we can set $U(k,p,q) = p$ by the Weighted Odlyzko Lemma (see Lemma B.1), giving us a bound of

$$\Pr(E_k \setminus E_{k-1}) \le (p+o(1))^{n-\mathfrak{f}}. \tag{45}$$

For $\frac{2^8 c_{\mathrm{LO}}^2 r}{p^2 q} < k \le (n-\mathfrak{f})\left(1-\dfrac{c}{\ln n}\right)+1$, we use Lemma A.1 to set $U(k,p,q) = \frac{c_{\mathrm{LO}}\sqrt{r}}{\sqrt{qk}}$. Since $\binom{n}{k}\binom{n}{k-1} \le \frac{2^{2n}}{n}$ we thus have

$$\Pr(E_k \setminus E_{k-1}) \le \frac{1}{n}2^{2n}\left(\frac{c_{\mathrm{LO}}^2 r}{qk}\right)^{\frac{n-k-\mathfrak{f}+1}{2}}.$$

As a function of k, this upper bound has strictly positive second derivative; thus, the largest upper bound will occur at one of the extremal values of $k = \frac{2^8 c_{\mathrm{LO}}^2 r}{p^2 q}$ or $k = (n-\mathfrak{f})\left(1-\frac{c}{\ln n}\right)+1$, and a bit of computation shows that

$$\Pr(E_k \setminus E_{k-1}) \le \frac{1}{n}O(p^{n-\mathfrak{f}}). \tag{46}$$

Summing the bounds in Inequalities (45) and (46) completes the proof. □

B The unexceptional case with $\mathfrak{f}$ fixed rows

This section is adapted from the proof of [11, Lemma 4.1], and proves Lemma 5.6 by setting $\mathfrak{f}=0$. Assume that $\mathfrak{f} \le \frac{c_{\mathfrak{f}}\epsilon_0 n}{r}$, and let m be the closest integer to $\frac{c_m \epsilon n}{r}$. Let $Z_1,\dots,Z_m$ be i.i.d. copies of the unexceptional row vector $Z^*_{i_0,k_0}$ from Definition 9.7, so $\epsilon_1 \Pr(Z_i \in V) > \Pr(X_i \in V)$ for all $\mathfrak{f}+1 \le i \le n$. We will need the following version of the Weighted Odlyzko Lemma:

Lemma B.1. [cf. [11, Lemma 4.3] or [4, Section 3.2]] *For $1 \le i$, let W_{i-1} be an $(\mathfrak{f}+i-1)$-dimensional subspace containing $X_1,\dots,X_{\mathfrak{f}}$ (which are fixed, linearly independent row vectors). Then*

$$\Pr(Z_i \in W_{i-1}) \le \left(p+\frac{\epsilon_0}{100}\right)^{\frac{n}{r}-\mathfrak{f}-i+1}.$$

Proof. Since W_{i-1} has dimension $\mathfrak{f}+i-1$, there exists a set of $\mathfrak{f}+i-1$ "determining" coordinates such that if a vector $V \in W_{i-1}$, then the $\mathfrak{f}+i-1$ "determining" coordinates determine the values of the remaining $n-\mathfrak{f}-i+1$ coordinates. Since the maximum probability that any of the n/r random coordinates in Z_i takes a given value is at most $1-\underline{\mu} = p+\frac{\epsilon_0}{100}$, and since there are at least $\frac{n}{r}-\mathfrak{f}-i+1$ of the random coordinates in Z_i that are not among the "determining" coordinates, we have the desired upper bound. □

Let $V_0 := \mathrm{Span}\{X_1, \ldots, X_\mathfrak{f}\}$, the space spanned by the $\mathfrak{f}$ fixed rows, and for $1 \leq i \leq m$ let $B_{V,i}$ be the event that $Z_1, \ldots, Z_m$ are linearly independent in $V \setminus V_0$. We have the following analog of Lemma 5.8 (and also [11, Lemma 4.4]):

Lemma B.2 (see Lemma 4.4 in [11])**.** *Let m, $\mathfrak{f}$, and $B_{V,m}$ be as defined above. Then,*

$$\Pr\left(B_{V,m}\right) \geq p^{o(n)} \left(\frac{\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1}\right)^m$$

Proof. Using Bayes' Identity, we have

$$\Pr(B_{V,m}) = \prod_{i=1}^{m} \Pr(B_{V,i} | B_{V,i-1}), \tag{47}$$

where $B_{V,0}$ denotes the full space of the Z_i. Conditioning on a particular instance of $Z_1, \ldots, Z_{i-1}$ in $B_{V,i-1}$, we have that

$$\Pr(B_{V,i} | B_{V,i-1}) = \Pr(Z_i \in V) - \Pr(Z_i \in W_{i-1}),$$

where W_{i-1} denotes the $(\mathfrak{f} + i - 1)$-dimensional space spanned by $X_1, \ldots, X_\mathfrak{f}$ and $Z_1, \ldots, Z_{i-1}$. We will now establish a uniform bound that does not depend on which particular instance of $Z_1, \ldots, Z_{i-1}$ in $B_{V,i-1}$ that we fixed by conditioning. By the definition of unexceptional, we have

$$\Pr(Z_i \in V) > \frac{1}{\epsilon_1} \max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V),$$

and by the Weighted Odlyzko Lemma (see Lemma B.1), we have

$$\Pr(Z_i \in W_{i-1}) \leq \left(p + \frac{\epsilon_0}{100}\right)^{\frac{n}{r} - \mathfrak{f} - i + 1} \leq \left(p + \frac{\epsilon_0}{100}\right)^{\frac{n}{r}(1 - (c_m + c_\mathfrak{f})\epsilon_0)}.$$

Using Taylor's Theorem with remainder (for example), one can show that

$$\left(p + \frac{\epsilon_0}{100}\right)^{\frac{n}{r}(1-(c_m+c_\mathfrak{f})\epsilon_0)} \leq \frac{1}{2n} \left(p + c_{\mathrm{MedDim}} \epsilon_0\right)^{n/r} \leq \frac{1}{n} \max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V),$$

so long as $c_{\mathrm{MedDim}} > \frac{1}{100} + c_m + c_\mathfrak{f} > \frac{1}{100} + (c_m + c_\mathfrak{f}) p \ln\left(\frac{1}{p}\right)$ and n is sufficiently large (the second inequality in the display above is the definition of medium combinatorial dimension).

Thus

$$\Pr(B_{V,i} | B_{V,i-1}) \geq \frac{1}{\epsilon_1} \left(\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)\right) \left(1 - \frac{\epsilon_1}{n}\right),$$

and plugging this estimate back into Inequality (47) we get

$$\Pr\left(B_{V,m}\right) \geq p^{o(n)} \left(\frac{\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1}\right)^m.$$

□

To conclude Lemma 9.8 (which implies Lemma 5.6 by setting $\mathfrak{f} = 0$), we will proceed as in the proof for [11, Lemma 4.1].

Let $Z_1, \ldots, Z_m$ be i.i.d. copies of $Z^*_{i_0,k_0}$ that are independent of the random rows $X_{\mathfrak{f}+1}, \ldots, X_n$. Using independence and Bayes' Identity we have

$$\Pr(A_V) = \Pr(A_V | B_{V,m}) = \frac{\Pr(A_V \wedge B_{V,m})}{\Pr(B_{V,m})} \leq \Pr(A_V \wedge B_{V,m}) p^{-o(n)} \left(\frac{\epsilon_1}{\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)} \right)^m.$$

Because the Z_i are linearly independent in $V \setminus V_0$, we know that there is a subset $I \subset \{\mathfrak{f}+1, \mathfrak{f}+2, \ldots, n\}$ of cardinality $|I| = m$, such that $\{Z_1, \ldots, Z_m\} \cup \{X_i : i \notin I\}$ spans V. Let $C_{V,I}$ be the event that $\{Z_1, \ldots, Z_m\} \cup \{X_i : i \notin I\}$ spans V. Then we have

$$\begin{aligned} \Pr(A_V \wedge B_{V,m}) &\leq \sum_{\substack{I \subset \{\mathfrak{f}+1, \ldots, n\} \\ |I| = m}} \Pr\left(C_{V,I} \wedge \{X_i \in V : i \in I\}\right) \\ &\leq \left(\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V) \right)^m \sum_{\substack{I \subset \{\mathfrak{f}+1, \ldots, n\} \\ |I| = m}} \Pr(C_{V,I}). \end{aligned}$$

Summing the above inequality over all unexceptional V (note that $\sum_V \Pr(C_{V,I}) \leq 1$) and combining with the bound for $\Pr(A_V)$ above gives us

$$\begin{aligned} \sum_{\text{unexceptional } V} \Pr(A_V) &\leq \left(\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V) \right)^m \binom{n-\mathfrak{f}}{m} p^{-o(n)} \left(\frac{\epsilon_1}{\max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V)} \right)^m \\ &\leq p^{-o(n)} 2^n \epsilon_1^m. \end{aligned}$$

This completes the proof of the estimate for unexceptional V.

References

[1] Aizenman, Michael; Germinet, Francois; Klein, Abel; Warzel, Simone. On Bernoulli Decompositions for Random Variables, Concentration Bounds, and Spectral Localization. arXiv:0707.0095v1 [math.PR]. July 1, 2007.

[2] Bollobás, Béla. *Random graphs. Second edition.* Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001.

[3] Bourgain, Jean, Katz, Nets, Tao, Terence. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.* **14** (2004), no. 1, 27–57.

[4] Kahn, Jeff; Komlós, János; Szemerédi, Endre. On the probability that a random ±1-matrix is singular. *J. Amer. Math. Soc.* **8** (1995), no. 1, 223–240.

[5] Komlós, János. On the determinant of (0, 1) matrices. *Studia Sci. Math. Hungar.* **2** (1967), 7–21.

[6] Komlós, János. Circulated manuscript, 1977. edited version available online at: `http://www.math.rutgers.edu/~komlos/01short.pdf`

[7] Martin, Greg; Wong, Erick B.. Almost all integer matrices have no integer eigenvalues. arXiv:0712.3060v1 [math.NT]. December 18, 2007.

[8] Rudelson, M.; Vershynin, R.. The Littlewood-Offord Problem and invertibility of random matrices. *Advances in Mathematics*, to appear. arXiv:math/0703503v1 [math.PR]. March 16, 2007.

[9] Slinko, Arkadii. A generalization of Komlós's theorem on random matrices *New Zealand J. Math.* **30** (2001), no. 1, 81–86.

[10] Tao, Terence; Vu, Van. On random ± 1 matrices: singularity and determinant. *Random Structures Algorithms* **28** (2006), no. 1, 1–23.

[11] Tao, Terence; Vu, Van. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.* **20** (2007), 603–628.

[12] Tao, Terence; Vu, Van. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics (No. 105). Cambridge University Press, Cambridge, 2006.

[13] Vu, Van; Wood, Melanie Matchett; Wood, Philip Matchett. Mapping Incidences. *submitted.*