

RANDOM INTEGRAL MATRICES: UNIVERSALITY OF SURJECTIVITY AND THE COKERNEL

HOI H. NGUYEN AND MELANIE MATCHETT WOOD

ABSTRACT. For a random matrix of entries sampled independently from a fairly general distribution in \mathbf{Z} we study the probability that the cokernel is isomorphic to a given finite abelian group, or when it is cyclic. This includes the probability that the linear map between the integer lattices given by the matrix is surjective. We show that these statistics are asymptotically universal (as the size of the matrix goes to infinity), given by precise formulas involving zeta values, and agree with distributions defined by Cohen and Lenstra, even when the distribution of matrix entries is very distorted. Our method is robust and works for Laplacians of random digraphs and sparse matrices with the probability of an entry non-zero only $n^{-1+\varepsilon}$.

1. INTRODUCTION

For square matrices $M_{n \times n}$ of random discrete entries, the problem to estimate the probability p_n of $M_{n \times n}$ being singular has attracted quite a lot of attention. In the 60's Komlós [16] showed $p_n = O(n^{-1/2})$ for entries $\{0, 1\}$ with probability each $1/2$. This bound was significantly improved by Kahn, Komlós, and Szemerédi in the 90's to $p_n \leq 0.999^n$ for ± 1 entries. About ten years ago, Tao and Vu [30] improved the bound for ± 1 entries to $p_n \leq (3/4 + o(1))^n$. The most recent record is due to Bourgain, Vu and Wood [3] who showed $p_n \leq (\frac{1}{\sqrt{2}} + o(1))^n$ for ± 1 entries and gave exponential bounds for more general entries as well. We also refer the reader to [27] by Rudelson and Vershynin for implicit exponential bounds. For sparse matrices having entries 0 with probability $1 - \alpha_n$, Basak and Rudelson [1] proved $p_n \leq e^{-c\alpha_n n}$ for $\alpha_n \geq C \log n/n$ and for rather general entries, including the adjacency matrices of sparse Erdős-Rényi random graphs.

When $M_{n \times n}$ has integral entries, these results imply that with very high probability the linear map $M_{n \times n} : \mathbf{Z}^n \rightarrow \mathbf{Z}^n$ is injective. Another important property of interest is surjectivity, it seems natural to wonder if with high probability $M_{n \times n} : \mathbf{Z}^n \rightarrow \mathbf{Z}^n$ is surjective (see [18, 21])? However, recent results of the second author show that the surjectivity probability goes to 0 with n (e.g. that is implied by [31, Corollary 3.4]). The main result of this paper will imply that when the matrix has more columns than rows, e.g. $M_{n \times (n+1)} : \mathbf{Z}^{n+1} \rightarrow \mathbf{Z}^n$, we have surjectivity with positive probability strictly smaller than one.

We make the following definition to restrict the types of entries our random matrices will have. We say a random integer ξ_n is α_n -balanced if for every prime p we have

$$\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n \equiv r \pmod{p}) \leq 1 - \alpha_n. \quad (1)$$

Our main result tells us not only whether $M_{n \times (n+u)}$ is surjective, but more specifically about the cokernel $\mathbf{Cok}(M_{n \times (n+u)})$, which is the quotient group $\mathbf{Z}^n/M_{n \times (n+u)}(\mathbf{Z}^{n+u})$ and gives the failure of surjectivity.

Theorem 1.1. *For integers $n, u \geq 0$, let $M_{n \times (n+u)}$ be an integral $n \times (n+u)$ matrix with entries i.i.d copies of an α_n -balanced random integer ξ_n , with $\alpha_n \geq n^{-1+\varepsilon}$ and $|\xi_n| \leq n^T$ for any fixed parameters $0 < \varepsilon < 1$ and $T > 0$ not depending on n . For any fixed finite abelian group B and $u \geq 0$,*

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B) = \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{k=u+1}^{\infty} \zeta(k)^{-1}. \quad (2)$$

Here $\zeta(s)$ is the Riemann zeta function. In particular, as $n \rightarrow \infty$, the map $M_{n \times (n+1)} : \mathbf{Z}^{n+1} \rightarrow \mathbf{Z}^n$ is surjective with probability approaching $\prod_{k=2}^{\infty} \zeta(k)^{-1} \approx 0.4358$. The one extra dimension mapping to \mathbf{Z}^n brought the surjectivity probability from 0 to ≈ 0.4358 .

Note that the product $\prod_{k=u+1}^{\infty} \zeta(k)^{-1}$ in (2) is non-zero for $u \geq 1$, but $\zeta(1)^{-1} = 0$. So Theorem 1.1 shows that every possible finite cokernel appears with positive probability when $u \geq 1$. (Note that when the matrix has full rank over \mathbf{R} , the cokernel must be finite.) Theorem 1.1 is a universality result because these precise positive probabilities do not depend on the distribution of ξ_n , the random entries of our matrices. As a simple example, if we take an $n \times (n+1)$ random matrix with entries all 0 or 1, whether we make entries 0 with probability $\frac{1}{100}$, $\frac{1}{2}$, or $1 - n^{-8/9}$, we obtain the exact same asymptotic probability of the map $\mathbf{Z}^{n+1} \rightarrow \mathbf{Z}^n$ being surjective. If we take entries from $\{-17, 0, 6, 7\}$ with respective probabilities $\frac{2}{3}, \frac{1}{n}, \frac{1}{6} - \frac{1}{n}, \frac{1}{6}$, the asymptotic probability of surjectivity is unchanged. Our theorem allows even more general entries as well.

Further, we prove the following.

Theorem 1.2. *Let $M_{n \times (n+u)}$ be as in Theorem 1.1. We have*

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \text{ is cyclic}) = \prod_{p \text{ prime}} (1 + p^{-(u+1)}(p-1)^{-1}) \prod_{k=u+2}^{\infty} \zeta(k)^{-1}.$$

Note that even when $u = 0$, the limiting probability here is positive. For $u = 0$, this probability has been seen in several papers studying the probability that a random lattice in \mathbf{Z}^n is co-cyclic (gives cyclic quotient), in cases when these lattices are drawn from the nicest, most uniform distributions, e.g. uniform on lattices up to index X with $X \rightarrow \infty$ [4, 23, 24], or with basis with uniform entries in $[-X, X]$ with $X \rightarrow \infty$ [29]. Stanley and Wang have asked whether the probability of having cyclic cokernel is universal (see [29, Remark 4.11 (2)] and [28, Section 4]). Theorem 1.2 proves this universality, showing that the same probability of cocyclicity occurs when the lattice is given by n random generators from a rather large class of distributions, including ones that are rather distorted mod p for each prime p .

Moreover, we show the same results hold if we replace $\mathbf{Cok}(M_{n \times (n+1)})$ with the total sandpile group of an Erdős-Rényi simple random digraph, proving a conjecture of Koplewitz [17, Conjecture 1] (see Theorem 1.6). This allows some dependence in the entries of our random matrices, since the diagonal of the graph Laplacian depends on the other entries in the matrix. In particular, this says that with asymptotic probability $\prod_{k=2}^{\infty} \zeta(k)^{-1} \approx 0.4358$ an Erdős-Rényi random digraph is *co-Eulerian*, which Farrell and Levine [8] define to be any of several equivalent definitions including a simple condition for when chip-firing configurations on the graph stabilize and the condition that recurrent states in the rotor-router model are in a single orbit. In contrast to the distribution of sandpile groups of Erdős-Rényi random graphs, where for each finite abelian group B , the sandpile group is B with asymptotic probability 0 [33, Corollary 9.3], for Erdős-Rényi random *digraphs*, we show that each finite abelian group appears with positive asymptotic probability as the total sandpile group. Moreover, the universality in our theorems proves that all of these positive limiting probabilities do not depend on the edge density of the random graph.

Previous work of the second author [34, Corollary 3.4] determined the probabilities of these $\mathbf{Cok}(M_{n \times (n+u)})$ having any given Sylow p -subgroup for a fixed prime p or finite set of primes p . The two significant advances of this work over previous work are (1) that we determine the distribution of the entire cokernel, not just the part of it related to a finite set of primes, and (2) that we allow our random matrix entries to be more distorted mod p as n increases, for example allowing sparse matrices where entries are non-zero with probability $n^{-1+\epsilon}$. Our hypothesis that $\alpha_n \geq n^{-1+\epsilon}$ is asymptotically best possible, in terms of the exponent of n (see the discussion following Equation (4)).

Our proofs require considering primes in three size ranges separately, and in each range we use different methods. Our work builds on methods from previous work, including that of Tao and Vu [30, 31, 32], the first author and Vu [22], Maples [19], the second author [33, 34], and the first author and Paquette [21]. The key ideas original to this paper are in our treatment of large primes, where we prove a result that lifts structured normal vectors from characteristic p to characteristic 0, crucially for p in a range much smaller than $n^{n/2}$.

1.3. Further results and connections to the literature. We also show asymptotic almost sure surjectivity when $u \rightarrow \infty$ with n , proving a conjecture of Koplewitz [18, Conjecture 2].

Theorem 1.4. Let $M_{n \times (n+u)}$ be as in Theorem 1.1. Then

$$\lim_{\min(u,n) \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq \{id\}) = 1. \quad (3)$$

Theorem 1.1 has several nice corollaries, including the $u \geq 1$ cases of Theorem 1.2 and the following.

Corollary 1.5. For any fixed $u \geq 0$

$$\lim_{n \rightarrow \infty} \mathbf{P}(M_{n \times (n+u)} : \mathbf{Z}^{n+u} \rightarrow \mathbf{Z}^n \text{ is surjective}) = \prod_{k=u+1}^{\infty} \zeta(k)^{-1}.$$

Also, for any fixed $u \geq 1$

$$\lim_{n \rightarrow \infty} \mathbf{P}(\det(M_{n \times (n+u)}) \text{ is square-free}) = \prod_{p \text{ prime}} (1 + p^{-u}(p-1)^{-1}) \prod_{k=u+1}^{\infty} \zeta(k)^{-1}.$$

To give a heuristic for why inverse zeta values arise in these probabilities, note that $M_{n \times (n+1)}$ is surjective if and only if its reduction to modulo p is surjective for all primes p . We then make two idealized heuristic assumptions on $M_{n \times (n+1)}$. (i) (uniformity assumption) Assume that for each prime p the entries of $M_{n \times (n+1)}$ are uniformly distributed modulo p . In this case, a simple calculation gives the probability for $M_{n \times (n+1)}$ being surjective modulo p is $\prod_{j=2}^n (1 - p^{-j})(1 - p^{-n-1})$. (ii) (independence assumption) We next assume that the statistics of $M_{n \times (n+1)}$ reduced to modulo p are asymptotically mutually independent for all primes p . Under these assumptions, as $n \rightarrow \infty$, the probability that $M_{n \times (n+1)}$ is surjective would be asymptotically the product of all of the surjectivity probability modulo p , which leads to the number $\prod_{k=2}^{\infty} \zeta(k)^{-1}$ as seen. The matrices in this paper do not have to satisfy either assumption, and indeed they can violate them dramatically. For example, if the matrix entries only take values 0 and 1, then they cannot be uniformly distributed mod any prime > 2 , and the matrix entries mod 3 are not only not independent from the entries mod 5, but they are in fact determined by the entries mod 5. The work of this paper is in showing that even for rather general random matrices, universality holds and gives the same cokernel distributions as for the simplest random matrices.

For our Theorem 1.1, we remark that for $u \geq 1$, the limiting probabilities $|B|^{-u} |\mathbf{Aut}(B)|^{-1} \prod_{k=u+1}^{\infty} \zeta(k)^{-1}$ in Theorem 1.1 do sum to 1 (use [6, Corollary 3.7 (i)] with $s = u$ and $k = \infty$). This gives, for each $u \geq 1$, a probability distribution on finite abelian groups. Cohen and Lenstra [6] introduced these distributions to conjecture that the $u = 1$ distribution is the distribution of class groups of real quadratic number fields (except for the Sylow 2-subgroup). Friedman and Washington [11] later proved that if $M_{n \times n}$ has independent entries taken from Haar measure on the p -adics \mathbf{Z}_p , then for a finite abelian p -group B we have

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times n}) \simeq B) = |\mathbf{Aut}(B)|^{-1} \prod_{k=u+1}^{\infty} (1 - p^{-k}).$$

The limit is proven by giving an explicit formula for the probability for each n . A similar argument shows that for $M_{n \times (n+u)}$ with independent entries taken from Haar measure on $\widehat{\mathbf{Z}}$, the profinite completion of \mathbf{Z} (these are exactly the matrices with entries that satisfy the two heuristic assumptions above), we have for every finite abelian group B that

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B) = |B|^{-u} |\mathbf{Aut}(B)|^{-1} \prod_{k=u+1}^{\infty} \zeta(k)^{-1}.$$

This is because as $\widehat{\mathbf{Z}} = \prod_p \mathbf{Z}_p$, this Haar measure is the product of the p -adic Haar measures. Building on work of Ekedahl [7], Wang and Stanley [29] find that the cokernels (equivalently, the Smith normal form) of random $n \times m$ matrices for *fixed* n and m and independent, uniform random integer entries in $[-X, X]$ as $X \rightarrow \infty$ match those for entries from Haar measure on $\widehat{\mathbf{Z}}$. While this agreement is easy to see for the Sylow subgroups at any finite set of primes (because X will eventually become larger than all of the primes), it was a substantial problem to prove this agreement for all primes at once.

Our approach to proving Theorem 1.1 and the $u = 0$ case of Theorem 1.2 involves considering three classes of primes (small, medium, and large) separately, and for each class the argument is rather different. For

small primes, we follow the general approach of [34]: finding the moments (which are an expected number of surjections) by dividing the a priori possible surjections into classes and obtaining, for each class, a bound on the number of possible surjections in it and a bound of the probability that any of those surjections are realized. Our advance over [34] is that we can allow sparser matrices, and to obtain this improvement we have to both refine the classes into which we divide surjections and the bounds we have for the probabilities for each class. For medium primes, our starting point is a theorem from [21], which carries out ideas of Maples [19] to show that, modulo a prime, under likely conditions, each time we add a column to our matrix, the probability that the new column is in the span of the previous columns (mod a prime) is near to the probability for a uniform random column. Our contribution is to show the bounds on “likely conditions” and “nearness of probability” can be turned into a bound on how the ranks of the final matrix compare to the ranks of a uniform random matrix. We do this via a rather general approach using a coupling of Markov chains. For large primes, our approach is new. We cannot control whether the rank of our matrix drops by 1 modulo any particular large prime, but considering columns being added one at a time, once the rank drops by 1 modulo a prime, we show that it is not likely to drop again. We do this by showing that the column spaces are unlikely to have a normal vector with many of its coefficients in a generalized arithmetic progression mod p , and then proving a new inverse Erdős-Littlewood-Offord result over finite fields for sparse random vectors based on the method from [22] by the first author and Vu. However, the probabilities of structured normal vectors mod p are still too large to add up over all p , and a key innovation of the paper is that for primes $> e^{n^{1-\varepsilon/3}}$ we show that having a non-zero structured normal vector mod p is equivalent to having one in characteristic 0. Fortunately, the bounds for p up to which we can sum the probabilities of structured normal vectors mod p , and the bounds for p where we can lift structured normal vectors overlap, and this allows us to control the probability of structured normal vectors at all primes.

In contrast to the result of Corollary [1.5], in the $u = 0$ case we are unable to determine $\mathbf{P}(\det(M_{n \times n}) \text{ is square-free})$, though from [34, Corollary 3.4] it follows that

$$\limsup_{n \rightarrow \infty} \mathbf{P}(\det(M_{n \times n}) \text{ is square-free}) \leq \zeta(2)^{-1} \prod_{k \geq 2} \zeta(k)^{-1}, \quad (4)$$

and we would conjecture the limit is equal to this value. We can obtain the limiting probability that $\mathbf{Cok}(M_{n \times n})$ is the product of a given finite abelian group B and a cyclic group (see Theorem [2.4]), and these are currently the most general classes of abelian groups for which we can obtain universality results for $n \times n$ matrices. Even for nicely distributed matrix entries and fixed n , the question of how often $\det(M_{n \times n})$ is square-free is very difficult (see [26, 2]).

Our main results work for $\alpha_n \geq n^{-1+\varepsilon}$, which is asymptotically best possible, in terms of the exponent of n . If the matrix entries are 0 with probability at least $1 - \log n/(n+u)$, then the matrix $M_{n \times (n+u)}$ has a row of all 0's with non-negligible probability, and thus cannot possibly be surjective or even have finite cokernel. We also refer the reader to [21] for some partial results where α_n is allowed to be as small as $O(\log n/n)$ and u is comparable to n . Much of the previous work that we build upon has required the matrix entries to be non-zero with probability bounded away from 0 as $n \rightarrow \infty$. It is perhaps surprising that even as the matrices have entries being 0 more and more frequently, the asymptotic probability that $M_{n \times (n+1)}$ is surjective does not change from $\approx .4358$ as long as $\alpha_n \geq n^{-1+\varepsilon}$.

Another advantage of our method (compared to existing results in the literature on classical random matrix theory) is that the bound on the matrix entries can be as large as any polynomial n^T of n . This can be relaxed somewhat by letting $T \rightarrow \infty$ slowly, but it cannot be lifted entirely as the example of Koplewitz shows [18, Section 4.4] (see also the discussion after Lemma [3.1]).

We now explain in more detail the extension of our results to a natural family of random matrices of dependent entries, namely to the Laplacian of random digraphs. More generally, let $M = M_{n \times n} = (x_{ij})_{1 \leq i,j \leq n}$ be a random matrix where $x_{ii} = 0$ and its off-diagonal entries are i.i.d. copies of an integral random variable ξ_n satisfying (1). A special case here is when M is the adjacency matrix of an Erdős-Rényi simple random digraph $\Gamma \in \overrightarrow{G}(n, q)$ where each directed edge is chosen independently with probability q satisfying $\alpha_n \leq q \leq 1 - \alpha_n$. Let $L_M = (L_{ij})$ be the Laplacian of M , that is

$$L_{ij} = \begin{cases} -x_{ij} & \text{if } i \neq j \\ \sum_{k=1}^n x_{ki} & \text{if } i = j. \end{cases}$$

We then denote S_M (or S_Γ in the case of digraphs) to be the cokernel of L with respect to the group \mathbf{Z}_0^n of integral vectors of zero entry-sum

$$S_M = \mathbf{Z}_0^n / L_M \mathbf{Z}^n.$$

When Γ is a graph, this group has been called the *sandpile group without sink* [8] and the *total sandpile group* [17] of the graph. The size of this group was has been called the *Pham Index* [8], and was introduced by Pham [25] in order to count orbits of the rotor-router operation. We will show that Theorems 1.1 and 1.2 extend to this interesting setting.

Theorem 1.6. *Let $0 < \varepsilon < 1$ and $T > 0$ be given. Let $M_{n \times n}$ be a integral $n \times n$ matrix with zero diagonal and off-diagonal entries i.i.d copies of an α_n -balanced random integer ξ_n , with $\alpha_n \geq n^{-1+\varepsilon}$ and $|\xi_n| \leq n^T$. Then for any finite abelian group B ,*

$$\lim_{n \rightarrow \infty} \mathbf{P}(S_{M_{n \times n}} \simeq B) = \frac{1}{|B| |\mathbf{Aut}(B)|} \prod_{k=2}^{\infty} \zeta(k)^{-1} \quad (5)$$

and

$$\lim_{n \rightarrow \infty} \mathbf{P}(S_{M_{n \times n}} \text{ is cyclic}) = \prod_{p \text{ prime}} (1 + (p^2(p-1))^{-1}) \prod_{k=3}^{\infty} \zeta(k)^{-1}. \quad (6)$$

In particular, every finite abelian group B appears with frequency given in (5) as a total sandpile of the random digraph $\vec{G}(n, q)$ with parameter $n^{-1+\varepsilon} \leq q \leq 1 - n^{-1+\varepsilon}$. In a paper about the sandpile (or chip-firing) and rotor-router models, Holroyd, Levine, Mészáros, Peres, Propp, and Wilson asked if there was an infinite family of non-Eulerian strongly connected digraphs such that the unicycles are in a single orbit of the rotor-router operation [13, Question 6.5]. Pham [25] then gave an infinite family with a single orbit, and asked if the probability of a single orbit for an Erdős-Rényi digraph in fact goes to 1. Koplewitz [17] gave an upper bound on this probability. We have now shown that the desired graphs with a single rotor-router orbit occur with asymptotic probability $\prod_{k=2}^{\infty} \zeta(k)^{-1} \approx 43.58\%$ (matching the upper bound from [17]) among Erdős-Rényi digraphs. Moreover, for every k , our result gives an explicit positive asymptotic probability for exactly k orbits.

Farrell and Levine show that this number of orbits is the size of the total sandpile group [8, Lemma 2.9, Theorem 2.10], and coined the term *co-Eulerian* for digraphs where the total sandpile group is trivial. Farrell and Levine also show that for a strongly connected digraph Γ the algebraic condition $S_\Gamma = \{id\}$ is equivalent to a more combinatorial condition [8, Theorem 1.2], i.e. in this graph a chip configuration σ on Γ stabilizes after a finite number of legal firings if and only if $|\sigma| \leq |E| - |V|$. Further, they prove that minimal length of a multi-Eulerian tour depends inversely on the size of the total sandpile group [9, Theorem 5], showing that $|S_\Gamma|$ measures “Eulerianness” of the graph.

Corollary 1.7. *Let $0 < \varepsilon < 1$ be given and let q be a given parameter such that $n^{-1+\varepsilon} \leq q \leq 1 - n^{-1+\varepsilon}$. Then*

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P}(\vec{G}(n, q) \text{ is co-Eulerian}) &= \\ \lim_{n \rightarrow \infty} \mathbf{P}(\vec{G}(n, q) \text{ is strongly connected, non-Eulerian, and co-Eulerian}) &= \prod_{k=2}^{\infty} \zeta(k)^{-1}. \end{aligned}$$

The corollary follows since $\vec{G}(n, q)$ is strongly connected and non-Eulerian asymptotically almost surely. Although our general method to prove Theorem 1.6 follows the proof method of Theorems 1.1 and 1.2, here the dependency of the entries in each column vector and the non-identical property of the columns pose new challenges. Among other things, for the medium primes we will need to prove a non-i.i.d. analog of the result of [21] that we used in the i.i.d. case. For small primes, when α_n is constant, our results specialize to those of Koplewitz [17], who determined the asymptotic probabilities of given Sylow p -subgroups of these total sandpile groups for finitely many primes p . However, as in our main theorem, we require a further refined method to deal with smaller α_n .

Note that for $M_{n \times n}$ with general i.i.d. α_n -balanced integer entries the results of [1] do not apply to bound the singularity probability. However, a recent result by Paquette and the first author [21] (following the preprint [19] of Maples) shows that the singularity probability p_n can also be bounded in this case by

$e^{-c\alpha_n n}$ with $\alpha_n \geq C \log n/n$ (see also Theorem 5.1). We prove the same bound for the singularity of digraph Laplacians in Corollary 8.16.

1.8. Outline of the paper. In Section 2, we state our results for each class of primes, and show that Theorems 1.1, 1.2, 1.4, and 1.6 follow from these results. We will present our main arguments for the i.i.d. case in Sections 3–7. The main arguments for the graph Laplacian case are in Section 8, building on the treatments in the i.i.d. case.

1.9. Notation. We write \mathbf{P} for probability and \mathbf{E} for expected value. For an event \mathcal{E} , we write $\bar{\mathcal{E}}$ for its complement. We write $\exp(x)$ for the exponential function e^x . We use $[n]$ to denote $\{1, \dots, n\}$. For a given index set $J \subset [n]$ and a vector $X = (x_1, \dots, x_n)$, we write $X|_J$ to be the subvector of X of components indexed from J . Similarly, if H is a subspace then $H|_J$ is the subspace spanned by $X|_J$ for $X \in H$. For a vector $w = (w_1, \dots, w_n)$ we let $\text{supp}(w) = \{i \in [n] | w_i \neq 0\}$. We will also write $X \cdot w$ for the dot product $\sum_{i=1}^n x_i w_i$. We say w is a *normal* vector for a subspace H if $X \cdot w = 0$ for every $X \in H$.

For $0 \leq u \leq n$, the matrix $M_{n \times (n-u)}$ is the submatrix of the first $n-u$ columns of $M_{n \times n}$. Sometimes we will write the Laplacian L_M as $L_{n \times n}$, and so $L_{n \times (n-u)}$ is the submatrix of the first $n-u$ columns of L_M . We also write \mathbf{Z}_0^n/p to denote the set of vectors of zero-entry sum in $(\mathbf{Z}/p\mathbf{Z})^n$.

For a finite abelian group G and a prime p , we write G_p for the Sylow p -subgroup of G . For a set P of primes, we write $G_P := \prod_{p \in P} G_p$.

Throughout this paper $C_i, K_i, c_i, \delta, \eta, \varepsilon, \lambda$, etc will denote positive constants. When it does not create confusion, the same letter may denote different constants in different parts of the proof. The value of the constants may depend on other constants we have chosen, but will never depend on the dimension n , which is regarded as an asymptotic parameter going to infinity. We consider many functions of n and other parameters, e.g. including $u, \{\xi_i\}_i, \alpha, \varepsilon, T, d, p, q$ (where for $\{\xi_i\}_i$ we mean their probability distributions). We say “ $f(n, \dots) \in O_S(g(n, \dots))$,” where S is a subset of the parameters, to mean for any values v_1, \dots, v_m of the parameters in S , there is exists a constant $K > 0$ depending on v_1, \dots, v_m , such that for all n sufficiently large given v_1, \dots, v_m , and *all* allowed values of the parameters not in S , that $|f(n, v_1, \dots, v_m, \dots)| \leq K g(n, v_1, \dots, v_m, \dots)$.

2. ORGANIZATION OF THE PROOF OF THEOREMS 1.1, 1.2, 1.4 AND 1.6

We will be mainly focusing on the i.i.d. case to prove Theorems 1.1, 1.2, and 1.4. The results for the Laplacian case will be shown in a similar fashion. We prove Theorems 1.1, 1.2, and 1.4 for $M_{n \times (n+u)}$ by checking if the Sylow- p subgroup of $\text{Cok}(M_{n \times (n+u)})$ is equal to B_p for each prime p (or is cyclic for each prime p). The argument will then break up into considering primes in three size ranges with totally different treatments.

For small primes, we prove the following generalization of [34, Corollary 3.4] to sparser matrices, which requires a refinement of the method of [34].

Proposition 2.1 (Small Primes). *Let $M_{n \times (n+u)}$ be as in Theorem 1.1. Let B be a finite abelian group. Let P be a finite set of primes including all those dividing $|B|$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\text{Cok}(M_{n \times (n+u)})_P \simeq B\right) = \frac{1}{|B|^u |\text{Aut}(B)|} \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-u}).$$

Proposition 2.1 is a special case of Theorem 4.1, which allows the matrices to be even sparser and have non-identical entries. This carries the main term of our estimates.

For medium primes, we combine a result of [21] with a comparison theorem on the evolving of the matrix ranks to obtain the following.

Proposition 2.2 (Medium Primes). *There are constants $c_0, \eta > 0$ such that the following holds. Let $M_{n \times (n+u)}$ be as in Theorem 1.1. Let p be a prime. Then,*

$$\mathbf{P}\left(M_{n \times (n+u)} \text{ mod } p \text{ is not full rank}\right) \leq 2p^{-\min(u+1, \eta n)} + O(e^{-c_0 \alpha_n n}) \quad (7)$$

and

$$\mathbf{P}\left(M_{n \times (n+u)} \text{ mod } p \text{ has rank} \leq n-2\right) \leq 2p^{-\min(2u+4, \eta n)} + O(e^{-c_0 \alpha_n n}). \quad (8)$$

Proposition 2.2 will follow from Theorem 5.1 where we allow the matrices to be sparser. (The big O allows us to require that n is large enough that $\alpha_n \geq n^{-1+\varepsilon} \geq C_0 \log n/n$.)

Even such a small error bound cannot be summed over all primes, and so for large primes we present a new approach that considers all large primes together.

Proposition 2.3 (Large Primes). *Let $d > 0$ and let $M_{n \times (n+u)}$ and ε be as in Theorem 1.1. Then,*

$$\mathbf{P}\left(\forall \text{ primes } p \geq e^{d\alpha_n n} : M_{n \times (n+1)} \bmod p \text{ has rank at least } n\right) \geq 1 - O_{d,T,\varepsilon}(n^{-\varepsilon}), \quad (9)$$

as well as

$$\mathbf{P}\left(\forall \text{ primes } p \geq e^{d\alpha_n n} : M_{n \times n} \bmod p \text{ has rank at least } n-1\right) \geq 1 - O_{d,T,\varepsilon}(n^{-\varepsilon}). \quad (10)$$

Proposition 2.3 is proven in Sections 6 and 7, and is the source of the lower bound on α_n in our theorems.

The heart of the paper is proving the three propositions above, as the main theorems follow simply from these, as we now show.

Proof of Theorems 1.1 and 1.4. We first prove Theorem 1.4, the case when $u \rightarrow \infty$, where we need to consider only medium and large primes. By Equation (7) of Proposition 2.2, there are $c_0, \eta > 0$ such that

$$\begin{aligned} \mathbf{P}\left(M_{n \times (n+u)} \bmod p \text{ not full rank for some } 2 \leq p \leq e^{c_0 \alpha_n n/2}\right) &\leq \sum_{2 \leq p \leq e^{c_0 \alpha_n n/2}} (2p^{-\min(u+1, \eta n)} + O(e^{-c_0 \alpha_n n})) \\ &= O\left(\frac{1}{2^{\min(u, \eta n)}} + e^{-c_0 \alpha_n n/2}\right). \end{aligned}$$

Combined with Equation (9) of Proposition 2.3 (applied to $d = c_0/2$) we obtain

$$\mathbf{P}\left(M_{n \times (n+u)} \bmod p \text{ is full rank for all } p \geq 2\right) \geq 1 - O_{T,\varepsilon}\left(\frac{1}{2^{\min(u, \eta n)}} + e^{-c_0 \alpha_n n/2} + n^{-\varepsilon}\right),$$

completing the proof of Equation (3).

We next turn to Equation (2). Let k_0 be fixed and $u \geq 1$ be fixed. By applying Equation (7) of Proposition 2.2 for n large enough given η and u we have

$$\begin{aligned} \mathbf{P}\left(M_{n \times (n+u)} \bmod p \text{ is not full rank for some } k_0 \leq p \leq e^{c_0 \alpha_n n/2}\right) &\leq \sum_{p=k_0}^{e^{c_0 \alpha_n n/2}} (2p^{-(u+1)} + O(e^{-c_0 \alpha_n n})) \\ &= O\left(\frac{1}{k_0} + e^{-c_0 \alpha_n n/2}\right). \end{aligned}$$

Combined with Equation (9) of Proposition 2.3 we obtain

$$\mathbf{P}\left(M_{n \times (n+u)} \bmod p \text{ is full rank for all } p \geq k_0\right) \geq 1 - O_{T,\varepsilon}\left(\frac{1}{k_0} + e^{-c_0 \alpha_n n/2} + n^{-\varepsilon}\right).$$

Now let k_0 be at least as large as the largest prime divisor of $|B|$, and let P be the collection of primes up to k_0 . By Proposition 2.1,

$$\mathbf{P}\left(\mathbf{Cok}(M_{n \times (n+u)})_P \simeq B\right) = \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \leq k_0} \prod_{k=1}^{\infty} (1 - p^{-k-u}) + o_{\{\xi_i\}_i, B, u}(1). \quad (11)$$

Putting the two bounds together,

$$\mathbf{P}\left(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B\right) \geq \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \leq k_0} \prod_{k=1}^{\infty} (1 - p^{-k-u}) - O_{T,\varepsilon}\left(\frac{1}{k_0} + e^{-c_0 \alpha_n n/2} + n^{-\varepsilon}\right) + o_{\{\xi_i\}_i, B, u}(1).$$

Taking the limit as $n \rightarrow \infty$, we obtain

$$\liminf_{n \rightarrow \infty} \mathbf{P}\left(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B\right) \geq \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \leq k_0} \prod_{k=1}^{\infty} (1 - p^{-k-u}) - O_{T,\varepsilon}\left(\frac{1}{k_0}\right).$$

As this is true for any fixed k_0 , we can take $k_0 \rightarrow \infty$ to obtain,

$$\liminf_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B) \geq \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_p \prod_{k=1}^{\infty} (1 - p^{-k-u}).$$

Since $\mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B) \leq \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)})_P \simeq B_P)$, Equation (11) gives

$$\limsup_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times (n+u)}) \simeq B) \leq \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \leq k_0} \prod_{k=1}^{\infty} (1 - p^{-k-u}),$$

completing the proof of (2). \square

When $u \geq 1$, the probabilities in Theorem 1.1 sum, over finite abelian groups B , to 1 [34, Lemma 3.2]. From this observation and Fatou's Lemma, we can conclude that the asymptotic probability of any property is just the sum over B with that property of the asymptotic probability of B . Thus to deduce Corollary 1.5 and the $u \geq 1$ cases of Theorem 1.2, it remains only to sum $|B|^{-u} |\mathbf{Aut}(B)|^{-1}$ over B that are of square-free order or cyclic, which is straightforward.

The proof of Theorem 1.2 is identical to the proof of Theorem 1.1, using Equations (8) and (10) in place of Equations (7) and (9), and the fact that $\mathbf{Cok}(M_{n \times n})$ is cyclic if and only if for every prime p , the matrix $M_{n \times n} \bmod p$ has rank at least $n-1$. In fact, the proof gives the following.

Theorem 2.4. *Let $M_{n \times n}$ be as in Theorem 1.1. Let B be a finite abelian group and let k_0 be larger than any prime divisor of $|B|$, and define $C_B = \{B \times C \mid C \text{ cyclic, } p \nmid |C| \text{ for } 1 < p < k_0\}$, the set of groups differing from B by a cyclic group with order only divisible by primes at least k_0 . Then, we have*

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M_{n \times n}) \in C_B) = \frac{1}{|\mathbf{Aut}(B)|} \prod_{\substack{p < k_0 \\ p \text{ prime}}} (1 - p^{-1}) \prod_{\substack{p \geq k_0 \\ p \text{ prime}}} (1 + (p^2 - p)^{-1}) \prod_{k=2}^{\infty} \zeta(k)^{-1}.$$

Now we turn to the Laplacian, where we will follow an almost identical outline (corresponding to the case $u = 1$ of our i.i.d. model $M_{n \times (n+u)}$). Indeed we will prove Theorem 1.6 by checking if the Sylow- p subgroup of S_M is equal to B_p for each prime p (or is cyclic for each prime p) in three size ranges. We prove the following proposition in Section 8.

Proposition 2.5. *Let $M_{n \times n}$ and ε be as in Theorem 1.6. There are constants $c_0, d > 0$ such that the following holds.*

- (Small Primes) *Let B be a finite abelian group. Let P be a finite set of primes including all those dividing $|B|$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}((S_{M_{n \times n}})_P \simeq B) = \frac{1}{|B| |\mathbf{Aut}(B)|} \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-1}). \quad (12)$$

- (Medium primes) *Let p be any prime. Then,*

$$\mathbf{P}(L_{M_{n \times n}} \bmod p \text{ has rank } \leq n-2) \leq 2p^{-2} + O(e^{-c_0 \alpha_n n}) \quad (13)$$

and

$$\mathbf{P}(L_{M_{n \times n}} \bmod p \text{ has rank } \leq n-3) \leq 2p^{-6} + O(e^{-c_0 \alpha_n n}). \quad (14)$$

- (Large primes) *We also have*

$$\mathbf{P}\left(\forall \text{ primes } p \geq e^{d\alpha_n n} : L_{M_{n \times n}} \bmod p \text{ has full rank in } \mathbf{Z}_0^n/p\right) \geq 1 - O_{d, T, \varepsilon}(n^{-\varepsilon}), \quad (15)$$

as well as

$$\mathbf{P}\left(\forall \text{ primes } p \geq e^{d\alpha_n n} : L_{M_{n \times (n-1)}} \bmod p \text{ has rank at least } n-2\right) \geq 1 - O_{d, T, \varepsilon}(n^{-\varepsilon}). \quad (16)$$

The deduction of Theorem 1.6 from the above results is similar to the deduction of Theorem 1.1 and Theorem 1.2 from Propositions 2.1, 2.2 and 2.3, and hence is omitted.

3. ODLYZKO'S LEMMA

In this section we give an elementary but extremely useful tool which is a variant of Odlyzko's lemma [15] (also [19, Lemma 2.2]). This result will be used in the arguments for small, medium, and large primes. We will focus on the i.i.d case and refer the reader to Lemma 8.3 for a similar result regarding the Laplacian.

Lemma 3.1. *Let \mathbf{F} be a field. For a deterministic subspace V of \mathbf{F}^n of dimension d and a random vector $X \in \mathbf{F}^n$ with i.i.d. entries taking any value with probability at most $1 - \alpha_n$,*

$$\mathbf{P}(X \in V) \leq (1 - \alpha_n)^{n-d}.$$

We give a short proof of this well-known result for completeness.

Proof. Assume that $V = \text{Span}(H_1, \dots, H_d)$, where $H_i = (h_{i1}, \dots, h_{in})$, and without loss of generality we assume the matrix $(h_{ij})_{1 \leq i, j \leq d}$ has rank d . Consider the event $X = (x_1, \dots, x_d, x_{d+1}, \dots, x_n) \in V$. Because $(h_{ij})_{1 \leq i, j \leq d}$ has rank d , there exist unique coefficients $c_1, \dots, c_d \in \mathbf{F}$ such that

$$(x_1, \dots, x_d) = \sum_i c_i (h_{i1}, \dots, h_{id}).$$

Hence conditioning on (x_1, \dots, x_d) , if $X = (x_1, \dots, x_d, x_{d+1}, \dots, x_n) \in V$ then for all $d+1 \leq j \leq n$

$$x_j = \sum_i c_i h_{ij}.$$

However the probability of each of these events is at most $1 - \alpha_n$, and so conditioning on (x_1, \dots, x_d) , the event $X \in V$ holds with probability at most $(1 - \alpha_n)^{n-d}$. \square

Corollary 3.2. *Let X_1, \dots, X_{n-k} be random vectors with i.i.d. entries taking any value with probability at most $1 - \alpha_n$. Then the probability that X_1, \dots, X_{n-k} are linearly independent in \mathbf{F}^n is at least $1 - \alpha_n^{-1}(1 - \alpha_n)^k$.*

Proof. Let $0 \leq i \leq n-k-1$ be minimal such that $X_{i+1} \in \text{span}(X_1, \dots, X_i)$. By Lemma 3.1, this event is bounded by $(1 - \alpha_n)^{n-i}$. Summing over $0 \leq i \leq n-k-1$, the probability under consideration is bounded by $\sum_{i=0}^{n-k-1} (1 - \alpha_n)^{n-i} < \alpha_n^{-1}(1 - \alpha_n)^k$. \square

In all three arguments, Lemma 3.1 (Odlyzko's lemma) will only suffice for the easy part of the argument, and a stronger, Littlewood-Offord style bound (Lemma 4.7, Theorem 5.2, Theorem 6.3, Theorem 7.3) will be required for the harder part of the argument. The details of the Littlewood-Offord style bound required are different in each argument, and thus are given in the corresponding sections. Note that Odlyzko's lemma is too weak to be used alone for our purposes, because it can produce a bound $1 - \alpha_n$, where we require bounds that go to 0 as $n \rightarrow \infty$. In this paper, α_n is possibly small. If, however, the matrix entries take values modulo large primes with probability at most $1 - \alpha_n$, and $1 - \alpha_n \rightarrow 0$ as $n \rightarrow \infty$, then we expect our arguments can all be considerably simplified and only Odlyzko's lemma would be necessary (and no Littlewood-Offord style bounds required). For example, such a simplification works to handle the case of entries chosen uniformly in an interval centered at 0 with size growing at any rate with n .

4. SMALL PRIMES

In this section, we prove the following theorem, which generalizes [34, Corollary 3.4] to smaller α_n and implies our Proposition 2.1. The method requires refinement from that of [34], and we discuss the differences below.

Theorem 4.1. *Let u be a non-negative integer and α_n a function of integers n such that for any constant $\Delta > 0$, for n sufficiently large we have $\alpha_n \geq \Delta \log n/n$. For every positive integer n , let $M(n)$ be a random matrix valued in $M_{n \times (n+u)}(\mathbf{Z})$ with independent α_n -balanced entries. Let B be a finite abelian group. Let P be a finite set of primes including all those dividing $|B|$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M(n))_P \simeq B) = \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-u}).$$

Note that the entries of the matrix do not have to be identical.

Throughout the section we write $\text{Hom}(A, B)$ and $\text{Sur}(A, B)$ for the set of homomorphisms and surjective homomorphisms, respectively, from A to B . We will always use a to denote a positive integer and $R = \mathbf{Z}/a\mathbf{Z}$. We then study finite abelian groups G whose exponent divides a , i.e. $aG = 0$. We write G^* for $\text{Hom}(G, R)$.

4.2. Set-up. We will study integral matrices by reducing them mod each positive integer. We let a be a positive integer. Let M be the random $n \times (n+u)$ matrix with entries in R that is the reduction of $M(n)$ from Theorem 4.1 modulo a . We let $X_1, \dots, X_{n+u} \in R^n$ be the columns of M , and x_{ij} the entries of M (so that the entries of X_j are x_{ij}). For a positive integer n , we let $V = R^n$ with standard basis v_i and $W = R^{n+u}$ with standard basis w_j (these will always implicitly depend on the integers we call a and n). Note for $\sigma \subset [n]$, V has distinguished submodules $V_{\setminus \sigma}$ generated by the v_i with $i \notin \sigma$. (So $V_{\setminus \sigma}$ comes from not using the σ coordinates.) We view $M \in \text{Hom}(W, V)$ and its columns X_j as elements of V so that $X_j = Mw_j = \sum_i x_{ij}v_i$. Let G be a finite abelian group with exponent dividing a . We have $\text{Cok}(M) = V/MW$.

We know from [34] that to understand the distribution of $\text{Cok}(M)$, it suffices to determine certain moments. To investigate the moments $\mathbf{E}(\#\text{Sur}(\text{Cok}(M), G))$ (see [5, Section 3.3] for more on why these are “moments”), we recognize that each such surjection lifts to a surjection $V \rightarrow G$ and so we have

$$\mathbf{E}(\#\text{Sur}(\text{Cok}(M), G)) = \sum_{F \in \text{Sur}(V, G)} \mathbf{P}(F(MW) = 0). \quad (17)$$

By the independence of columns, we have

$$\mathbf{P}(F(MW) = 0) = \prod_{j=1}^m \mathbf{P}(F(X_j) = 0).$$

So we aim to estimate these probabilities $\mathbf{P}(F(X_j) = 0)$, which will give us our desired moments.

4.3. Finding the moments. We will first estimate $\mathbf{P}(F(X_j) = 0)$ for the vast majority of F , which satisfy the following helpful property.

Definition 4.4. We say that $F \in \text{Hom}(V, G)$ is a *code* of distance w , if for every $\sigma \subset [n]$ with $|\sigma| < w$, we have $FV_{\setminus \sigma} = G$. In other words, F is not only surjective, but would still be surjective if we throw out (any) fewer than w of the standard basis vectors from V . (If a is prime so that R is a field, then this is equivalent to whether the transpose map $F : G^* \rightarrow V^*$ is injective and has image $\text{im}(F) \subset V^*$ a linear code of distance w , in the usual sense.)

First we recall a lemma from [34] that lets us see how a code F acts on a single column from our matrix. The following statement is slightly stronger than [34, Lemma 2.1], but one can see this statement follows directly from the proof of [34, Lemma 2.1].

Lemma 4.5. *Let a, n be positive integers, G a finite abelian group of exponent dividing a , and X the reduction mod a of a random vector in \mathbf{Z}^n with independent, α -balanced entries. Let $F \in \text{Hom}(V, G)$ be a code of distance w and $A \in G$. We have*

$$|\mathbf{P}(FX = A) - |G|^{-1}| \leq \frac{|G| - 1}{|G|} \exp(-\alpha w/a^2).$$

We then will put these estimates for columns together using this simple inequality.

Lemma 4.6 ([34, Lemma 2.3]). *If we have an integer $m \geq 2$ and real numbers $x \geq 0$ and y such that $|y|/x \leq 2^{1/(m-1)} - 1$ and $x + y \geq 0$, then*

$$x^m - 2mx^{m-1}|y| \leq (x + y)^m \leq x^m + 2mx^{m-1}|y|.$$

The below is a refinement of [34, Lemma 2.4] that allows sparse matrices.

Lemma 4.7 (Bound for codes). *Let $a \geq 1$ and $u \geq 0$ be integers, G be a finite abelian group of exponent dividing a , the sequence $\{\alpha_n\}_n$ be as in Theorem 4.1, and $\delta > 0$. Then there are $c_1, K_1 > 0$ such that the*

following holds. Let $\bar{M}(n)$ be the reduction modulo a of random matrices $M(n)$ as in Theorem 4.1. For every positive integer n , and $F \in \text{Hom}(V, G)$ a code of distance δn , and $A \in \text{Hom}(W, G)$, we have

$$|\mathbf{P}(F\bar{M}(n) = A) - |G|^{-n-u}| \leq \frac{K_1 n^{-c_1}}{|G|^{n+u}}.$$

Proof. Choose $\Delta > a^2\delta^{-1}$ and n large enough (depending on Δ and $\{\alpha_i\}_i$) so that $\alpha_n \geq \Delta \log n/n$. Then for n large enough given $\delta, \Delta, u, a, |G|$, we have

$$\exp(-(\Delta \log n/n)\delta n/a^2)|G| = \exp(-\Delta \delta \log n/a^2)|G| \leq \frac{\log 2}{n+u-1} \leq 2^{1/(n+u-1)} - 1.$$

So for such n we can combine Lemma 4.5 and Lemma 4.6 (with $x = |G|^{-1}$ and $y = \pm \exp(-(\Delta \log n/n)\delta n/a^2)$) to obtain

$$|\mathbf{P}(FM = A) - |G|^{-n-u}| \leq 2(n+u) \exp(-\Delta \delta \log n/a^2)|G|^{-n-u+1}.$$

We take $c_1 < \Delta \delta/a^2 - 1$ and then for n sufficiently large given $u, \Delta, \delta, a, c_1, |G|, \{\alpha_i\}_i$, we have

$$|\mathbf{P}(FM = A) - |G|^{-n-u}| \leq \frac{2(n+u)}{n^{1+(\Delta \delta/a^2-1)}}|G|^{-n-u+1} \leq n^{-c_1}|G|^{-n-u}.$$

We choose K_1 large enough so that $\frac{K_1 n^{-c_1}}{|G|^{n+u}} \geq 2$ for n that are not as large as needed above, and the lemma follows. \square

So far, we have a good estimate for $\mathbf{P}(FM = 0)$ when F is a code. Unfortunately, it is not sufficient to divide F into codes and non-codes. We need a more delicate division of F based on the subgroups of G . In [34], a notion of *depth* was used to divide the F into classes. Here we require a slightly finer notion (that we call *robustness*) to deal with the sparser matrices. Both notions can be approximately understood as separating the F based on what largest size subgroup they are a code for. For an integer D with prime factorization $\prod_i p_i^{e_i}$, let $\ell(D) = \sum_i e_i$.

Definition 4.8. Given $\delta > 0$, we say that $F \in \text{Hom}(V, G)$ is *robust* (or, more precisely, δ -robust) for a subgroup H of G if H is minimal such that

$$\#\{i \in [n] | Fv_i \notin H\} \leq \ell([G : H])\delta n.$$

Note that $H = G$ satisfies the above inequality, so every $F \in \text{Hom}(V, G)$ is robust for some subgroup H of G . An F might be robust for more than one subgroup.

Lemma 4.9. Let $\delta > 0$, and a, n be positive integers, and G be a finite abelian group of exponent dividing a . Let $F \in \text{Hom}(V, G)$ be robust for H . Let $\pi := \{i \in [n] | Fv_i \notin H\}$. Then F restricted to $V_{\setminus \pi}$ is a code of distance δn in $\text{Hom}(V_{\setminus \pi}, H)$.

Proof. Suppose not. Then there exists a $\sigma \subset [n] \setminus \pi$ such that $|\sigma| < \delta n$ and $FV_{\setminus (\pi \cup \sigma)}$ lies in some proper subgroup H' of H . In particular, the set of i such that $Fv_i \notin H'$ is contained in $\pi \cup \sigma$. Since

$$|\pi \cup \sigma| \leq \ell([G : H])\delta n + \delta n \leq \ell([G : H'])\delta n,$$

we then have a contradiction on the minimality of H . \square

We then bound the number of F that are robust for a certain group H , and with certain given behavior outside of H . The separation of F into classes based on their behavior outside of H did not appear in [34], but is necessary here to deal with sparser matrices.

Lemma 4.10 (Count of robust F for a subgroup H). Let $\delta > 0$, and $a, n \geq 1$ be integers, and G be finite abelian group of exponent dividing a . Let H be a subgroup of G of index $D > 1$ and let $H = G_{\ell(D)} \subset \dots \subset G_2 \subset G_1 \subset G_0 = G$ be a maximal chain of proper subgroups. Let $p_j = |G_{j-1}/G_j|$. The number of $F \in \text{Hom}(V, G)$ such that F is robust for H and for $1 \leq j \leq \ell(D)$, there are w_j elements i of $[n]$ such that $Fv_i \in G_{j-1} \setminus G_j$, is at most

$$|H|^{n - \sum_j w_j} \prod_{j=1}^{\ell(D)} \binom{n}{w_j} |G_{j-1}|^{w_j}.$$

Note that by the definition of robustness we have that $w_j \leq \ell([G : H])\delta n$, or else there are no such F .

Proof. There are at most $\binom{n}{w_j}$ ways to choose the i such that $Fv_i \in G_{j-1} \setminus G_j$ and then at most $|G_j|^{w_j}$ ways to choose the Fv_i . Then there are $|H|$ choices for each remaining Fv_i . \square

Now for F robust for a subgroup H , we will get a bound on $\mathbf{P}(FM = 0)$, where the larger the H , the better the bound. This is a more delicate bound than [34, Lemma 2.7] that it is replacing, and in particular takes into account the behavior of F outside of H .

Lemma 4.11 (Probability bound for columns given robustness). *Let $\delta > 0$, and $a, n \geq 1$ be integers, and G be finite abelian group of exponent dividing a . Let $F \in \text{Hom}(V, G)$ be robust for a proper subgroup H of G and let $D := [G : H]$. Let $H = G_{\ell(D)} \subset \dots \subset G_2 \subset G_1 \subset G_0 = G$ be a maximal chain of proper subgroups. Let $p_j = |G_{j-1}/G_j|$. For $1 \leq j \leq \ell(D)$, let w_j be the number of $i \in [n]$ such that $Fv_i \in G_{j-1} \setminus G_j$. Let $X \in R^n$ be a random vector with independent entries that are the reduction mod a of α -balanced random integers. Then for all n ,*

$$\mathbf{P}(FX = 0) \leq \left(D|G|^{-1} + \exp(-\alpha\delta n/a^2) \right) \prod_{j=1}^{\ell(D)} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha w_j/p_j^2) \right).$$

Proof. Assume that $X = (x_1, \dots, x_n)$. Let σ_j be the collection of indices $i \in [n]$ such that $Fv_i \in G_{j-1} \setminus G_j$. Let $\sigma = \bigcup_{j=1}^{\ell(D)} \sigma_j$. Then,

$$\begin{aligned} \mathbf{P}(FX = 0) &= \mathbf{P} \left(\sum_{i \in \sigma_1} (Fv_i) x_i \in G_1 \right) \mathbf{P} \left(\sum_{i \in \sigma_1 \cup \sigma_2} (Fv_i) x_i \in G_2 \mid \sum_{i \in \sigma_1} (Fv_i) x_i \in G_1 \right) \times \dots \\ &\quad \times \mathbf{P} \left(\sum_{i \in \sigma_1 \cup \dots \cup \sigma_{\ell(D)}} (Fv_i) x_i \in H \mid \sum_{i \in \sigma_1 \cup \dots \cup \sigma_{\ell(D)-1}} (Fv_i) x_i \in G_{\ell(D)-1} \right) \\ &\quad \times \mathbf{P} \left(\sum_{i \notin \sigma} (Fv_i) x_i = - \sum_{i \in \sigma} (Fv_i) x_i \mid \sum_{i \in \sigma} (Fv_i) x_i \in H \right). \end{aligned}$$

For $1 \leq j \leq \ell(D)$, we will bound the j th factor above by conditioning on the x_i with $i \in \sigma_1 \cup \dots \cup \sigma_{j-1}$ and then looking at images in G_{j-1}/G_j . Note for $i \in \sigma_j$, we have that the reduction of Fv_i is non-zero in G_{j-1}/G_j . So F restricted to the σ_j coordinates in the reduction to G_{j-1}/G_j is a code of length w_j . We then apply Lemma 4.5 to this case to obtain

$$\mathbf{P} \left(\sum_{i \in \sigma_1 \cup \dots \cup \sigma_j} (Fv_i) x_i \in G_j \mid \sum_{i \in \sigma_1 \cup \dots \cup \sigma_{j-1}} (Fv_i) x_i \in G_{j-1} \right) \leq p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha w_j/p_j^2).$$

Note that σ is the set of i such that $Fv_i \notin H$. By the definition of robust, $|\sigma| < \ell(D)\delta n$. By Lemma 4.9, the restriction of F to $V_{\setminus \sigma}$ is a code of distance δn in $\text{Hom}(V_{\setminus \sigma}, H)$. So conditioning on the X_i with $i \in \sigma$, we can estimate the conditional probability above using Lemma 4.5:

$$\mathbf{P} \left(\sum_{i \notin \sigma} (Fv_i) x_i = - \sum_{i \in \sigma} (Fv_i) x_i \mid \sum_{i \in \sigma} (Fv_i) x_i \in H \right) \leq |H|^{-1} + \exp(-\alpha\delta n/a^2).$$

The lemma follows. \square

Now we can combine the estimates we have for $\mathbf{P}(FM = 0)$ for various types of F with the bounds we have on the number of F of each type to obtain our main result on the moments of cokernels of random matrices.

Theorem 4.12. *Let $u \geq 0$ be an integer, G be a finite abelian group, and the sequence $\{\alpha_n\}_n$ be as in Theorem 4.1. Then there are c_2, K_2 such that the following holds. For every positive integer n and random matrix $M(n)$ as in Theorem 4.1, we have*

$$|\mathbf{E}(\#\text{Sur}(\text{Cok}(M(n)), G)) - |G|^{-u}| \leq K_2 n^{-c_2}.$$

Proof. Let a be the exponent of G . By Equation (17), we need to estimate $\sum_{F \in \text{Sur}(V, G)} \mathbf{P}(FM(n) = 0)$. Fix a proper subgroup H of G . We will apply Lemma 4.11 and use the notation from that lemma, along with

Lemma 4.10. We then have

$$\begin{aligned}
& \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ is robust for } H}} \mathbf{P}(FM(n) = 0) \\
& \leq \sum_{\substack{0 \leq w_1, \dots, w_{\ell(D)} \leq \ell(D)\delta n \\ w_1 \neq 0}} |H|^{n-\sum_j w_j} \prod_{j=1}^{\ell(D)} \binom{n}{w_j} |G_{j-1}|^{w_j} \prod_{j=1}^{\ell(D)} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha_n w_j / p_j^2) \right)^{n+u} \\
& \quad \times \left(D|G|^{-1} + \exp(-\alpha_n \delta n / a^2) \right)^{n+u} \\
& = |H|^n \left(D|G|^{-1} + \exp(-\alpha_n \delta n / a^2) \right)^{n+u} \prod_{j=1}^{\ell(D)} \sum_{\substack{w_j=0 \\ w_1 \neq 0}} |H|^{-w_j} \binom{n}{w_j} |G_{j-1}|^{w_j} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha_n w_j / p_j^2) \right)^{n+u}.
\end{aligned}$$

We have $w_1 \neq 1$ since F is a surjection. Now we apply Lemma A.3 from the Appendix to bound the sums. The D_1, d_1 from Lemma A.3 will be $|G_{j-1}|/|H|$ and p_j^{-1} respectively. We choose the Δ' of Lemma A.3 so that $\Delta' > 2/(1-p_j^{-1})$ for all j . For n sufficiently large (in terms of $\{\alpha_i\}_i, \Delta', G$), we have $\alpha_n \geq p_j^2 \Delta' \log n / n$ for all j . Lemma A.3 then gives us that, for δ sufficiently small (given G), and n sufficiently large (given $G, \Delta', \delta, \{\alpha_i\}_i$), we have

$$\sum_{w_j=1}^{\ell(D)\delta n} \binom{n}{w_j} \left(\frac{|G_{j-1}|}{|H|} \right)^{w_j} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha_n w_j / p_j^2) \right)^{n+u} \leq 3n^{((1-p_j^{-1})\Delta'/2-1)}.$$

Let $\Delta > a^2 \delta^{-1}$ and $\Delta > 2p^3/(p-1)$ for every prime $p \mid a$. For n also sufficiently large (given Δ and $\{\alpha_i\}_i$) that $\alpha_n \geq \Delta \log n / n$, and we have

$$|H|^n \left(D|G|^{-1} + \exp(-\alpha_n \delta n / a^2) \right)^{n+u} \leq |H|^{-u} \left(1 + |H| \exp(-\Delta \delta \log n / a^2) \right)^{n+u}.$$

For $n+u \geq 2$, and n sufficiently large (given δ, Δ, u, G) such that

$$|H| \exp(-\Delta \delta \log n / a^2) = |H| n^{-\Delta \delta / a^2} \leq \frac{\log 2}{n+u-1} \leq 2^{1/(n+u-1)} - 1.$$

By Lemma 4.6,

$$\left(1 + |H| \exp(-\Delta \delta \log n / a^2) \right)^{n+u} \leq 1 + 2(n+u) |H| \exp(-\Delta \delta \log n / a^2).$$

Putting it altogether we have

$$\begin{aligned}
& \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ is robust for } H}} \mathbf{P}(FM(n) = 0) \\
& \leq |H|^n \left(D|G|^{-1} + \exp(-\alpha_n \delta n / a^2) \right)^{n+u} \prod_{j=1}^{\ell(D)} \sum_{\substack{w_j=0 \\ w_1 \neq 0}} |H|^{-w_j} \binom{n}{w_j} |G_{j-1}|^{w_j} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp(-\alpha_n w_j / p_j^2) \right)^{n+u} \\
& \leq |H|^{-u} \left(1 + 2(n+u) |H| \exp(-\Delta \delta \log n / a^2) \right) 3n^{((1-p_1^{-1})\Delta'/2-1)} \prod_{j=2}^{\ell(D)} \left(1 + 3n^{((1-p_j^{-1})\Delta'/2-1)} \right).
\end{aligned}$$

We sum this over proper subgroups H of G to bound, for δ sufficiently small (given G), and $\Delta' > 2/(1-p_j^{-1})$ for all j , and $\Delta > a^2 \delta^{-1}$, and $\Delta > 2p^3/(p-1)$ for every prime $p \mid a$, for n sufficiently large given $G, \delta, \Delta, \Delta', u, \{\alpha_i\}_i$,

$$\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} \mathbf{P}(FM(n) = 0) \leq K_3 n^{-c_3}.$$

where K_3 is a constant depending on $G, \delta, \Delta, \Delta', u$ and $c_3 > 0$ (depending on a, Δ').

Also, from the proof of [34, Theorem 2.9], we can choose δ small enough (given G) so that we have for all n

$$\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} |G|^{-n-u} \leq K_4 e^{-(\log 1.5)n}$$

for some K_4 depending on u, G, δ . We also have (e.g. see the proof of [34, Theorem 2.9]) for all n ,

$$\sum_{F \in \text{Hom}(V, G) \setminus \text{Sur}(V, G)} |G|^{-n-u} \leq K_5 e^{-\log(2)n}$$

for some K_5 depending on G . Using Lemma 4.7 we have that for all n ,

$$\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ code of distance } \delta n}} |\mathbf{P}(FM(n) = 0) - |G|^{-n-u}| \leq K_1 n^{-c_1}.$$

We now make a choice of δ that is sufficiently small for the two requirements above (given G), and we choose Δ and Δ' as required above, so that for all n sufficiently large (given $G, \delta, \Delta, \Delta', u, \{\alpha_i\}_i$),

$$\begin{aligned} \left| \sum_{F \in \text{Sur}(V, G)} \mathbf{P}(FM(n) = 0) - |G|^{-u} \right| &= \left| \sum_{F \in \text{Sur}(V, G)} \mathbf{P}(FM(n) = 0) - \sum_{F \in \text{Hom}(V, G)} |G|^{-n-u} \right| \\ &\leq \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ code of distance } \delta n}} \left| \mathbf{P}(FM(n) = 0) - |G|^{-n-u} \right| + \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} \mathbf{P}(FM(n) = 0) + \sum_{\substack{F \in \text{Hom}(V, G) \\ F \text{ not code of distance } \delta n}} |G|^{-n-u} \\ &\leq K_1 n^{-c_1} + K_3 n^{-c_3} + K_4 e^{-(\log 1.5)n} + K_5 e^{-\log(2)n}. \end{aligned}$$

We choose $c_2 \leq \min(c_1, c_3, \log(1.5))$ (which depends on $G, u, \{\alpha_i\}_i$). We choose K_2 so that $K_2 \geq K_1 + K_3 + K_4 + K_5$, and also $K_2 \geq |G|^n n^{c_2}$ for any n not sufficiently large for the requirements above (so K_2 depends on $G, u, \{\alpha_i\}_i$). The theorem follows. \square

We now conclude the proof of Theorem 4.1. For each fixed $u \geq 1$, we construct a random abelian group according to Cohen and Lenstra's distribution mentioned in the introduction. Independently for each p , we have a random finite abelian p -group Y_p such that for each p -group B

$$\mathbf{P}(Y_p = B) = \frac{\prod_{k=1}^{\infty} (1 - p^{-k-u})}{|B|^u |\mathbf{Aut}(B)|}.$$

Let P be a set of primes dividing a given number a , we then define a random group Y by taking the group product $\prod_{p \in P} Y_p$.

Lemma 4.13 ([34, Lemma 3.2]). *For every finite abelian group G with exponent dividing a we have*

$$\mathbf{E}(\# \text{Sur}(Y, G)) = |G|^{-u}.$$

From Theorem 4.12 we have seen that Y and $\mathbf{Cok}(M(n))$ have asymptotic matching “moments” with respect to all groups G of exponent dividing a . To pass this information back to distribution, we then use the following result on the moment problem for finite abelian groups.

Theorem 4.14. *Let X_n and Y_n be sequences of random finitely generated abelian groups. Let a be a positive integer and \mathcal{A} be the set of isomorphism classes of abelian groups with exponent dividing a . Suppose that for every $G \in \mathcal{A}$ we have a number $M_G \leq |\wedge^2 G|$ such that $\lim_{n \rightarrow \infty} \mathbf{E}(\# \text{Sur}(X_n, G)) = \lim_{n \rightarrow \infty} \mathbf{E}(\# \text{Sur}(Y_n, G)) = M_G$. Then we have that for every $H \in \mathcal{A}$*

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n \otimes (\mathbf{Z}/a\mathbf{Z}) \simeq H) = \lim_{n \rightarrow \infty} \mathbf{P}(Y_n \otimes (\mathbf{Z}/a\mathbf{Z}) \simeq H).$$

The proof of Theorem 4.14 can be found in the proof of [34, Theorem 3.1], which is stated for $M_G = |\wedge^2 G|$ but only ever uses the inequality $M_G \leq |\wedge^2 G|$. (Here $\wedge^2 G$ is a particular quotient of the tensor product $G \otimes G$, and all we need here is $|\wedge^2 G| \geq 1$.) To prove Theorem 4.1, assume that the exponent of the

group B under consideration has prime factorization $\prod_{p \in P} p^{e_p}$. Theorem 4.14, applied to the sequence $X_n = \mathbf{Cok}(M(n))$ and $Y_n = Y$ with $a = \prod_{p \in P} p^{e_p+1}$, implies that

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{Cok}(M(n)) \otimes (\mathbf{Z}/a\mathbf{Z}) \simeq B) = \mathbf{P}(Y \otimes (\mathbf{Z}/a\mathbf{Z}) \simeq B) = \frac{1}{|B|^u |\mathbf{Aut}(B)|} \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-u}).$$

The proof is then complete because $\mathbf{Cok}(M(n)) \otimes (\mathbf{Z}/a\mathbf{Z}) \simeq B$ if and only if $\mathbf{Cok}(M(n))_P \simeq B$.

5. MEDIUM PRIMES

In this section we prove the following, which we apply to medium primes for the proof of our main results.

Theorem 5.1. *There are constants $c_0, \eta, C_0, K_0 > 0$ such that we have the following. Let $n, u \geq 0$ be integers, p be a prime, and let $M_{n \times (n+u)}$ be a random matrix $n \times (n+u)$ with independent i.i.d. entries $\xi_n \in \mathbf{Z}/p\mathbf{Z}$. We further assume we have a real number α_n such that*

$$\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n = r) = 1 - \alpha_n \leq 1 - \frac{C_0 \log n}{n}.$$

Then we have

$$\mathbf{P}(\text{rank}(M_{n \times (n+u)}) \leq n-1) \leq 2p^{-\min(u+1, \eta n-1)} + K_0 e^{-c_0 \alpha_n n}$$

and

$$\mathbf{P}(\text{rank}(M_{n \times (n+u)}) \leq n-2) \leq 2p^{-\min(2u+4, \eta n-1)} + K_0 e^{-c_0 \alpha_n n}.$$

The proof of Theorem 5.1 has two main ingredients. First, we have a result from [21] that says that the first $n-k$ columns of $M_{n \times (n+u)}$ are likely to generate a subspace V such that the probability of the next column being in V is near to the probability of a uniform random column mod p being in V . (This result was originally stated in [19] by Maples, but [21] gives a corrected proof using the ideas of [19] and [30].)

Theorem 5.2 ([21, Theorems A.1 and A.4]). *There are constants $c, \eta, C_0, K > 0$ such that the following holds. Let $n, u \geq 0$ be integers with $u \leq \eta n$, p be a prime, and let $M_{n \times (n+u)}$ be a random matrix $n \times (n+u)$ with independent i.i.d. entries $\xi_n \in \mathbf{Z}/p\mathbf{Z}$. We further assume we have a real number α_n such that*

$$\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n = r) = 1 - \alpha_n \leq 1 - \frac{C_0 \log n}{n}.$$

For $-u < k \leq \eta n$, let X_{n-k+1} be the $(n-k+1)$ st column of $M_{n \times (n+u)}$, and W_{n-k} be the subspace spanned by the first $n-k$ columns of $M_{n \times (n+u)}$. Then there is an event \mathcal{E}_{n-k} on the σ -algebra generated by the first $n-k$ columns of $M_{n \times (n+u)}$, of probability at least $1 - 3e^{-c\alpha_n n}$, such that for any k_0 with $\max(0, k) \leq k_0 \leq \eta n$

$$\left| \mathbf{P}(X_{n-k+1} \in W_{n-k} | \mathcal{E}_{n-k} \wedge \text{codim}(W_{n-k}) = k_0) - p^{-k_0} \right| \leq K e^{-c\alpha_n n}.$$

We also refer the reader to Theorem 8.2 for a similar statement for the Laplacian with a complete proof. Note that for a uniform random $X \in (\mathbf{Z}/p\mathbf{Z})^n$, we have $\mathbf{P}(X \in V) = p^{-\text{codim}(V)}$. Thus, as long as we avoid certain rare bad events, as we consider more and more columns of our random matrices, the probability that the next column is in the span of the previous columns is close to what it would be if we were using uniform random matrices. The following result, proven in Section B in the Appendix, allows us to use that information to conclude that the rank distribution of our matrices is close to that of uniform random matrices. This theorem says that if sequences of random variables x_i and y_i have similar transition probabilities going from x_i to x_{i+1} and y_i to y_{i+1} , at least under conditions that are likely to be true, then the distributions of x_n and y_n must stay close.

Theorem 5.3. *Let $x_1, \dots, x_n, g_0, \dots, g_{n-1}$ be a sequence of random variables, and let $x_0 = 0$. Let y_1, \dots, y_n be a sequence of random variables, and let $y_0 = 0$. We assume each x_i, y_i takes on at most countably many values, and $g_i \in \{0, 1\}$. Suppose that for $0 \leq i \leq n-1$,*

$$\begin{aligned} \mathbf{P}(y_{i+1} = s | y_i = r) &= \mathbf{P}(x_{i+1} = s | x_i = r \text{ and } g_i = 1) + \delta(i, r, s) \\ &\text{for all } r \text{ and } s \text{ s.t. } \mathbf{P}(y_i = r) \mathbf{P}(x_i = r \text{ and } g_i = 1) \neq 0. \end{aligned}$$

Then for all $n \geq 0$ and any set A of values taken by x_n and y_n , we have

$$\begin{aligned} & |\mathbf{P}(x_n \in A) - \mathbf{P}(y_n \in A)| \\ & \leq \frac{1}{2} \sum_{i=0}^{n-1} \sum_r \sum_s |\delta(i, r, s)| \mathbf{P}(x_i = r) + \sum_{i=0}^{n-1} \mathbf{P}(g_i \neq 1), \end{aligned}$$

where r is summed over $\{r \mid \mathbf{P}(x_i = r) \neq 0 \text{ and } \mathbf{P}(y_i = r) \neq 0\}$ and s is summed over $\{s \mid \mathbf{P}(x_{i+1} = s) \neq 0 \text{ or } \mathbf{P}(y_{i+1} = s) \neq 0\}$.

We remark that our error bounds come from the δ 's and the complement of $g_i = 1$. The explicit form here will be extremely useful because in the sparse case δ and $\mathbf{P}(g_i \neq 0)$ are not small.

Proof of Theorem 5.1. We take η, C_0 as in Theorem 5.2. Since

$$\mathbf{P}(\text{rank}(M_{n \times (n+u+1)}) \leq m) \leq \mathbf{P}(\text{rank}(M_{n \times (n+u)}) \leq m),$$

it suffices to prove the theorem for $u \leq \lfloor \eta n \rfloor - 1$. Let X_m be the m -th column of $M_{n \times (n+u)}$, and W_m the subspace generated by X_1, \dots, X_m . Let $x_0 = 0$ and for $1 \leq i \leq \lfloor \eta n \rfloor + u$, define the random variable

$$x_i = \begin{cases} k_0 & \text{if } \text{rank}(W_{n-\lfloor \eta n \rfloor+i}) = n - \lfloor \eta n \rfloor + i - k_0 \text{ and } 0 \leq k_0 \leq u+1 \\ * & \text{if } \text{rank}(W_{n-\lfloor \eta n \rfloor+i}) \leq n - \lfloor \eta n \rfloor + i - u - 2. \end{cases}$$

In other words, x_i measures the deficiency $(n - \lfloor \eta n \rfloor + i) - \text{rank}(W_{n-\lfloor \eta n \rfloor+i})$ if this difference is not larger than $u+1$.

Let y_i be the analogous function for a uniform random matrix mod p for $1 \leq i \leq \lfloor \eta n \rfloor + u$. Let g_0 be the indicator function of the event that requires both $\text{rank}(W_{n-\lfloor \eta n \rfloor}) = n - \lfloor \eta n \rfloor$ and $\mathcal{E}_{n-\lfloor \eta n \rfloor}$ from Theorem 5.2. Let g_i be the indicator function for the event $\mathcal{E}_{n-\lfloor \eta n \rfloor+i}$ from Theorem 5.2, so from that theorem we have for $i \geq 1$ that $\mathbf{P}(g_i = 1) \geq 1 - 3e^{-c\alpha_n n}$.

We will apply Theorem 5.3 to the sequences x_i, y_i and g_i defined above. For this, we will estimate the error terms $\delta(i, b, a)$ for various values of i, a and b . First, note that if

$$\text{rank}(W_{n-\lfloor \eta n \rfloor+i}) \leq n - \lfloor \eta n \rfloor + i - u - 2,$$

then

$$\text{rank}(W_{n-\lfloor \eta n \rfloor+i+1}) \leq n - \lfloor \eta n \rfloor + i + 1 - u - 2.$$

So for $i \geq 1$,

$$\mathbf{P}(y_{i+1} = * | y_i = *) = \mathbf{P}(x_{i+1} = * | x_i = * \wedge g_i = 1) = 1.$$

Therefore, for $i \geq 1$ and all a we have

$$\delta(i, *, a) = 0.$$

Next, Theorem 5.2 gives that for $i \geq 1$ and $0 \leq k_0 \leq u+1$ (as $u+1 \leq \eta n$),

$$\delta(i, k_0, k_0) = |\mathbf{P}(y_{i+1} = k_0 | y_i = k_0) - \mathbf{P}(x_{i+1} = k_0 | x_i = k_0 \wedge g_i = 1)| \leq K e^{-c\alpha_n n}.$$

Furthermore, if $x_i = k_0$, the only possibility for x_{i+1} is either k_0 or $k_0 + 1$ (which should be interpreted as $*$ if $k_0 = u+1$). It then follows that for $i \geq 1$ and all k_0, ℓ , we have

$$\delta(i, k_0, \ell) = |\mathbf{P}(y_{i+1} = \ell | y_i = k_0) - \mathbf{P}(x_{i+1} = \ell | x_i = k_0 \wedge g_i = 1)| \leq K e^{-c\alpha_n n}.$$

At the initial position $i = 0$ we have (by definition) $y_0 = 0$ and so

$$\mathbf{P}(y_1 = 0 | y_0 = 0) = \mathbf{P}(y_1 = 0) = \prod_{j=0}^{n-\lfloor \eta n \rfloor} (1 - p^{-(\lfloor \eta n \rfloor + j)})$$

and Theorem 5.2 gives

$$\mathbf{P}(x_1 = 0 | x_0 = 0 \text{ and } g_0 = 1) \geq 1 - p^{-\lfloor \eta n \rfloor} - K e^{-c\alpha_n n}.$$

Thus for any ℓ ,

$$\begin{aligned}\delta(0, 0, \ell) &\leq p^{-\lfloor \eta n \rfloor} + K e^{-c\alpha_n n} + 1 - \prod_{j=0}^{n-\lfloor \eta n \rfloor} (1 - p^{-(\lfloor \eta n \rfloor + j)}) \\ &\leq p^{-\lfloor \eta n \rfloor} + K e^{-c\alpha_n n} + p^{-\lfloor \eta n \rfloor} / (1 - p^{-1}).\end{aligned}$$

We can apply Lemma 3.1 to find the $\mathbf{P}(X_{m+1} \notin W_m | \text{rank}(W_m) = m)$ for all $0 \leq m \leq n - \lfloor \eta n \rfloor - 1$. Taking union bound (see Corollary 3.2), we obtain

$$\mathbf{P}(\text{rank}(W_{n-\lfloor \eta n \rfloor}) = n - \lfloor \eta n \rfloor) \geq 1 - \alpha_n^{-1} (1 - \alpha_n)^{\lfloor \eta n \rfloor + 1}.$$

So

$$\mathbf{P}(g_0 = 0) \leq \alpha_n^{-1} (1 - \alpha_n)^{\lfloor \eta n \rfloor + 1} + 3e^{-c\alpha_n n}.$$

We now apply Theorem 5.3. The n from that theorem will be what we call $\lfloor \eta n \rfloor + u$ here. We conclude that for $k_0 = u$ or $u + 1$,

$$\begin{aligned}& \left| \mathbf{P}(x_{\lfloor \eta n \rfloor + u} = k_0) - \mathbf{P}(y_{\lfloor \eta n \rfloor + u} = k_0) \right| \\ & \leq \frac{1}{2} (\lfloor \eta n \rfloor + u) K e^{-c\alpha_n n} \cdot 2 + (\lfloor \eta n \rfloor + u) \cdot 3e^{-c\alpha_n n} \\ & \quad + \left(p^{-\lfloor \eta n \rfloor} + K e^{-c\alpha_n n} + p^{-\lfloor \eta n \rfloor} / (1 - p^{-1}) \right) + \left(\alpha_n^{-1} (1 - \alpha_n)^{\lfloor \eta n \rfloor + 1} + 3e^{-c\alpha_n n} \right).\end{aligned}$$

Here the first two terms are from the $i \geq 1$ summands in each sum, the $(\lfloor \eta n \rfloor + u)$ is from the sum over i , the sum over r of $\mathbf{P}(x_i = r)$ is 1, and the 2 is from the sum over s (for each r there are at most 2 values of s with non-zero $\delta(i, r, s)$). The second two terms are from the $i = 0$ summands.

Thus for $k_0 = u$ or $u + 1$, using $u \leq \eta n$,

$$|\mathbf{P}(x_{\lfloor \eta n \rfloor + u} = k_0) - \mathbf{P}(y_{\lfloor \eta n \rfloor + u} = k_0)| \leq 2(K + 3)\eta n e^{-c\alpha_n n} + 3e^{-c\alpha_n n} + 3p^{-\eta n + 1} + \alpha_n^{-1} (1 - \alpha_n)^{\eta n}.$$

Since (e.g. by [10])

$$\mathbf{P}(y_{\lfloor \eta n \rfloor + u} = u) = \prod_{j=1}^n (1 - p^{-j-u}) \geq 1 - \sum_{j \geq 1} p^{-j-u} = 1 - p^{-1-u} / (1 - p^{-1}),$$

and

$$\begin{aligned}\mathbf{P}(y_{\lfloor \eta n \rfloor + u} \geq u - 1) &= \left(1 + p^{-2-u} \left(\frac{1 - p^{-(n-1)}}{1 - p^{-1}} \right) \right) \prod_{j=2}^n (1 - p^{-j-u}) \\ &\geq 1 - \frac{p^{-4-2u}}{1 - p^{-1}} - \frac{p^{-(n-1)-2-u}}{1 - p^{-1}}.\end{aligned}$$

we have that

$$\mathbf{P}(\text{rank}(W_{n+u}) \leq n - 1) \leq 2p^{-1-u} + 2(K + 3)\eta n e^{-c\alpha_n n} + 3e^{-c\alpha_n n} + 3p^{-\eta n + 1} + \alpha_n^{-1} (1 - \alpha_n)^{\eta n}$$

and

$$\mathbf{P}(\text{rank}(W_{n+u}) \leq n - 2) \leq 2p^{-4-2u} + 2p^{-n} + 2(K + 3)\eta n e^{-c\alpha_n n} + 3e^{-c\alpha_n n} + 3p^{-\eta n + 1} + \alpha_n^{-1} (1 - \alpha_n)^{\eta n}.$$

Since $u \leq \eta n$, for some K_0 depending on K, c, η, C_0 , for all n we have

$$\mathbf{P}(\text{rank}(W_{n+u}) \leq n - 1) \leq 2p^{-1-u} + K_0 e^{-\min(c/2, \eta \log(2)/2, \log(2))\alpha_n n}$$

and

$$\mathbf{P}(\text{rank}(W_{n+u}) \leq n - 2) \leq 2p^{-4-2u} + K_0 e^{-\min(c/2, \eta \log(2)/2, \log(2))\alpha_n n}.$$

The result follows with $c_0 = \min(c/2, \eta \log(2)/2, \log(2))$. \square

6. LARGE PRIMES

In this section and the next we prove Proposition 2.3.

Notation for Sections 6 and 7: Throughout this and the next section, we fix $T > 0$ and for each positive integer n we let ξ_n be an α_n -balanced random integer with $|\xi_n| \leq n^T$. We define $M_{n \times (n+1)}$ to be the integral $n \times (n+1)$ matrix with entries i.i.d copies of ξ_n . We do not make a global assumption on the size of α_n , but we will need different assumptions on α_n for the various results in these two sections. We further fix $d > 0$. Let X_1, \dots, X_{n+1} be the columns of $M_{n \times (n+1)}$. We write $M_{n \times k}$ for the submatrix of $M_{n \times (n+1)}$ composed of the first k columns. Let W_k be the submodule of \mathbf{Z}^n spanned by X_1, \dots, X_k . We write X_k/p and W_k/p for their reductions mod p (and more generally use this notation to denote the reduction of an object from \mathbf{Z} to $\mathbf{Z}/p\mathbf{Z}$). We let

$$n_0 := n - \left\lfloor \frac{3 \log n}{\alpha_n} \right\rfloor. \quad (18)$$

Let

$$\mathfrak{P}_n := \left\{ p \text{ prime, } p \geq e^{d\alpha_n n} \right\}.$$

Let $\mathcal{E}_{\neq 0}$ be the event that $\det(M_{n \times n}) \neq 0$. As mentioned in the introduction section, from [21] and also by taking the limit as $p \rightarrow \infty$ in Theorem 5.1, we have

$$\mathbf{P}(\mathcal{E}_{\neq 0}) \geq 1 - K_0 e^{-c_0 \alpha_n n}$$

for absolute constants c_0, K_0 . Our strategy is as follows. We consider the columns of the matrix one at a time, and check if they are in the span of the previous columns modulo p for each prime in \mathfrak{P}_n . We cannot control whether this happens, as \mathfrak{P}_n contains too many primes, but each p for which this happens is put on a “watch list” (called \mathfrak{W}_k) and necessarily divides the determinant of $M_{n \times n}$. If the watch list grows too large, since all the primes in the watch list are large, then too large a number divides the determinant, and $M_{n \times n}$ must be singular. However, we have already bounded the probability of that occurring. Otherwise, if our watch list is not too large, for each prime in the watch list, we can bound the probability that the next column is in the span of the previous columns mod that prime.

Let \mathfrak{W}_k be the set of primes $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-1$. Let \mathcal{C}_k be the event that $|\mathfrak{W}_k| \leq (2T+1) \log n / (2d\alpha_n)$ (the watch list is under control). Note that any $p \in \mathfrak{W}_k$ for $k \leq n$ must divide $\det(M_{n \times n})$. By Hadamard’s bound, $|\det(M_{n \times n})| \leq n^{n/2} n^{Tn}$, and so in particular, when $\bar{\mathcal{C}}_k$ occurs (“the watch list is out of control”) then $\det(M_{n \times n}) = 0$. (Recall we write $\bar{\mathcal{E}}$ for the complement of an event \mathcal{E} .) Let \mathcal{D}_k be the event that there is a $p \in \mathfrak{W}_k$ such that $\text{rank}(W_k/p) \leq k-2$ (the rank drops), this is the event we want to avoid.

We will show $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k)$ is large. The goal is to conclude that $\mathbf{P}(\bar{\mathcal{C}}_n \vee \bar{\mathcal{D}}_n)$ is large, and since we know that $\mathbf{P}(\bar{\mathcal{C}}_n)$ is small, we can conclude that $\mathbf{P}(\bar{\mathcal{D}}_n)$ is large, as desired. Note that since $\mathfrak{W}_k \subset \mathfrak{W}_{k+1}$, we have that $\bar{\mathcal{C}}_k \subset \bar{\mathcal{C}}_{k+1}$. Thus

$$\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \bar{\mathcal{C}}_k) = 1. \quad (19)$$

It remains to estimate $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \mathcal{C}_k \wedge \bar{\mathcal{D}}_k)$. We condition on the exact values of X_1, \dots, X_k where $\mathcal{C}_k \wedge \bar{\mathcal{D}}_k$ holds, and so there are at most $(2T+1) \log n / (2d\alpha_n)$ primes $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-1$ and no prime $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-2$. In $\bar{\mathcal{D}}_{k+1}$, as long as for each $p \in \mathfrak{W}_k$, we have $X_{k+1}/p \notin W_k/p$. Consider one prime $p \in \mathfrak{W}_k$, and let V be the value of W_k/p that the conditioned X_1, \dots, X_k give. From Lemma 3.1, $\mathbf{P}(X_{k+1}/p \in V) \leq (1 - \alpha_n)^{n-(k-1)}$. Thus,

$$\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \mathcal{C}_k \wedge \bar{\mathcal{D}}_k) \geq 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) (1 - \alpha_n)^{n-(k-1)}.$$

In particular, by (19) we conclude that

$$\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k) \geq 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) (1 - \alpha_n)^{n-(k-1)}.$$

Then inductively, we have

$$\mathbf{P}(\bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k) \geq 1 - \sum_{i=1}^{k-1} \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) (1 - \alpha_n)^{n-(i-1)} = 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) \frac{(1 - \alpha_n)^{n-k+2}}{\alpha_n}.$$

We defined $n_0 := n - \lfloor \frac{3 \log n}{\alpha_n} \rfloor$ above and so if we let $k = n_0$ then if we assume $\alpha_n \geq n^{-1}$, we have that

$$\mathbf{P}(\bar{\mathcal{C}}_{n_0} \vee \bar{\mathcal{D}}_{n_0}) \geq 1 - O_{d,T}(n^{-1/2}).$$

Certainly as k gets very close to n , W_k has very small codimension, and so Odlyzko's bound will not continue to be strong enough. Thus for the remaining k we will have to use a different bound.

6.1. Proof of Proposition 2.3 when $\alpha_n \geq n^{-1/6+\varepsilon}$. First, in this section, we will prove Proposition 2.3 for the denser case $\alpha_n \geq n^{-1/6+\varepsilon}$. For these larger α_n we can present a simpler proof than in the case when α_n might be as small as $n^{-1+\varepsilon}$. Odlyzko's bound is sharp for some spaces, e.g. the hyperplane of vectors with first coordinate 0, and so if we need to improve on Odlyzko's bound we cannot expect to do it for all spaces at once. The overall strategy is to see that apart from some bad subspaces, we can improve on Odlyzko's bound, and we can also prove that it is unlikely that W_k is one of those bad spaces. At this level of generality, this description fits the small and medium primes sections. However, the specifics are very different, because the small and medium primes sections treat one prime at a time, and now we are in a regime where there are just too many possible primes to add the probability of W_k being bad over all the primes (e.g. adding $\mathbf{P}(\bar{\mathcal{E}}_{n-k})$ from Theorem 5.2 over all primes up to $n^{n/2+Tn}$ gives too big a result). On the other hand, we do not need the same strength of improvement over Odlyzko's bound that Theorem 5.2 provides, because the bound on the probability of $X_{k+1}/p \in W_k/p$ only has to be added over the small number of primes in the watch list. The following lemma balances these requirements, and its proof will be delayed till the end of this subsection.

Lemma 6.2. *Suppose that $\alpha_n \geq 6 \log n/n$. Then there is a set of submodules \mathcal{S} of \mathbf{Z}^n such that*

$$\mathbf{P}(W_{n_0} \in \mathcal{S}) \geq 1 - e^{-\alpha_n n/8}$$

and for any prime $p \geq e^{d\alpha_n n}$, and any submodule $H \in \mathcal{S}$, for any proper subspace H' of $(\mathbf{Z}/p\mathbf{Z})^n$ containing H/p ,

$$\mathbf{P}(X/p \in H') = O_{d,T}\left(\frac{\sqrt{\log n}}{\alpha_n \sqrt{n}}\right),$$

where X is any column of $M_{n \times n}$.

Now, we will also condition on \mathcal{G} , which we define to be the event that $W_{n_0} \in \mathcal{S}$ (i.e., W_{n_0} is \mathcal{G} ood). We then have

$$\mathbf{P}((\bar{\mathcal{C}}_{n_0} \vee \bar{\mathcal{D}}_{n_0}) \wedge \mathcal{G}) \geq 1 - e^{-\alpha_n n/8} + O_{d,T}(n^{-1/2}),$$

Now let $n_0 \leq k \leq n$. As before, since $\bar{\mathcal{C}}_k \subset \bar{\mathcal{C}}_{k+1}$, we have

$$\mathbf{P}((\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1}) \wedge \mathcal{G} | \bar{\mathcal{C}}_k \wedge \mathcal{G}) = 1.$$

It remains to estimate $\mathbf{P}((\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1}) \wedge \mathcal{G} | \bar{\mathcal{C}}_k \wedge \bar{\mathcal{D}}_k \wedge \mathcal{G})$. Again, we condition on exact values of X_1, \dots, X_k such that $\mathcal{C}_k, \bar{\mathcal{D}}_k, \mathcal{G}$ hold. Then $\bar{\mathcal{D}}_{k+1}$ holds unless for some $p \in \mathfrak{W}_k$ we have $X_{k+1}/p \in W_k/p$. Since \mathcal{C}_k holds, we have a bound on the size of \mathfrak{W}_k , and since \mathcal{G} holds, we can use Lemma 6.2 to bound the probability that X_{k+1}/p is in W_k/p . (Note that even if $k = n$, for $p \in \mathfrak{W}_k$, we have that W_k/p is a proper subspace of $(\mathbf{Z}/p\mathbf{Z})^n$, and that $n^{-1/6} \geq 2(\log n)/n$.) We conclude that

$$\mathbf{P}((\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1}) \wedge \mathcal{G} | \bar{\mathcal{C}}_k \wedge \bar{\mathcal{D}}_k \wedge \mathcal{G}) \geq 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n}\right) O_{d,T}\left(\frac{\sqrt{\log n}}{\alpha_n \sqrt{n}}\right).$$

Inductively, starting from $k = k_0$ we then have

$$\mathbf{P}((\bar{\mathcal{C}}_n \vee \bar{\mathcal{D}}_n) \wedge \mathcal{G}) \geq 1 - \lfloor \frac{3 \log n}{\alpha_n} \rfloor O_{d,T}\left(\frac{\log^{1.5} n}{\alpha_n^2 \sqrt{n}}\right) - e^{-\alpha_n n/8} + O_{d,T}(n^{-1/2}).$$

So then,

$$\mathbf{P}(\bar{\mathcal{D}}_n) \geq 1 - O_{d,T}\left(\frac{\log^{2.5} n}{\alpha_n^3 \sqrt{n}}\right) - K_0 e^{-c_0 \alpha_n n} - e^{-\alpha_n n/8} + O_{d,T}(n^{-1/2}). \quad (20)$$

To this end, since $\alpha_n \geq n^{-1/6+\varepsilon}$, we have that

$$\mathbf{P}(\bar{\mathcal{D}}_n) \geq 1 - O_{d,T,\varepsilon}(n^{-\varepsilon}),$$

which is exactly Equation (10). Equation (9) follows similarly, with $\lfloor 3 \log n / \alpha_n \rfloor$, the number of steps in the induction, replaced by $\lfloor 3 \log n / \alpha_n \rfloor + 1$.

The key, of course, is to now verify Lemma 6.2, which is the heart of the proof of Proposition 2.3 for $\alpha_n \geq n^{-1/6+\varepsilon}$. We need a way to bound the probability that $X_{k+1}/p \in W_k/p$ that works for all $p \in \mathfrak{W}_k$ and is effective for large k . For this, we introduce a version of the classical Erdős-Littlewood-Offord result in $\mathbf{Z}/p\mathbf{Z}$.

Theorem 6.3 (forward Erdős-Littlewood-Offord, for non-sparse vectors). *Let $X \in (\mathbf{Z}/p\mathbf{Z})^n$ be a random vector whose entries are i.i.d. copies of a random variable ν_n satisfying $\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\nu_n = r) \leq 1 - \alpha_n$. Suppose that $w \in (\mathbf{Z}/p\mathbf{Z})^n$ has at least n' non-zero coefficients, and $\alpha_n \geq 4/n'$. Then we have*

$$|\mathbf{P}(X \cdot w = r) - \frac{1}{p}| \leq \frac{2}{\sqrt{\alpha_n n'}}.$$

A proof of this result due to Maples (based on an argument by Halász) can be seen in [19, Theorem 2.4] (see also [21, Theorem A.21].) To use Theorem 6.3 we need to know it is unlikely that W_k has normal vectors with few non-zero entries. First, we will see this is true over \mathbf{R} . The approach is standard: there are few sparse vectors and by the Odlyzko's bound each is not that likely to be normal to W_k .

Lemma 6.4 (Sparse normal vectors over \mathbf{R} unlikely). *Suppose $\alpha_n \geq \log n/n$ and $k \geq n/2$. For n sufficiently large (in an absolute sense), with probability at least $1 - e^{-\alpha_n n/8}$, the random subspace X_1, \dots, X_k does not have a non-trivial normal vector with less than $\alpha_n n / (32 \log n)$ non-zero entries.*

Proof. (of Lemma 6.4) Let $l = \lfloor \alpha_n n / (32 \log n) \rfloor$. With a loss of a multiplicative factor $\binom{n}{l}$ in probability, we assume that there exists a vector $w = (w_1, \dots, w_l, 0, \dots, 0)$ which is normal to X_1, \dots, X_k . Let $M_{l \times k}$ be the matrix with columns given by the first l coordinates of each of X_1, \dots, X_k , which has rank at most $l-1$. With a loss of a multiplicative factor l in probability, we assume that the first row of $M_{l \times k}$ belongs to the subspace H generated by the other $l-1$ rows. However, as H has codimension at least $k-l$, Lemma 3.1 implies a bound $(1 - \alpha_n)^{k-l}$ for this event. Putting together, the event under consideration is bounded by

$$\binom{n}{l} \times l \times (1 - \alpha_n)^{k-l} \leq e^{(\log(32 \log n / \alpha_n) + 1) \alpha_n n / (32 \log n)} \times e^{\log(\alpha_n n / (32 \log n))} \times e^{-\alpha_n (k-l)}.$$

We then have that the exponent of e in the above bound is

$$\leq \left(\frac{\log \log n}{\log n} + \frac{\log(\alpha_n^{-1})}{\log n} + \frac{\log(32)}{\log n} + \frac{1}{\log n} \right) \alpha_n n / 32 + \log(\alpha_n n) - \alpha_n n / 4 \leq -\alpha_n n / 8$$

for n sufficiently large so that

$$\left(\frac{\log \log n}{\log n} + \frac{\log(32)}{\log n} + \frac{1}{\log n} \right) \leq 1$$

and so that $\log(\alpha_n n) \leq \alpha_n n / 16$ (which happens for $\alpha_n n \geq 22$, which is implied by $\log n \geq 22$). \square

We could prove a similar lemma to Lemma 6.4 over $\mathbf{Z}/p\mathbf{Z}$ for each p , but we could not sum the probabilities $e^{-\alpha_n n/8}$ of sparse normal vectors over any meaningful range of primes $> e^{d_{\text{con}}}$. However, now we will prove a deterministic lemma, that lets us lift normal vectors with few non-zero entries from characteristic p , for large p , to characteristic 0. Then there is only one bad event to avoid instead of one for each p . This aspect of our argument is unlike previous approaches and uses critically a lower bound on p .

Lemma 6.5 (Lifting sparse normal vectors from $\mathbf{Z}/p\mathbf{Z}$ to \mathbf{R}). *Let k, l, n be positive integers, and M a $l \times k$ matrix with integer entries $|M_{ij}| \leq n^T$. If p is a prime larger than $e^{(k \log k)/2 + kT \log n}$, then the rank of M over \mathbf{Q} is equal to the rank of M/p over $\mathbf{Z}/p\mathbf{Z}$. This has the following corollaries.*

- (1) *If $Z_1, \dots, Z_k \in \mathbf{Z}^l$ are vectors with entries $|Z_{ij}| \leq n^T$, and $Z_1/p, \dots, Z_k/p$ are linearly dependent in $(\mathbf{Z}/p\mathbf{Z})^l$, then Z_1, \dots, Z_k are also linearly dependent in \mathbf{Z}^l .*
- (2) *Let $Z_1, \dots, Z_l \in \mathbf{Z}^m$ be vectors with entries $|Z_{ij}| \leq n^T$. If there is a non-zero vector $w \in (\mathbf{Z}/p\mathbf{Z})^m$ with at most k non-zero entries that is normal to $Z_1/p, \dots, Z_l/p$, then there is a non-zero vector $w' \in \mathbf{Z}^m$ with at most k non-zero entries and normal to Z_1, \dots, Z_l .*
- (3) *The kernel of the map $M : \mathbf{Z}^k \rightarrow \mathbf{Z}^l$ surjects onto the kernel of the map $M : (\mathbf{Z}/p\mathbf{Z})^k \rightarrow (\mathbf{Z}/p\mathbf{Z})^l$.*

Proof. (of Lemma 6.5) The rank is the greater r such that an $r \times r$ minor has non-zero determinant. Since by Hadamard's bound, the determinant of a $r \times r$ minor is at most $e^{(r \log r)/2 + rT \log n}$, for primes $p > e^{(k \log k)/2 + kT \log n}$ and $r \leq k$, a determinant of an $r \times r$ minor vanishes in \mathbf{Z} if and only if it vanishes mod p . Thus we conclude the main statement of the lemma. For the first corollary, consider the $\ell \times k$ matrix with Z_1, \dots, Z_k as columns. The vectors are linearly dependent if and only if the matrix has rank less than k . For the second corollary, assume that $\sigma = \text{supp}(w) \subset [m]$ and $Z_1|_\sigma, \dots, Z_l|_\sigma$ are the restrictions of Z_1, \dots, Z_l over the components in σ . By definition, the k row vectors of the matrix formed by $Z_1|_\sigma, \dots, Z_l|_\sigma$ are dependent when reduced mod p , and thus these vectors are dependent over \mathbf{Z} . This gives a non-zero $w' \in \mathbf{Z}^m$ with $|\text{supp}(w')| \leq k$ that is normal to Z_1, \dots, Z_l . For the third corollary, we can express M under the Smith normal form $M = S_1 D S_2$, where $S_1 \in \text{GL}_\ell(\mathbf{Z})$ and $S_2 \in \text{GL}_k(\mathbf{Z})$ and D is an integral diagonal matrix. Then since the ranks of M and D agree over \mathbf{Q} and over $\mathbf{Z}/p\mathbf{Z}$, we conclude that D has the same rank over \mathbf{Q} or $\mathbf{Z}/p\mathbf{Z}$. This implies that the only diagonal entries of D that are divisible by p are the ones that are 0. From this it follows that the kernel of $D : \mathbf{Z}^k \rightarrow \mathbf{Z}^l$ surjects onto the kernel of $D : (\mathbf{Z}/p\mathbf{Z})^k \rightarrow (\mathbf{Z}/p\mathbf{Z})^l$. Multiplication by S_2^{-1} on the left takes these kernels of D to the corresponding kernels of M , and the statement follows. \square

Putting this all together, we can now prove Lemma 6.2. The choices of parameters are rather delicate here, e.g. we could obtain more non-zero coordinates of a normal vector than Lemma 6.4 provides, but then we could not use Lemma 6.5 to lift those non-zero coordinates.

Proof of Lemma 6.2. We let $k = c\alpha_n n / \log n$, where $c < 1/32$ is a sufficiently small constant (in terms of d, T) such that

$$e^{(k \log k)/2 + kT \log n} < e^{d\alpha_n n}.$$

Since $\alpha_n \geq 6 \log n / n$, it follows that $n_0 \geq n/2$, and we can apply Lemma 6.4 and find that for sufficiently large n , with probability at least $1 - e^{-\alpha_n n/8}$, W_{n_0} does not have a normal vector with less than k entries. Let \mathcal{S} be the set of submodules of \mathbf{Z}^n that do not have a normal vector with less than k non-zero coordinates. Then by Lemma 6.5, for each prime $p \geq e^{d\alpha_n n}$, if $W_{n_0} \in \mathcal{S}$, then the space W_{n_0}/p (and thus any space containing this space) does not have a non-trivial normal vector with less than k non-zero coordinates. Since $\alpha_n \geq 6 \log n / n$, for n sufficiently large in terms of d, c , we have that $e^{d\alpha_n n} \geq \sqrt{\alpha_n k}$. Thus by Theorem 6.3, for any prime $p \geq e^{d\alpha_n n}$ the following holds. Let H be a subspace of $(\mathbf{Z}/p\mathbf{Z})^n$ that does not have a non-trivial normal vector with less than k non-zero entries, and then for any proper subspace H' of $(\mathbf{Z}/p\mathbf{Z})^n$ containing H and with normal vector w ,

$$\mathbf{P}(X \in H') \leq \mathbf{P}(X \cdot w = 0) \leq \frac{3}{\sqrt{\alpha_n k}} = \frac{3\sqrt{\log n}}{\alpha_n \sqrt{cn}}.$$

\square

6.6. Proof of Proposition 2.3 in general. Equation (20) shows exactly why $n^{-1/6+\varepsilon}$ is the threshold exponent for α_n such that the above method can work, as the error bound has an $\alpha_n^3 n^{1/2}$ in the denominator. Thus, to obtain results that can work for smaller α_n , we need a further improvement on Odlyzko's bound, which requires that we consider further bad subspaces besides those with sparse normal vectors. We have the following upgrade to Lemma 6.2, whose proof is rather more involved than that of Lemma 6.2 will be completed in the next section, and again, is the heart of the proof.

Lemma 6.7. *There is an absolute constant $c_2 > 0$ such that the following holds. Suppose that $\alpha_n \geq n^{-1+\varepsilon}$. There is a set \mathcal{S}' of submodules of \mathbf{Z}^n , such that*

$$\mathbf{P}(W_{n_0} \in \mathcal{S}') \geq 1 - e^{-c_2 \alpha_n n}$$

and for n sufficiently large given d, ε, T , any prime $p \geq e^{d\alpha_n n}$, any submodule $H \in \mathcal{S}'$ and any proper subspace H' of $(\mathbf{Z}/p\mathbf{Z})^n$ containing H/p ,

$$\mathbf{P}(X/p \in H') \leq n^{-3},$$

where X is any column of $M_{n \times n}$.

The rest of the proof goes the same as after Lemma 6.2, replacing \mathcal{S} with \mathcal{S}' . We conclude that for $\alpha_n \geq n^{-1+\varepsilon}$, we have that $\mathbf{P}(\bar{\mathcal{D}}_k) \geq 1 - O_{d,T,\varepsilon}(n^{-1/2})$, for $k = n, n+1$, which proves Proposition 2.3. We have not attempted to optimize the error, or even record in Proposition 2.3 the error this argument proves (as we wanted to give a weaker statement that could be proved by the simpler argument above when the α_n were not too small).

7. PROOF OF LEMMA 6.7: ENUMERATION OF STRUCTURES

Instead of only avoiding sparse normal vectors, in Lemma 6.7 we will avoid normal vectors with more general structure. We now need to make some definitions necessary to describe this structure.

7.1. Additive structures in abelian groups. Let G be an (additive) abelian group.

Definition 7.2. A set Q is a *generalized arithmetic progression* (GAP) of rank r if it can be expressed as in the form

$$Q = \{a_0 + x_1 a_1 + \cdots + x_r a_r \mid M_i \leq x_i \leq M'_i \text{ and } x_i \in \mathbf{Z} \text{ for all } 1 \leq i \leq r\}$$

for some elements a_0, \dots, a_r of G , and for some integers M_1, \dots, M_r and M'_1, \dots, M'_r .

It is convenient to think of Q as the image of an integer box $B := \{(x_1, \dots, x_r) \in \mathbf{Z}^r \mid M_i \leq x_i \leq M'_i\}$ under the linear map

$$\Phi : (x_1, \dots, x_r) \mapsto a_0 + x_1 a_1 + \cdots + x_r a_r.$$

Given Q with a representation as above

- the numbers a_i are *generators* of Q , the numbers M_i and M'_i are *dimensions* of Q , and $\text{Vol}(Q) := |B|$ is the *volume* of Q associated to this presentation (i.e. this choice of a_i, M_i, M'_i);
- we say that Q is *proper* for this presentation if the above linear map is one to one, or equivalently if $|Q| = |B|$;
- If $-M_i = M'_i$ for all $i \geq 1$ and $a_0 = 0$, we say that Q is *symmetric* for this presentation. For later use, for a symmetric progression Q and for $t > 0$ we also define that

$$Q_t := \{x_1 a_1 + \cdots + x_r a_r \mid -tM_i \leq x_i \leq tM_i \text{ and } x_i \in \mathbf{Z} \text{ for all } 1 \leq i \leq r\}.$$

The following inverse-type idea, which was first studied by Tao and Vu about ten years ago (see for instance [32]), will allow us prove bounds much sharper than Theorem 6.3.

Theorem 7.3 (inverse Erdős-Littlewood-Offord). *Let $\varepsilon < 1$ and C be positive constants. Let n be a positive integer. Assume that p is a prime that is larger than $C'n^C$ for a sufficiently large constant C' depending on ε and C . Let ν be a random variable taking values in $\mathbf{Z}/p\mathbf{Z}$ which is α_n -balanced, that is $\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\nu = r) \leq 1 - \alpha_n$ where $\alpha_n \geq n^{-1+\varepsilon}$. Assume $w = (w_1, \dots, w_n) \in (\mathbf{Z}/p\mathbf{Z})^n$ such that*

$$\rho(w) := \sup_{a \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\nu_1 w_1 + \cdots + \nu_n w_n = a) \geq n^{-C},$$

where ν_1, \dots, ν_n are iid copies of ν . Then for any $n^{\varepsilon/2} \alpha_n^{-1} \leq n' \leq n$ there exists a proper symmetric GAP Q in $\mathbf{Z}/p\mathbf{Z}$ of rank $r = O_{C,\varepsilon}(1)$ which contains all but n' coordinates of w , where

$$|Q| \leq \max \left\{ 1, O_{C,\varepsilon}(\rho(w)^{-1}/(\alpha_n n')^{r/2}) \right\}.$$

When α_n is a constant, we then recover a variant of [22, Theorem 2.5]. The new, but not too surprising, aspects here are that the result works for small α_n and for $\mathbf{Z}/p\mathbf{Z}$ for large enough p . A proof of Theorem 7.3 will be presented in Appendix C by modifying the approach of [22]. We remark that it is in the proof of Theorem 7.3 where the requirement $\alpha_n n' \geq n^{\varepsilon/2}$ is crucial (which henceforth requires α_n to be at least $n^{\varepsilon/2-1}$) to guarantee polynomial growth of certain sumsets (see [47]). We see that $Q = \{0\}$ includes the special case of sparse w . Theorem 7.3 is much sharper than Theorem 6.3, and relates the volume of the GAP involved to the bound for $\rho(w)$.

We let

$$n' := \lceil n^{\varepsilon/2} \alpha_n^{-1} \rceil$$

and $m = n - n'$ for the rest of this section, and we will apply Theorem 7.3 with this choice of n' and $C = 3$. Thus it will be convenient to let C_ε be the maximum of C' and the constants from the $O_{C,\varepsilon}$ notation

bounding the rank of volume of $|Q|$ in Theorem 7.3 applied with $C = 3$. We call a GAP Q *well-bounded* if it is of rank $\leq C_\varepsilon$ and $|Q| \leq C_\varepsilon n^3$. We call a vector w *structured* if it is non-zero, and there exists a symmetric well-bounded GAP Q such that all but n' coordinates of w belong to Q . Note that it is not always true that $\rho(w) = n^{-O(1)}$ if w is structured in this sense.

Our general approach is to see that it is not too likely for W_{n_0}/p to have structured normal vectors, where we recall n_0 from (18). We need to handle the case of $r = 0$ separately from the case of $r \geq 1$, as in the latter case we will use the $(\alpha_n n')^{r/2}$ term crucially. Now, we will give a very different approach to proving W_{n_0}/p is unlikely to have sparse vectors than we used in Lemma 6.4, as Lemma 6.4 is too weak for small α_n . The method of Lemma 7.4 will actually give better results as α_n gets smaller, while Lemma 6.4 gets worse. This method will automatically control sparse vectors for all large primes at once, without any lifting from characteristic p to characteristic 0. Notably, the bound we get from Lemma 7.4 will be the largest term in our error.

Lemma 7.4 (Extremely sparse normal vectors). *There are absolute constants c_1, C_0 such that the following holds. Let $\beta_n := 1 - \max_{x \in \mathbf{Z}} \mathbf{P}(\xi_n = x)$, and assume $\beta_n \geq C_0 \log n/n$ and $\alpha_n \geq 6 \log n/n$. For $n \geq 2$, the following happens with probability at most $e^{-c_1 \beta_n n/2}$: for some prime $p > 2n^T$, the space W_{n_0}/p has a non-zero normal vector with at most $144\beta_n^{-1}$ non-zero coordinates.*

Proof. (of Lemma 7.4) In fact, we will show that the following holds with probability at least $1 - e^{-c_1 \beta_n n/2}$. For any $1 \leq t \leq 144\beta_n^{-1}$, and any $\sigma \in \binom{[n]}{t}$, there are at least two columns X_i, X_j whose restriction $(X_j - X_i)|_\sigma$ has exactly one non-zero entry. We first show that this will suffice to prove the lemma. Since $(X_j - X_i)|_\sigma$ has a unique non-zero entry, and all its entries are at most $2n^T$ in absolute value, for any prime $p > 2n^T$ we have that $(X_j/p - X_i/p)|_\sigma$ has exactly one non-zero entry. Suppose we had a normal vector w to W_{n_0}/p with $1 \leq t \leq 144\beta_n^{-1}$ non-zero entries, and let σ be the indices of those entries. Since $w|_\sigma$ is normal to $(X_j/p - X_i/p)|_\sigma$, that would imply that one of the σ coordinates of w is zero, which contradicts the choice of σ .

Now we prove the claim from the beginning of the proof. Our method is similar to that of [11 Lemma 3.2] and [21 Claim A.9]. For $k \in \{1, 3, \dots, 2\lfloor(n_0-1)/2\rfloor+1\}$, consider the vectors $Y_i = X_{k+1} - X_k$. The entries of these vectors are iid copies of the symmetrized random variable $\psi = \xi - \xi'$, where ξ', ξ are independent and have distribution ξ_n . With $1 - \beta'_n := \mathbf{P}(\psi = 0)$, then $\beta_n \leq \beta'_n \leq 2\beta_n$ as this can be seen by

$$(1 - \beta_n)^2 \leq \max_x \mathbf{P}(\xi = x)^2 \leq \sum_x \mathbf{P}(\xi = x)^2 = \mathbf{P}(\psi = 0) \leq \max_x \mathbf{P}(\xi = x) = 1 - \beta_n. \quad (21)$$

Now let p_σ be the probability that all $Y_i|_\sigma, i \in \{1, 3, \dots, 2\lfloor(n_0-1)/2\rfloor+1\}$ fail to have exactly one non-zero entry (in \mathbf{Z}), then by independence of the columns and of the entries

$$p_\sigma = (1 - t\beta'_n(1 - \beta'_n)^{t-1})^{\lfloor(n_0+1)/2\rfloor} \leq (1 - t\beta'_n e^{-(t-1)\beta'_n})^{n_0/2} \leq e^{-nt\beta'_n e^{-(t-1)\beta'_n}/4}.$$

(Recall since $\alpha_n \geq 6 \log n/n$ we have $n_0 \geq n/2$.) Notice that as $1 \leq t \leq 144\beta_n^{-1}$, we have $e^{-(t-1)\beta'_n}/4 \geq c_1$ for some positive constant c_1 , and hence

$$e^{-nt\beta'_n e^{-(t-1)\beta'_n}/4} \leq (e^{-c_1 n \beta'_n})^t \leq n^{-c_1 C_0 t/2} e^{-c_1 n \beta_n/2},$$

for any $C_0 > 0$. Thus

$$\sum_{1 \leq t \leq 144\beta_n^{-1}} \sum_{\sigma \in \binom{[n]}{t}} p_\sigma \leq \sum_{1 \leq t \leq 144\beta_n^{-1}} \binom{n}{t} n^{-c_1 C_0 t/2} e^{-c_1 n \beta_n/2} \leq \sum_{1 \leq t \leq 144\beta_n^{-1}} (n^t n^{-c_1 C_0 t/2}) e^{-c_1 n \beta_n/2} < e^{-c_1 n \beta_n/2},$$

provided that $n \geq 2$ and C_0 is sufficiently large in terms of c_1 . \square

The downside of Lemma 7.4 is that it is rather weak for constant α_n . So it needs be combined with an improvement of Lemma 6.4. For the improvement, we use Littlewood-Offord (Theorem 6.3) in place of Odlyzko's bound. However, that substitution only makes sense once we have k non-zero coordinates in our normal vector and $\alpha_n k$ is at least a constant. Luckily, Lemma 7.4 provides us with exactly that. This strategy is analogous to that used in the proof of [21 Proposition A.8].

Lemma 7.5 (Moderately sparse normal vectors). *There exist absolute constants c_0, C_0 such that the following holds. Let $\beta_n := 1 - \max_{x \in \mathbf{Z}} \mathbf{P}(\xi_n = x)$. Assume $\alpha_n \geq \frac{C_0 \log n}{n}$ and let p be a prime $> 2n^T$. The following happens with probability at most $(2/3)^{n/4}$: the space W_{n_0}/p has a non-zero normal vector w with $144\beta_n^{-1} \leq |\text{supp}(w)| \leq c_0 n$.*

Note Lemma 7.5 only bounds the probability of sparse normal vectors modulo one p at a time, unlike Lemma 7.4, which controls sparse normal vectors modulo all sufficiently big primes.

Proof. (of Lemma 7.5) For $\sigma \subset [n]$ with $144\beta_n^{-1} \leq t = |\sigma| \leq c_0 n$, consider the event that W_{n_0}/p is normal to a vector w with $\text{supp}(w) = \sigma$ but not to any other vector of smaller support size. With a loss of a multiplicative factor $\binom{n}{t}$ in probability, we assume that $\sigma = \{1, \dots, t\}$. Consider the submatrix $M_{t \times n_0}$ of $M_{n \times n}$ consisting of the first t rows and first n_0 columns of $M_{n \times n}$. Since the restriction $w|_{\sigma}$ of w to the first t coordinates is normal to all the columns of $M_{t \times n_0}/p$, the matrix $M_{t \times n_0}/p$ has rank $t-1$ (if $p=0$, we mean rank over \mathbf{R}). With a loss of a multiplicative factor $\binom{n_0}{t-1}$ in probability, we assume that the column space of $M_{t \times n_0}/p$ is spanned by its first $t-1$ columns.

Note that for $p > 2n^T$, the value of ξ_n is determined by its value mod p , and so $\beta_n = 1 - \max_{x \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n/p = x)$. If we fix X_1, \dots, X_{t-1} such that $W_{t-1}|_{\sigma}/p$ has a normal vector with all t coordinates non-zero, then by Theorem 6.3, the probability that $X_i|_{\sigma}/p \in W_{t-1}|_{\sigma}/p$ for all $t \leq i \leq n_0$ is at most

$$\left(\frac{1}{p} + \frac{2}{\sqrt{\beta_n t}}\right)^{n_0-t+1} \leq \left(\frac{1}{p} + \frac{2}{\sqrt{\beta_n t}}\right)^{(1-2c_0)n} \leq \left(\frac{2}{3}\right)^{n/2}.$$

The first inequality follows as long as $C_0 \geq 3/c_0$ as then we have $c_0 n \geq 3 \log n / \alpha_n$ and $n_0 \geq n - c_0 n$. Thus the total probability of the event in the lemma is at most

$$\sum_{144\beta_n^{-1} \leq t \leq c_0 n} \binom{n}{t}^2 \left(\frac{2}{3}\right)^{n/2} \leq \left(\frac{2}{3}\right)^{n/4}.$$

provided that c_0 is sufficiently small absolutely. \square

Now we will show that the probability of having a structured normal vector for a GAP of rank $r \geq 1$ (which was defined in the discussion following Theorem 7.3) is extremely small.

Lemma 7.6 (Structured, but not sparse, normal vectors). *Let $\alpha_n \geq n^{-1+\varepsilon}$. Let p be a prime $p \geq C_{\varepsilon} n^3$. The following event happens with probability $O_{\varepsilon}(p^{C_{\varepsilon} n^{-\varepsilon n/5}})$: the space W_{n_0}/p has a structured normal vector w , and W_{n_0}/p does not have a non-zero normal vector w' such that $|\text{supp}(w')| \leq c_0 n$ with c_0 from Lemma 7.5.*

Very roughly speaking, aside from the choices of parameters for the GAPs that might contain the most elements of w , and of the exceptional elements after applying Theorem 7.3, the key estimate leading to Lemma 7.6 is that

$$(\rho^{-1}/\sqrt{\alpha_n n'})^n \rho^{n_0} = O(n^{-\varepsilon n/5})$$

as long as $n^{-O(1)} \leq \rho \leq O(n^{-\varepsilon/2})$. We now present the details.

Proof. (of Lemma 7.6) Throughout the proof, we assume n is sufficiently large given ε . Suppose we have such a w . By Theorem 6.3 and $|\text{supp}(w)| > c_0 n$, as long as $\alpha_n \geq 4/(c_0 n)$, we have $\rho(w) \leq p^{-1} + 2/\sqrt{\alpha_n c_0 n} \leq (1 + 2c_0^{-1/2})n^{-\varepsilon/2}$, since $p \geq n^{\varepsilon/2}$ and $\alpha_n \geq n^{-1+\varepsilon}$.

Let Q be a symmetric GAP in $\mathbf{Z}/p\mathbf{Z}$ of rank at most r and volume at most V , such that for some subset $\tau \subset [n]$ of size n' we have for $j \in ([n] \setminus \tau)$ that $w_j \in Q$. Let R_1, \dots, R_n denote the rows of the matrix M formed by the columns $X_1/p, \dots, X_{n_0}/p$. For n sufficiently large (in terms of ε) such that $c_0 n \geq n'$, we see that the R_j for $j \in \tau$ must be linearly independent (or else there would be a normal vector to W_{n_0}/p with at most $c_0 n$ non-zero coefficients).

First, we will determine how many possible choices there are for the data of Q , τ , σ , and the w_j for $j \in \tau$, without any attempt to be sharp. Then, given those data, we will determine the probability that X_1, \dots, X_{n_0} could produce the situation outlined above with those data.

So we have at most p^r choices of generators for Q and at most V^r choices of dimensions (to obtain a lower rank GAP we just take some dimensions to be 0). There are at most 2^n choices of τ , and at most 2^n choices of σ . There are at most V^m choices of w_j for $j \in ([n] \setminus \tau)$.

Given Q , τ , σ , and the w_j for $j \in \tau$, we condition on the X_i for $i \in \sigma$. Then the τ entries of w are determined by the w_j for $j \in ([n] \setminus \tau)$ and the X_i for $i \in \sigma$ as follows. From $w \cdot X_i/p = 0$ for $i \in \sigma$, it follows that

$$\sum_{j \in \tau} w_j R_j|_{\sigma} = - \sum_{j \in ([n] \setminus \tau)} w_j R_j|_{\sigma}. \quad (22)$$

Since the $R_j|_{\sigma}$ (meaning row R_j restricted to the σ entries) for $j \in \tau$ are linearly independent and $|\tau| = |\sigma|$, we conclude that the w_j for $j \in ([n] \setminus \tau)$ and X_i for $i \in \sigma$ determine at most one possible choice for the w_j for $j \in \tau$.

For simplicity, we will proceed in two cases. First, we will determine how likely it is for W_{n_0}/p to have a normal vector w as in the lemma statement such that $\rho(w) \leq n^{-9}$. From the well-boundedness of Q , we have that $r \leq C_{\varepsilon}$ and $V \leq C_{\varepsilon}n^3$. Thus the total number of choices for Q , τ , σ , and the w_j for $j \in \tau$ is at most $p^{C_{\varepsilon}}(C_{\varepsilon}n^3)^{C_{\varepsilon}+m}4^n$. Once we condition on the X_i for $i \in \sigma$, the vector w is determined by our choices, and the probability that $w \cdot X_i/p = 0$ for $i \in ([n_0] \setminus \sigma)$ is at most $n^{-9(n_0-n')}$. Thus the total probability that W_{n_0}/p has a normal vector w as in the lemma statement such that $\rho(w) \leq n^{-9}$ is at most

$$p^{C_{\varepsilon}}(C_{\varepsilon}n^3)^{C_{\varepsilon}+m}4^n n^{-9(n_0-n')} = O_{\varepsilon}(p^{C_{\varepsilon}}n^{-n}).$$

Next we will determine how likely it is for W_{n_0}/p to have a normal vector w as in the lemma statement such that $\rho(w) > n^{-9}$. However, instead of counting the Q from the lemma statement, we are going to count the Q provided by Theorem 7.3. More specifically, we divide $[n^{-9}, (1+2c_0^{-1/2})n^{-\varepsilon/2}]$ into dyadic subintervals $I_{\ell} = [\rho_{\ell}, 2\rho_{\ell}]$ and we suppose that $\rho(w) \in I_{\ell}$. Let $\rho = \rho(w)$. We can apply Theorem 7.3 with $C = 3$ for $p \geq C_{\varepsilon}n^3$. Then there exists a symmetric GAP Q of rank $r \leq C_{\varepsilon}$ with $|Q| \leq \max((C_{\varepsilon}(\rho^{-1}/(\alpha_n n')^{r/2}), 1)$, and a subset $\tau \subset [n]$ of n' indices such that for $j \in ([n] \setminus \tau)$, we have $w_j \in Q$. Note that $r = 0$ would imply that $|\text{supp}(w)| \leq \lceil n^{\varepsilon/2}\alpha_n^{-1} \rceil$, which contradicts the fact that $|\text{supp}(w)| > c_0 n$. Also, since $\rho^{-1} \geq (1+2c_0^{-1/2})^{-1}n^{\varepsilon/2}$, we have that $C'_{\varepsilon}\rho^{-1}/(\alpha_n n')^{1/2} \geq 1$ for some constant $C'_{\varepsilon} \geq C_{\varepsilon}$ only depending on ε . So $r \geq 1$, and

$$|Q| = O_{\varepsilon}(\rho^{-1}/(\alpha_n n')^{1/2}). \quad (23)$$

Since ℓ was chosen so that $\rho(w) \leq 2\rho_{\ell}$, we have that the probability that $w \cdot X_i/p = 0$ for $i \in ([n_0] \setminus \sigma)$ is at most $(2\rho_{\ell})^{n_0-|\sigma|}$. Thus the total probability that there is a w as in the lemma statement such that $\rho(w) > n^{-9}$ is at most

$$\begin{aligned} \sum_{\ell=1}^{O(\log n)} p^{C_{\varepsilon}} O_{\varepsilon}(\rho_{\ell}^{-1}/(\alpha_n n')^{1/2})^{C_{\varepsilon}+m} 4^n (2\rho_{\ell})^{n_0-n'} &\leq \sum_{\ell=1}^{O(\log n)} p^{C_{\varepsilon}} e^{O_{\varepsilon}(n)} O_{\varepsilon}((\alpha_n n')^{-1/2})^m (\rho_{\ell}^{-1})^{n-n_0} \\ &\leq O_{\varepsilon}(p^{C_{\varepsilon}} n^{-\varepsilon n/5}). \end{aligned} \quad (24)$$

For these inequalities, we use facts including $\rho_{\ell}^{-1} \leq n^9$ and $n-n_0 = \lfloor \frac{3 \log n}{\alpha_n} \rfloor \leq 3n^{1-\varepsilon} \log n$, and $(\alpha_n n')^{-1/2} \leq n^{-\varepsilon/4}$, and $m = n - \lceil n^{\varepsilon/2}\alpha_n^{-1} \rceil \geq n - \lceil n^{1-\varepsilon/2} \rceil$. \square

As good as the bounds in Lemmas 7.5 and 7.6 are, they still cannot be summed over all primes p that might divide the determinant of $M_{n \times n}$. So at some point, we need to lift the structured normal vectors from characteristic p to characteristic 0. Unlike in Section 6.1, when we could lift non-sparse normal vectors for all large primes, our structured vectors here have more noise and we cannot lift until the primes are even larger. The following lemma does this lifting and is the only place we use that the coefficients of the X_i are bounded. Instead of counting structured vectors in characteristic 0 (or modulo a prime $> n^{n/2}$), for which we would need some bound on their coefficients (e.g., see the commensurability results [20, Lemma 9.1] and [30, Theorem 5.2(iii)]), we prove in the following lemma that we can also reduce structured vectors in characteristic 0 to structured vectors modulo a prime around $e^{n^{1-\varepsilon/3}}$. This allows us to transfer structured vectors modulo p for our largest range of p to structured vectors for a single prime p_0 that is of reasonably controlled size.

We say a submodule of \mathbf{Z}^n is *admissible* if it is generated by vectors with coordinates at most n^T in absolute value. (In particular, W_k is always admissible.)

Lemma 7.7 (Lifting and reducing structured vectors). *Let A be an admissible submodule of \mathbf{Z}^n , and p be a prime $\geq e^{n^{1-\varepsilon/3}}$, and n be sufficiently large given ε and T . Then A has a structured normal vector (for a GAP with integral generators) if and only if A/p has a structured normal vector.*

Proof. (of Lemma 7.7) We will first prove the “if” direction. Assume that the first $m = n - n'$ entries of the normal vector $w = (w_1, \dots, w_n)$ belong to a symmetric well-bounded GAP Q with r generators a_1, \dots, a_r in $\mathbf{Z}/p\mathbf{Z}$, and $w_j = \sum_{l=1}^r x_{jl}a_l$ for $1 \leq j \leq m$. Let M be the matrix with entries at most n^T in absolute value whose columns generate A . Let R_1, \dots, R_n be the rows of M . We have the equality modulo p

$$0 = \sum_{j=1}^m w_j R_j + \sum_{j=m+1}^n w_j R_j = \sum_{l=1}^r a_l \left(\sum_{j=1}^m x_{jl} R_j \right) + \sum_{j=m+1}^n w_j R_j.$$

Now for $1 \leq l \leq r$, let $Z_l := \sum_{j=1}^m x_{jl} R_j$. We have $|x_{jl}| \leq |Q| \leq C_\varepsilon n^3$. The entries of Z_l are then bounded by $C_\varepsilon n^{T+4}$, which is $\leq n^{T+5}$ for n sufficiently large given ε , while the entries of R_{m+1}, \dots, R_n are bounded by n^T . Let M' be the matrix whose columns are $Z_1, \dots, Z_r, R_{m+1}, \dots, R_n$. The above identity then implies that $(a_1, \dots, a_r, w_{m+1}, \dots, w_n)^t$ is in the kernel of M' . Lemma 6.5 (3) applied to M' , with $k = r + n'$ implies that as long as $p \geq e^{(k \log k)/2 + k(T+5) \log n}$ (which is satisfied because $p \geq e^{n^{1-\varepsilon/3}}$, and $r \leq C_\varepsilon$, and $n' \leq n^{1-\varepsilon/2} + 1$, and n is sufficiently large given ε and T), then there exist integers a'_l, w'_j , reducing mod p to a_l, w_j , for $1 \leq l \leq r$ and $m + 1 \leq j \leq n$, such that

$$\sum_{k=1}^r a'_k Z_k + \sum_{j=m+1}^n w'_j R_j = 0.$$

Let $w' = (w'_1, \dots, w'_n)$ where $w'_j = \sum_{l=1}^r x_{jl}a'_l$ for $1 \leq j \leq m$. By definition the w'_j for $1 \leq j \leq m$ belong to the symmetric GAP with generators a'_l and with the same rank and dimensions as Q , and w' is normal to A . Further w' is non-zero since it reduces to w mod p .

The “only if” direction appears easier at first—if we start with a structured normal vector, we can reduce the generators of the GAP and the normal vector mod p for any prime p . However, the difficulty is that for general primes p it is possible for the generators a_l of the GAP to be not all 0 mod p , but yet the resulting normal vector w to be 0 mod p . Given A , we choose w minimal (e.g. with $\sum_i |w_i|$ minimal) so that the first $m = n - n'$ entries (without loss of generality) of the normal vector $w = (w_1, \dots, w_n)$ to A belong to a symmetric well-bounded GAP Q with r generators a_1, \dots, a_r in \mathbf{Z} , and $w_j = \sum_{l=1}^r x_{jl}a_l$ for $1 \leq j \leq m$ and w is non-zero. Let M_x be the $n \times (r + n')$ matrix with entries x_{jl} in the first m rows and r columns, the $n' \times n'$ identity matrix in the last n' rows and columns, and zeroes elsewhere. So for $a := (a_1, \dots, a_r, w_{m+1}, \dots, w_n)^t$, we have $M_x a = w^t$.

Certainly by minimality of w at least some coordinate of w is not divisible by p (else we could divide the a_l and w_j all by p and produce a smaller structured normal w). Suppose, for the sake of contradiction that all of the coordinates of w are divisible by p . The entries of M_x are bounded by $C_\varepsilon n^3$, so, as above, for $p \geq e^{n^{1-\varepsilon/3}}$, by Lemma 6.5 (3) we have that $\ker M_x|_{\mathbf{Z}^{r+n'}}$ surjects onto $\ker M_x/p$. So a/p is in the kernel of M_x/p , and choose some lift $a' := (a'_1, \dots, a'_r, w'_{m+1}, \dots, w'_n)^t \in \mathbf{Z}^n$ of a/p in the kernel of M_x . Then $a - a' \in p\mathbf{Z}^n$, and $M_x(\frac{1}{p}(a - a')) = \frac{1}{p}w$. Note that $\frac{1}{p}w$ is non-zero integral normal vector to A , and the equality $M_x(\frac{1}{p}(a - a')) = \frac{1}{p}w$ shows that all but n' of the coordinates of $\frac{1}{p}w$ belong to a symmetric well-bounded GAP with integral generators and the same rank and volume as Q , contradicting the minimality of w . Thus we conclude that w/p is non-zero and thus a structured normal vector of A/p for GAP Q/p . \square

We now conclude the main result of this section.

Proof of Lemma 6.7. We let \mathcal{S}' be the set of submodules H of \mathbf{Z}^n such that for all primes $p > e^{d\alpha_n n}$, the vector space H/p has no structured normal vector w . We assume throughout the proof that n is sufficiently large given ε, T, d . First, we will bound $\mathbf{P}(W_{n_0} \notin \mathcal{S}')$. By Lemma 7.7, for $p \geq e^{n^{1-\varepsilon/4}}$, if W_{n_0}/p has a structured normal vector, then W_{n_0} has a structured normal vector, and then W_{n_0}/p' has a structured normal vector for every prime p' with $e^{n^{1-\varepsilon/3}} \leq p' < e^{n^{1-\varepsilon/4}}$ (of which there is at least 1).

So it suffices to bound the condition that W_{n_0}/p has a structured normal vector for p is a prime $C_\varepsilon n^3 \leq p < e^{n^{1-\varepsilon/4}}$.

We will include in our upper bound the probability that W_{n_0}/p has a non-zero normal vector w with $|\text{supp}(w)| \leq c_0 n$ for some prime $p < e^{n^{1-\varepsilon/4}}$, which is at most $e^{-c_1 \alpha_n n/2} + e^{n^{1-\varepsilon/4}} (2/3)^{n/4}$ by Lemmas 7.4 and 7.5. Then, otherwise, by Lemma 7.6, it is with probability at most $e^{n^{1-\varepsilon/4}} O_{\varepsilon, T}(e^{C_\varepsilon n^{1-\varepsilon/4}} n^{-\varepsilon n/5})$ that, for some prime $p < e^{n^{1-\varepsilon/4}}$, the space W_{n_0}/p has a structured normal vector w . We conclude that $\mathbf{P}(W_{n_0} \in \mathcal{S}') \geq 1 - e^{-c_2 \alpha_n n}$ for some absolute constant c_2 .

If $H \in \mathcal{S}'$ and H' is a proper subspace of $(\mathbf{Z}/p\mathbf{Z})^n$ containing H/p , then H' has some non-zero normal vector w (also normal to H/p). Let $p > e^{d\alpha_n n}$ be a prime. If $\rho(w) \leq n^{-3}$, then since $\mathbf{P}(X/p \in H') \leq \mathbf{P}(X/p \cdot w = 0)$ we have $\mathbf{P}(X/p \in H') \leq n^{-3}$. Otherwise, if $\rho(w) > n^{-3}$, we apply Theorem 7.3 with $C = 3$ and find a symmetric well-bounded GAP containing all but n' coordinates of w , which contradicts the definition of \mathcal{S}' . \square

8. LAPLACIAN OF RANDOM DIGRAPHS: PROOF OF THEOREM 1.6

As laid out in Section 2, it suffices to prove Proposition 2.5 and this task consists of three parts, in the first part we modify the method of Section 4 to justify Equation (12) for the small primes, in the second part we provide a complete proof for Equation (13) and (14) regarding the medium primes by improving the method of [19, 21], and in the last part we modify the method of Sections 6 and 7 to prove Equation (15) and (16) for the large primes.

For $1 \leq i \leq n$, we say that a random vector $X = (x_1, \dots, x_n) \in \mathbf{Z}_0^n$, the set of vectors of zero entry sum in \mathbf{Z}^n , has type \mathcal{T}_i if $x_i = -(x_1 + \dots + x_{i-1} + x_{i+1} + \dots + x_n)$ and $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ are i.i.d. copies of ξ_n from (1). Recall that $L_{M_{n \times n}}$ is a random matrix with independent columns X_i sampled from \mathcal{T}_i . Sometimes we will also denote this matrix by $L_{n \times n}$ for short.

I. Proof of Equation (12) of Proposition 2.5: treatment for small primes. In this subsection we modify the approach of Section 4 toward the Laplacian setting. We first prove the analog of Theorem 4.12 for the Laplacian. We will use the same approach as in [33, Theorem 6.2] to consider an auxiliary matrix that lets us carry the argument from the i.i.d. case to the Laplacian case. Let a be the exponent of G . Let $R = \mathbf{Z}/a\mathbf{Z}$ and $V = (\mathbf{Z}/a\mathbf{Z})^n$. We let M' be an $n \times n$ random matrix with coefficients in R with entries X_{ij} distributed as $(M_{n \times n})_{ij}$ for $i \neq j$ and with X_{ii} distributed uniformly in R , with all entries independent. Let $F_0 \in \text{Hom}(V, R)$ be the map that sends each standard basis element to 1. Now, M' and $L_{M_{n \times n}}$ do not have the same distribution, as the column sums of M' can be anything and the column sums of $L_{M_{n \times n}}$ are zero, i.e. $F_0 L_{M_{n \times n}} = 0$. However if we condition on $F_0 M' = 0$, then we find that this conditioned distribution of M' is the same as the distribution of $L_{M_{n \times n}}$. Given M' and conditioning on the off diagonal entries, we see that the probability that $F_0 M' = 0$ is a^{-n} (for any choice of off diagonal entries). So any choice of off diagonal entries is equally likely in $L_{M_{n \times n}}$ as in M' conditioned on $F_0 X = 0$.

So for $F \in \text{Hom}(V, G)$, we have

$$\mathbf{P}(F L_{M_{n \times n}} = 0) = \mathbf{P}(F M' = 0 | F_0 M' = 0) = \mathbf{P}(F M' = 0 \text{ and } F_0 M' = 0) a^n.$$

Let $\tilde{F} \in \text{Hom}(V, G \oplus R)$ be the sum of F and F_0 . Let $Z \subset V$ denote the vectors whose coordinates sum to 0, i.e.

$$Z = \{v \in V \mid F_0 v = 0\}.$$

Let $\text{Sur}^*(V, G)$ denote the maps from V to G that are a surjection when restricted to Z . We wish to estimate

$$\begin{aligned} \mathbf{E}(\# \text{Sur}(S_{M_{n \times n}}, G)) &= \mathbf{E}(\# \text{Sur}(Z / L_{M_{n \times n}} R^n, G)) \\ &= \sum_{F \in \text{Sur}(Z, G)} \mathbf{P}(F L_{M_{n \times n}} = 0) \\ &= \frac{1}{|G|} \sum_{F \in \text{Sur}^*(V, G)} \mathbf{P}(F L_{M_{n \times n}} = 0) \\ &= |G|^{-1} a^n \sum_{F \in \text{Sur}^*(V, G)} \mathbf{P}(\tilde{F} M' = 0). \end{aligned}$$

Note that if $F : V \rightarrow G$ is a surjection when restricted to Z , then \tilde{F} is a surjection from V to $G \oplus R$.

Now we need a slight variant on Lemma 4.10 to bound $F \in \text{Hom}(V, G)$ such that \tilde{F} is robust for a subgroup H of G .

Lemma 8.1 (Count of robust F for a subgroup H). *Let $\delta > 0$, and $a, n \geq 1$ be integers, and G be finite abelian group of exponent dividing a . Let H be a subgroup of $G \oplus R$ of index $D > 1$ and let $H = G_{\ell(D)} \subset \dots \subset G_2 \subset G_1 \subset G_0 = G \oplus R$ be a maximal chain of proper subgroups. Let $p_j = |G_{j-1}/G_j|$. For n sufficiently large given G , the number of $F \in \text{Hom}(V, G \oplus R)$ such that F composed with the projection onto R is 1 for each standard basis vector, and F is robust for H and for $1 \leq j \leq \ell(D)$, there are w_j elements i of $[n]$ such that $Fv_i \in G_{j-1} \setminus G_j$ is at most*

$$a^{-n} |H|^{n - \sum_j w_j} \prod_{j=1}^{\ell(D)} \binom{n}{w_j} |G_{j-1}|^{w_j}.$$

We note that for n sufficiently large in terms of G , the condition on the projection onto R implies that H surjects in the projection to R , and otherwise the proof of Lemma 8.1 is analogous to that of Lemma 4.10. (See also [33, Lemma 5.3].) We can then apply Lemma 4.11 as written to the maps \tilde{F} with range $G \oplus R$ and the matrix M' . The proof now follows the proof of Theorem 4.12, except that we are estimating

$$|G|^{-1} a^n \sum_{F \in \text{Sur}^*(V, G)} \mathbf{P}(\tilde{F} M' = 0).$$

The two sums of $|G|^{-n}$ over various F are replaced by sums of $|G|^{-n} a^{-n}$, but proofs of the same bounds can be found in the proof of [33, Theorem 6.2]. We deduce

$$|\mathbf{E}(\# \text{Sur}(S_{M_n \times n}, G)) - |G|^{-1}| \leq K_2 n^{-c_2},$$

and then deduce Equation (12) of Proposition 2.5, just as we proved Theorem 4.1 from Theorem 4.12.

II. Proof of Equations (13) and (14) of Proposition 2.5: treatment for the medium primes.

In this subsection we fix a prime p and will work with $\mathbf{Z}/p\mathbf{Z}$. As such, if not specified otherwise, all of the vectors and subspaces in this subsection are modulo p . For brevity, instead of X_i/p or W_i/p , we just write X_i or W_i . The co-dimensions (coranks) of subspaces, if not otherwise specified, are with respect to \mathbf{Z}_0^n/p . Although our main result, Theorem 8.2, works for any subspace W_{n-k} generated by $n-k$ columns of $L_{n \times n}$, for simplicity we assume $W_{n-k} = \langle X_1, \dots, X_{n-k} \rangle$. We show the following variant of Theorem 5.2.

Theorem 8.2. *There are sufficiently small constants $c, \eta > 0$ and sufficiently large constants $C_0, K > 0$ such that the following holds. Let p be a prime, and let $L_{n \times n}$ be a random matrix with independent columns X_i sampled from \mathcal{T}_i respectively, where we assume that*

$$\max_{r \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n = r) = 1 - \alpha_n \leq 1 - \frac{C_0 \log n}{n}. \quad (25)$$

Then for $1 \leq k \leq \eta n$ there exists an event \mathcal{E}_{n-k} on the σ -algebra generated by X_1, \dots, X_{n-k} , all of probability at least $1 - e^{-c\alpha_n n}$, such that for any k_0 with $k-1 \leq k_0 \leq \eta n$

$$|\mathbf{P}_{X_{n-k+1}}(X_{n-k+1} \in W_{n-k} | \mathcal{E}_{n-k} \wedge \text{codim}(W_{n-k}) = k_0) - p^{-k_0}| \leq K e^{-c\alpha_n n}.$$

Combining with Theorem 5.3 and with appropriate choices of c_0 and K_0 we then deduce the part of Proposition 2.5 for medium primes, analogous to the proof of Theorem 5.1.

Now we give a proof of Theorem 8.2. Our overall approach is similar to the proof of [21, Theorems A.1 and A.4] (which is built on approaches in [19, 30]), but for the Laplacian we cannot apply these results because the column vectors, as well as the entries in each column, are not identically distributed any more.

We would like to emphasize that in our argument below the positive constants $c, \beta, \delta, \eta, \lambda$ are sufficiently small and allowed to depend only on the constant C_0 in the bound (25) of α_n . We first introduce a version of Lemma 3.1 and Corollary 3.2 for α_n -dense random variables in the Laplacian setting.

Lemma 8.3. *For a deterministic subspace V of \mathbf{Z}_0^n/p (or \mathbf{Z}_0^n) of dimension d and for any i*

$$\mathbf{P}_{X \in \mathcal{T}_i}(X \in V) \leq (1 - \alpha_n)^{n-d-1}.$$

As a consequence, X_1, \dots, X_{n-k} are linearly independent in \mathbf{Z}_0^n/p with probability at least $1 - n(1 - \alpha_n)^{k-1}$.

Proof. It suffices to verify the first part. But for this we just project the vectors onto the coordinates of indices different from i , and then use Lemma 3.1. \square

We will also need the following variant of Theorem 6.3.

Theorem 8.4 (forward Erdős-Littlewood-Offord for the Laplacian). *Suppose that $w = (w_1, \dots, w_n) \in (\mathbf{Z}/p\mathbf{Z})^n$ does not have any component w_j with multiplicity larger than $n - m$, then for any i*

$$\sup_r |\mathbf{P}_{X \in \mathcal{T}_i}(X \cdot w = r) - \frac{1}{p}| \leq \frac{2}{\sqrt{\alpha_n m}}.$$

We remark that the classical Erdős-Littlewood-Offord in characteristic zero implies that if $w = (w_1, \dots, w_n) \in \mathbf{Z}^n$ does not have any component w_j with multiplicity larger than $n - m$, then for any i

$$\sup_{r \in \mathbf{Z}} |\mathbf{P}_{X \in \mathcal{T}_i}(X \cdot w = r)| \leq \frac{2}{\sqrt{\alpha_n m}}.$$

Proof. (of Theorem 8.4) Assume that $X = (x_1, \dots, x_n) \in \mathcal{T}_i$ for some $1 \leq i \leq n$. Then

$$x_1 w_1 + \dots + x_n w_n = x_1(w_1 - w_i) + \dots + x_{i-1}(w_{i-1} - w_i) + x_{i+1}(w_{i+1} - w_i) + \dots + x_n(w_n - w_i).$$

By the assumption, at least m entries $w_1 - w_i, \dots, w_n - w_i$ are non-zero. Because $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ are i.i.d., we then can apply Theorem 6.3. \square

8.5. Sparse subspace. Let $0 < \delta, \eta$ be small constants (independently from α_n). Given a vector space $H \leq (\mathbf{Z}/p\mathbf{Z})^n$, we call H δ -sparse if there is a non-zero vector w with $|\text{supp}(w)| \leq \delta n$ such that $w \perp H$.

Lemma 8.6 (random subspaces are not sparse, Laplacian case). *There exist absolute constant c' and C' such that the following holds with $\alpha_n \geq \frac{C_0 \log n}{n}$. Let $\varepsilon, \delta, \eta$ be constants such that $0 < \varepsilon < 1/12$ and $0 \leq \delta, \eta \leq \varepsilon$.*

- (characteristic p) For ξ_n satisfying Equation (25), and for $0 \leq k < \eta n$

$$\mathbf{P}_{X_1, \dots, X_{n-k}}(W_{n-k}/p \text{ is not } \delta\text{-sparse}) \geq 1 - e^{-c' \alpha_n n}.$$

- (characteristic zero) For ξ_n satisfying Equation (1), and for $0 \leq k < \eta n$

$$\mathbf{P}_{X_1, \dots, X_{n-k}}(W_{n-k} \text{ is not } \delta\text{-sparse in } \mathbf{Z}^n) \geq 1 - e^{-c' \alpha_n n}.$$

This result is actually a special case of Lemma 8.21 and 8.22, which will be discussed in due course. In connection to Theorem 8.4, it is more useful to connect the sparseness property to the one of having an entry of high multiplicity.

Claim 8.7. *Assume that the random subspace W_{n-k}/p does not accept any normal vector in $(\mathbf{Z}/p\mathbf{Z})^n$ of support size at most δ , then it does not accept any normal vector with an entry of multiplicity between $n - \delta n$ and $n - 1$ either. The same holds in the the characteristic zero case \mathbf{Z}^n .*

Proof. This is because of the invariance property that if $w = (w_1, \dots, w_n)$ is normal to W_{n-k} then so is any shifted vector $(w_1 - w_0, \dots, w_n - w_0)$ to W_{n-k} . \square

To conclude our treatment for the sparse case, given constants $\varepsilon, \eta, \delta$ and the parameter α_n from (25), let $\mathcal{E}_{k, \text{dense}} = \mathcal{E}_{k, \text{dense}}(\varepsilon, \eta, \delta)$ denote the event in the σ -algebra generated by X_1, \dots, X_{n-k} considered in Lemma 8.6, then

$$\mathbf{P}(\mathcal{E}_{k, \text{dense}}) \geq 1 - e^{-c' \alpha_n n}. \quad (26)$$

As such we can simply condition on this event without any significant loss. In our next move, we will choose $\lambda > 0$ to be a sufficiently small constant and show that it is highly unlikely that W_{n-k}/p is some non δ -sparse subspace (modulo V) of co-dimension k_0 with $k - 1 \leq k_0 \leq \eta n$ such that

$$e^{-\lambda \alpha_n n} < \max_{i, X \in \mathcal{T}_i} |\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}|.$$

Let us simply call V *bad* if this holds. For motivation, instead of bounding the probability that W_{n-k}/p is bad, let us simplify it to bounding the probability that X_1, \dots, X_{n-k} all belong to a bad subspace V . For this we will use the “swapping method” from [15, 30], and this was also adapted by Maples in [19] for the modulo p case. Roughly speaking, by letting the random variable ξ_n be lazier at zero, the associated

random vector Y with this lazy random variable will stick to V more often than X does (see Lemma 8.12), say $|\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}| \leq 0.51|\mathbf{P}(Y \in V) - \frac{1}{p^{k_0}}|$. Hence if $|\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}|$ is large enough, say larger than $\frac{16}{p^{k_0}}$, then we have that $\mathbf{P}(X \in V) \leq (2/3)\mathbf{P}(Y \in V)$, which in turn leads to a very useful bound $\mathbf{P}(X_1, \dots, X_{n-k} \in V) \leq (2/3)^{n-k}\mathbf{P}(Y_1, \dots, Y_{n-k} \in V)$. In what follows we will try to exploit this crucial exponential gain toward the Laplacian setting and toward the event that X_1, \dots, X_{n-k} actually span a bad subspace.

8.8. Semi-saturated subspace. Given $0 < \alpha, \delta, \lambda < 1$. We call a subspace $V \leq \mathbf{Z}_0^n/p$ of co-dimension $k_0 \leq \eta n$ (with respect to \mathbf{Z}_0^n/p) *semi-saturated* (or *semi-sat* for short) with respect to these parameters if V is not δ -sparse and

$$e^{-\lambda\alpha_n n} < \max_{i, X \in \mathcal{T}_i} |\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}| \leq \frac{16}{p^{k_0}}. \quad (27)$$

Here we assume

$$e^{-\lambda\alpha_n n} < \frac{16}{p^{k_0}}.$$

If this condition is not satisfied (such as when p is sufficiently large), then the semi-saturated case can be omitted. Our main result of this part can be viewed as a structural theorem which says that semi-saturated subspaces can be “captured” by a set of significantly fewer than p^n vectors.

Lemma 8.9. *For all $\beta > 0$ and $\delta > 0$ there exists $0 < \lambda = \lambda(\beta, \delta) < 1$ in the definition of semi-saturation and a deterministic set $\mathcal{R} \subset (\mathbf{Z}/p\mathbf{Z})^n$ of non δ -sparse vectors and of size $|\mathcal{R}| \leq (2\beta^\delta)^n p^n$ such that every semi-saturated V is normal to a vector $R \in \mathcal{R}$. In fact the conclusion holds for any subspace V satisfying the LHS of (27).*

Proof. (of Lemma 8.9) Without loss of generality, assume that $e^{-\lambda\alpha_n n} < |\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}|$ where $X \in \mathcal{T}_1$. Equivalently, with $J = \{2, \dots, n\}$

$$e^{-\lambda\alpha_n n} < |\mathbf{P}(X|_J \in V|_J) - \frac{1}{p^{k_0}}|.$$

By [19, Proposition 2.5] (see also [21, Lemma A.12]), there exists a deterministic set $\mathcal{R}' \subset (\mathbf{Z}/p\mathbf{Z})^{n-1}$ of non δ -sparse vectors and of size $|\mathcal{R}'| \leq (2\beta^\delta)^{n-1} p^{n-1}$ such that $V|_J$ is normal to a vector $R \in \mathcal{R}'$. We then define \mathcal{R} by appending a first coordinate to the vectors of \mathcal{R}' to make them have zero entry-sum. \square

Let $\mathcal{F}_{n-k, k_0, \text{semi-sat}}$ be the event that $\text{codim}(W_{n-k}) = k_0$ and W_{n-k} is semi-saturated.

Lemma 8.10 (random subspaces are not semi-saturated, Laplacian case). *Let $\beta, \delta > 0$ be parameters such that $\beta^\delta < 17^{-2}/2$. With $\lambda = \lambda(\beta, \delta)$ from Lemma 8.9 we have*

$$\mathbf{P}(\mathcal{F}_{n-k, k_0, \text{semi-sat}}) \leq e^{-n}.$$

In particular, with $\mathcal{E}_{n-k, \text{semi-sat}}$ denotes the event complements $\wedge_{k-1 \leq k_0 \leq \eta n} \overline{\mathcal{F}}_{n-k, k_0, \text{semi-sat}}$ in the σ -algebra generated by X_1, \dots, X_{n-k} , then

$$\mathbf{P}(\mathcal{E}_{n-k, \text{semi-sat}}) \geq 1 - e^{-n/2}. \quad (28)$$

Proof. (of Lemma 8.10) We have

$$\mathbf{P}(\mathcal{F}_{n-k, k_0, \text{semi-sat}}) = \sum_{V \text{ semi-sat}, \text{ codim}(V) = k_0} \mathbf{P}(W_{n-k} = V) \leq \sum_{V \text{ semi-sat}, \text{ codim}(V) = k_0} \mathbf{P}(X_1, \dots, X_{n-k} \in V).$$

Now for each fixed $V \leq \mathbf{Z}_0^n/p$ that is semi-saturated of co-dimension k_0 , by definition

$$\mathbf{P}(X_{n+k+j} \in V) \leq \max_{i, X \in \mathcal{T}_i} |\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}| + \frac{1}{p^{k_0}} \leq 17p^{-k_0}.$$

So

$$\mathbf{P}(X_1, \dots, X_{n-k} \in V) \leq 17^{n-k} p^{-k_0(n-k)}.$$

We next use Lemma 8.9 to count the number $N_{semi-sat}$ of semi-saturated subspaces V . Each V is determined by its annihilator V^\perp in \mathbf{Z}_0^n/p (of cardinality p^{k_0}). For V^\perp we can choose a first vector $v_1 \in \mathcal{R}$, and then $v_2, \dots, v_{k_0} \in \mathbf{Z}_0^n/p$ (linearly independently). By double counting, we obtain an upper bound

$$N_{semi-sat} = O\left((2\beta^\delta)^n p^n \frac{(p^{n-1})^{k_0-1}}{|V^\perp|^{k_0-1}}\right) = O\left((2\beta^\delta)^n p^{nk_0 - k_0^2 + k_0}\right).$$

Putting together,

$$\begin{aligned} \mathbf{P}(\mathcal{F}_{n-k, k_0, semi-sat}) &\leq \sum_{V \text{ semi-sat, codim}(V)=k_0} \mathbf{P}(X_1, \dots, X_{n-k} \in V) = O\left((2\beta^\delta)^n p^{nk_0 - k_0^2 + k_0} 17^{n-k} p^{-k_0(n-k)}\right) \\ &= O\left(17^{n-k} (2\beta^\delta)^n p^{k_0} p^{k_0(k-k_0)}\right) = O\left(17^{n-k} (2\beta^\delta)^n p^{2k_0}\right), \end{aligned}$$

where we noted that $k_0 \geq k-1$. Now recall that $e^{-\lambda\alpha_n n} \leq 16p^{-k_0}$, and so

$$\mathbf{P}(\mathcal{F}_{n-k, k_0, semi-sat}) = O(17^{n-k} (2\beta^\delta)^n p^{2k_0}) = O(17^{n+1-k} (2\beta^\delta)^n e^{2\lambda\alpha_n n}).$$

We then choose β so that $2\beta^\delta < 17^{-2}$ and with $\lambda < 1/2$ we have

$$\mathbf{P}(\mathcal{F}_{n-k, k_0, semi-sat}) \leq e^{-n}.$$

□

Having worked with subspaces V where $\max_{i, X \in \mathcal{T}_i} |\mathbf{P}(X \in V) - p^{-k_0}|$ are still small, we now turn to the remaining case to apply the swapping method.

8.11. Unsaturated subspace. Let V be a subspace of codimension k_0 in \mathbf{Z}_0^n/p for some $k-1 \leq k_0 \leq \eta n$. We say that V is *unsaturated* (or *unsat.* for short) if V is not δ -sparse and

$$\max(e^{-d\alpha n}, 16p^{-k_0}) < \max_{i, X \in \mathcal{T}_i} |\mathbf{P}(X \in V) - p^{-k_0}|.$$

In particular this implies that

$$\max_{i, X \in \mathcal{T}_i} \mathbf{P}(X \in V) \geq \max\{17p^{-k_0}, \frac{16}{17}e^{-d\alpha n}\}.$$

In this case, for each $1 \leq i \leq n$ we say that V has *type* i if

$$\mathbf{P}_{X \in \mathcal{T}_i}(X \in V) = \max_{1 \leq j \leq n, X \in \mathcal{T}_j} \mathbf{P}(X \in V).$$

By taking union bound, it suffices to work with unsaturated subspaces of type 1. So in what follows $X \in \mathcal{T}_1$. The following is from [19, Lemma 2.8] (see also [21, Lemma A.15]).

Lemma 8.12. *There is a α'_n -balanced probability distribution ν on $\mathbf{Z}/p\mathbf{Z}$ with $\alpha'_n = \alpha_n/64$ such that if $Y = (y_1, \dots, y_n) \in (\mathbf{Z}/p\mathbf{Z})^n$ is a random vector with i.i.d. coefficients y_2, \dots, y_n distributed according to ν and $y_1 = -(y_2 + \dots + y_n)$ then for any unsaturated proper subspace V in \mathbf{Z}_0^n/p*

$$|\mathbf{P}(X \in V) - \frac{1}{p^{k_0}}| \leq (\frac{1}{2} + o(1))|\mathbf{P}(Y \in V) - \frac{1}{p^{k_0}}|.$$

(To be more precise, [19, Lemma 2.8] and [21, Lemma A.15] stated for vectors of i.i.d. entries, but for Lemma 8.12 we just need to truncate the first coordinate from all vectors.) For short, we will say that the vector Y from Lemma 8.12 has type \mathcal{T}'_1 . It follows from the definition of unsaturation and from Lemma 8.12 that

$$\mathbf{P}_{X \in \mathcal{T}_1}(X \in V) \leq \frac{2}{3} \mathbf{P}_{Y \in \mathcal{T}'_1}(Y \in V).$$

Definition 8.13. Let V be a subspace in $(\mathbf{Z}/p\mathbf{Z})^n$. Let $d_{comb} \in \{1/n, \dots, n^2/n\}$. We say that V of type 1 has combinatorial codimension d_{comb} if

$$(1 - \alpha)^{d_{comb}} \leq \mathbf{P}_{X \in \mathcal{T}_1}(X \in V) \leq (1 - \alpha_n)^{d_{comb}-1/n}. \quad (29)$$

Now as we are in the unsaturated case, $\mathbf{P}(X \in V) \geq \frac{16}{17}e^{-\lambda\alpha_n n}$, and so

$$d_{\text{comb}} \leq 2\lambda n. \quad (30)$$

In what follows we will fix d_{comb} from the above range, noting that d is sufficiently small, and there are only $O(n^2)$ choices of d_{comb} .

Let be fixed any $0 < \delta_1 < \delta_2 < 1/3$ such that

$$16(\delta_2 - \delta_1)(1 + \log \frac{1}{\delta_2 - \delta_1}) < \delta_1. \quad (31)$$

Set

$$r = \lfloor \delta_1 n \rfloor \text{ and } s = n - k - \lfloor \delta_2 n \rfloor.$$

Let $Y_1, \dots, Y_r \in \mathcal{T}'_1$ be random vectors with entries distributed by ν obtained by Lemma 8.12, and $Z_1, \dots, Z_s \in \mathcal{T}_1$ bee i.i.d. copies of a type 1 vector generated by μ . Note that in what follows the subspaces V are of given combinatorial dimension d_{comb} as in Equation (29) and (30).

Lemma 8.14 (random subspaces are not unsaturated, Laplacian case).

$$\mathbf{P}(X_1, \dots, X_{n-k} \text{ span an unsat. } V \text{ of type 1 of dim. between } r+s \text{ and } n-k) \leq (3/2)^{-\delta_1 n/4}.$$

Note that the event considered here is significantly harder to control than the event discussed in the paragraph preceding Subsection 8.8. This is also the place where [19] treated incorrectly by relying on [19, Proposition 2.3] (although our situation here is more technical with vectors of dependent and extremely sparse entries.) To prove Lemma 8.14 we will actually show

Lemma 8.15. *Assume that V is any subspace of type 1 and of dimension between $r+s$ and $n-k$ and $d_{\text{comb}} \leq 2\lambda n$. Then we have*

$$\mathbf{P}(X_1, \dots, X_{n-k} \text{ span } V) \leq (3/2)^{-r/2} \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}(Y_1, \dots, Y_r, Z_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V),$$

where $(i_1, \dots, i_{n-k-r-s})$ ranges over all subsets of size $n-k-r-s$ of $\{1, \dots, n-k\}$.

To conclude Lemma 8.14 we just use $(3/2)^{-r/2} \binom{n-k}{r+s} \leq (3/2)^{-\delta_1 n/4}$ (basing on Equation (31)) and the fact that for each fixed $(i_1, \dots, i_{n-k-r-s})$

$$\sum_{V \leq (\mathbf{Z}/p\mathbf{Z})^n, \text{ type 1, codim}(V) \geq k} \mathbf{P}(Y_1, \dots, Y_r, Z_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V) \leq 1.$$

Proof. (of lemma 8.15) We use the swapping method from [19, 30]. First of all, by independence between X_i, Y_j, Z_l ,

$$\begin{aligned} & \mathbf{P}(X_1, \dots, X_{n-k} \text{ span } V) \times \mathbf{P}(Y_1, \dots, Y_r, Z_1, \dots, Z_s \text{ linearly independent in } V) \\ &= \mathbf{P}(X_1, \dots, X_{n-k} \text{ span } V \wedge Y_1, \dots, Y_r, Z_1, \dots, Z_s \text{ linearly independent in } V). \end{aligned} \quad (32)$$

Roughly speaking, the linear independence of Y_1, \dots, Z_s is to guarantee that we then can add a few other X_i to form a new linear span of V , and by this way we can free the other X_i from the role of spanning (see for instance Equation (34)). We next estimate $\mathbf{P}(Y_1, \dots, Y_r, Z_1, \dots, Z_s \text{ linearly independent in } V)$. By product rule,

$$\begin{aligned} & \mathbf{P}(Z_1, \dots, Z_s, Y_1, \dots, Y_r \text{ linearly independent in } V) \\ &= \mathbf{P}(Y_r \in V) \times \mathbf{P}(Y_{r-1} \in V, Y_{r-1} \notin \langle Y_r \rangle | Y_r \in V) \times \dots \times \mathbf{P}(Y_1 \in V, Y_1 \notin \langle Y_2, \dots, Y_r \rangle | Y_2, \dots, Y_r \text{ lin. in } V) \times \\ & \times \mathbf{P}(Z_s \in V, Z_s \notin \langle Y_1, \dots, Y_r \rangle | Y_1, \dots, Y_r \text{ lin. in } V) \times \dots \times \\ & \times \mathbf{P}(Z_1 \in V, Z_1 \notin \langle Z_2, \dots, Z_r, Y_1, \dots, Y_r \rangle | Z_2, \dots, Z_s, Y_1, \dots, Y_r \text{ lin. in } V). \end{aligned}$$

We first estimate the terms on Y_i . By Lemma 8.3

$$\mathbf{P}\left(Y_i \in V, Y_i \notin \langle Y_{i+1}, \dots, Y_r \rangle | Y_{i+1}, \dots, Y_r \text{ lin. in } V\right) \geq \mathbf{P}(Y_i \in V) - (1 - \alpha'_n)^{n-(r-i)-1}.$$

This then can be estimated from below by

$$\begin{aligned} \mathbf{P}(Y_i \in V) - (1 - \alpha'_n)^{n-(r-i)-1} &\geq \frac{3}{2} \mathbf{P}(X_i \in V) - (1 - \alpha'_n)^{n-(r-i)-1} \geq \frac{3}{2} (1 - \alpha)^{d_{comb}} - (1 - \alpha'_n)^{n-(r-i)-1} \\ &\geq \frac{3}{2} (1 - \alpha_n)^{d_{comb}} (1 - (1 - \alpha_n)^{n/256 - d_{comb}}), \end{aligned}$$

where we used that $\alpha'_n = \alpha_n/64$ and $n - r \geq (1 - \delta_1)n \geq n/2$. Similarly,

$$\begin{aligned} \mathbf{P}\left(Z_i \in V, Z_i \notin \langle Z_{i+1}, \dots, Z_s, Y_1, \dots, Y_r \rangle | Z_{i+1}, \dots, Z_s, Y_1, \dots, Y_r\right) &\geq \mathbf{P}(Z_i \in V) - (1 - \alpha')^{n-(r+s-i)-1} \\ &\geq (1 - \alpha_n)^{d_{comb}} - (1 - \alpha'_n)^{n-(r+s-i)-1} \geq (1 - \alpha_n)^{d_{comb}} - (1 - \alpha_n)^{n/256} \end{aligned}$$

where we used that $r + s = n - k - (\lfloor \delta_2 n \rfloor - \lfloor \delta_1 n \rfloor) \geq n/2$. Putting together

$$\begin{aligned} \mathbf{P}\left(Y_1, \dots, Z_s \text{ linearly independent in } V\right) &\geq (3/2)^r (1 - \alpha_n)^{(r+s)d_{comb}} \left(1 - (1 - \alpha_n)^{n/256 - d_{comb}}\right)^{r+s} \\ &\geq (3/2)^{r-1} (1 - \alpha_n)^{(r+s)d_{comb}}, \end{aligned} \tag{33}$$

where we used $d_{comb} \leq 2\lambda n$ and λ is sufficiently small.

Now we estimate the probability $\mathbf{P}(X_1, \dots, X_{n-k} \text{ span } V \wedge Y_1, \dots, Y_r, Z_1, \dots, Z_s \text{ linearly independent in } V)$. Since $Y_1, \dots, Y_r, Z_1, \dots, Z_s$ are linearly independent in V and X_{k+1}, \dots, X_n span V , there exist $n - k - r - s$ vectors $X^{(i_1)}, \dots, X^{(i_{n-k-r-s})}$, which together with $Y_1, \dots, Y_r, Z_1, \dots, Z_s$, span V , and the remaining vectors $X^{i_{n-k-r-s+1}}, \dots, X^{i_{n-k}}$ belong to V . Thus,

$$\begin{aligned} &\mathbf{P}\left(X_1, \dots, X_{n-k} \text{ span } V \wedge Y_1, \dots, Y_r, Z_1, \dots, Z_s \text{ linearly independent in } V\right) \\ &\leq \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}\left(Y_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V \wedge X^{i_{n-k-r-s+1}}, \dots, X^{i_{n-k}} \in V\right) \\ &\leq \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}\left(Y_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V\right) \mathbf{P}\left(X^{i_{n-k-r-s+1}}, \dots, X^{i_{n-k}} \in V\right) \\ &\leq \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}\left(Y_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V\right) (1 - \alpha)^{(r+s)(d_{comb} - 1/n)}, \end{aligned} \tag{34}$$

where in the last step we used the upper bound $(1 - \alpha)^{d_{comb} - 1/n}$ for each $\mathbf{P}(X^{(i)} \in V)$.

Putting (32), (33) and (34) together,

$$\begin{aligned} \mathbf{P}\left(X_1, \dots, X_{n-k} \text{ span } V\right) &= \frac{\mathbf{P}\left(X_1, \dots, X_{n-k} \text{ span } V \wedge Y_1, \dots, Z_s \text{ linearly independent in } V\right)}{\mathbf{P}\left(Y_1, \dots, Z_s \text{ linearly independent in } V\right)} \\ &\leq (3/2)^{-r+1} (1 - \alpha_n)^{-(r+s)d_{comb}} \times \\ &\quad \times \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}\left(Y_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V\right) (1 - \alpha_n)^{(r+s)(d_{comb} - 1/n)} \\ &\leq (3/2)^{-r/2} \sum_{(i_1, \dots, i_{n-k-r-s})} \mathbf{P}\left(Y_1, \dots, Z_s, X^{(i_1)}, \dots, X^{(i_{n-k-r-s})} \text{ span } V\right). \end{aligned}$$

□

Remark that $r + s = n - k - (\lfloor \delta_2 n \rfloor - \lfloor \delta_1 n \rfloor) < n - k - \eta n$ if η is sufficiently small. As a consequence, if we let $\mathcal{E}_{k,unsat}$ denote the complement of the event in Lemma 8.14 in the σ -algebra generated by X_1, \dots, X_{n-k} then

$$\mathbf{P}(\mathcal{E}_{n-k,unsat}) \geq 1 - (3/2)^{-\delta_1 n/4}. \tag{35}$$

We now conclude the proof of Theorem 8.2. Let $\mathcal{E}_{n-k,dense}, \mathcal{E}_{n-k,semi-sat}, \mathcal{E}_{n-k,unsat}$ be the events introduced in (26), (28), (35) and let \mathcal{E}_{n-k} be their intersection. If we choose $c \leq \min\{c', \lambda\}$ then by definition, on these events, if $\text{codim}(W_{n-k}) = k_0$ then for a random vector X of any type \mathcal{T}_i

$$|\mathbf{P}(X \in W_{n-k}) - \frac{1}{p^{k_0}}| \leq e^{-c\alpha_n n},$$

completing the proof.

Finally, we conclude this section with an interesting consequence of Theorem 8.2 in light of singularity bounds for random matrices from [1, 3, 15, 16, 27, 30].

Corollary 8.16 (Non-singularity of the Laplacian). *There exist absolute constants $c_0, K_0 > 0$ such that the following holds. Assume that the i.i.d. entries are distributed according to a random variable ξ_n taking integral values and such that*

$$\max_{x \in \mathbf{Z}} \mathbf{P}(\xi_n = x) = 1 - \alpha_n \leq 1 - \frac{C_0 \log n}{n}, \text{ for a sufficiently large constant } C_0.$$

Then with probability at least $1 - K_0 e^{-c_0 \alpha_n n}$ the matrix $L_{n \times (n-1)}$ of any $n-1$ columns of $L_{n \times n}$ has rank $n-1$ in \mathbf{R}^n .

Note that we do not require ξ_n to be bounded at all, and our sparseness is almost best possible.

Proof. (of Corollary 8.16) We assume $L_{n \times (n-1)}$ to be the matrix of the first $n-1$ columns. For primes p sufficiently large given n , we will show that $L_{n \times (n-1)}/p$ has rank $n-1$ with probability at least $1 - e^{-c\alpha_n n}$. Given n , we choose a p large enough that $\mathbf{P}(\xi_n \equiv x \pmod{p}) \leq 1 - C_0(\log n)/n$. By Lemma 8.3, it suffices to bound the probability that $L_{n \times (n-1)}/p$ has rank between $n - \eta n$ and $n - 2$. For this we can deduce from Theorem 8.2 that if $1 \leq k \leq \eta n$ for some sufficiently small η , then

$$\mathbf{P}(\text{rank}(L_{n \times (n-1)}/p) = (n-1) - k) = O\left(n^k(p^{-k^2} + e^{-c\alpha_n n})\right). \quad (36)$$

Indeed, the event $\text{rank}(L_{n \times (n-1)}/p) = (n-1) - k$ implies that there exist k column vectors X_{i_1}, \dots, X_{i_k} which belong to the subspace of \mathbf{Z}_0^n/p of dimension $(n-1) - k$ generated by the remaining column vectors X_i for $i \neq i_1, \dots, i_k$. With a loss of a factor of $\binom{n-1}{k}$ in probability, we assume that $\{i_1, \dots, i_k\} = \{n-k, \dots, n-1\}$. We then use Theorem 8.2 (with the vectors and subspaces below being mod p)

$$\begin{aligned} & \mathbf{P}(X_{n-k}, \dots, X_{n-1} \in W_{n-k-1} \wedge \text{codim}(W_{n-k-1}) = k) \\ &= \mathbf{P}(X_{n-k}, \dots, X_{n-1} \in W_{n-k-1} \wedge \mathcal{E}_{n-k-1} \wedge \text{codim}(W_{n-k-1}) = k) + O(e^{-c\alpha_n n}) \\ &\leq \mathbf{P}(X_{n-k}, \dots, X_{n-1} \in W_{n-k-1} | \mathcal{E}_{n-k-1} \wedge \text{codim}(W_{n-k-1}) = k) + O(e^{-c\alpha_n n}) \\ &\leq \left(p^{-k} + O(e^{-c\alpha_n n})\right)^k + O(e^{-c\alpha_n n}) = O(p^{-k^2} + e^{-c\alpha_n n}), \end{aligned}$$

proving (36), and hence the corollary by taking p large enough, given n . \square

III. Proof of Equations (15) and (16) of Proposition 2.5: treatment for large primes. Now we modify the approach of Section 6 and 7 to the Laplacian setting. Let $d > 0$ be a constant and

$$\mathfrak{P}_n = \left\{ p \text{ prime}, p \geq e^{d\alpha_n n} \right\}.$$

Let $\mathcal{E}_{\neq 0}^{(L)}$ be the event that $L_{n \times (n-1)}$ has rank $n-1$ in \mathbf{R}^n . It follows from Corollary 8.16 that

$$\mathbf{P}(\mathcal{E}_{\neq 0}^{(L)}) \geq 1 - K_0 e^{-c_0 \alpha_n n}.$$

Our strategy is similar to the proof of Proposition 2.3. Recall that W_k is the submodule of \mathbf{Z}_0^n spanned by X_1, \dots, X_k . Let \mathfrak{W}_k be the set of primes $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-1$. Let \mathcal{C}_k be the event that $|\mathfrak{W}_k| \leq (2T+1) \log n / (2d\alpha_n)$ (the watch list is not too big). Note that any $p \in \mathfrak{W}_k$ for $k \leq n$ must divide $\det(L_{n \times (n-1)})$. By Hadamard's bound, $|\det(L_{n \times (n-1)})| \leq n^{n/2} n^{Tn}$, and so in particular, when \mathcal{C}_k occurs

then $\det(L_{n \times (n-1)}) = 0$. Let \mathcal{D}_k be the event that there is a $p \in \mathfrak{W}_k$ such that $\text{rank}(W_k/p) \leq k-2$, this is the event we want to avoid for all p .

We will show that $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \mathcal{C}_k \wedge \bar{\mathcal{D}}_k)$, and hence $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k)$, are large. The goal is to conclude that $\mathbf{P}(\bar{\mathcal{C}}_n \vee \bar{\mathcal{D}}_n)$ is large, and since we know that $\mathbf{P}(\bar{\mathcal{C}}_n)$ is small, we can conclude that $\mathbf{P}(\bar{\mathcal{D}}_n)$ is large, as desired. Now to estimate $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \mathcal{C}_k \wedge \bar{\mathcal{D}}_k)$ we will condition on the exact values of X_1, \dots, X_k where $\mathcal{C}_k \wedge \bar{\mathcal{D}}_k$ holds, and so there are at most $(2T+1) \log n / (2\lambda\alpha_n)$ primes $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-1$ and no prime $p \in \mathfrak{P}_n$ such that $\text{rank}(W_k/p) \leq k-2$. In this case $\bar{\mathcal{D}}_{k+1}$, as long as for each $p \in \mathfrak{W}_k$, we have $X_{k+1}/p \notin W_k/p$. Consider one prime $p \in \mathfrak{W}_k$, and let V be the value of W_k/p that the conditioned X_1, \dots, X_k give. From Lemma 8.3, $\mathbf{P}(X_{k+1}/p \in V) \leq (1 - \alpha_n)^{(n-1)-(k-1)} = (1 - \alpha)^{n-k}$. Thus,

$$\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \mathcal{C}_k \wedge \bar{\mathcal{D}}_k) \geq 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) (1 - \alpha_n)^{n-k}.$$

In particular, we have the same lower bound for $\mathbf{P}(\bar{\mathcal{C}}_{k+1} \vee \bar{\mathcal{D}}_{k+1} | \bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k)$, and then inductively, we have

$$\mathbf{P}(\bar{\mathcal{C}}_k \vee \bar{\mathcal{D}}_k) \geq 1 - \sum_{i=1}^{k-1} \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) (1 - \alpha_n)^{n-i} = 1 - \left(\frac{(2T+1) \log n}{2d\alpha_n} \right) \frac{(1 - \alpha_n)^{n-k+1}}{\alpha_n}.$$

Set $n_0 := n - \lfloor 3 \log n / \alpha_n \rfloor$. Then by using $\alpha_n \geq n^{-1+\varepsilon}$ we have that

$$\mathbf{P}(\bar{\mathcal{C}}_{n_0} \vee \bar{\mathcal{D}}_{n_0}) \geq 1 - O_{d,T}(n^{-1/2}).$$

We then have the following analog of Lemma 6.2.

Lemma 8.17. *Suppose that $\alpha_n \geq 6 \log n / n$. Then there is a set of submodules \mathcal{S}_L of \mathbf{Z}_0^n such that*

$$\mathbf{P}(W_{n_0} \in \mathcal{S}_L) \geq 1 - e^{-\alpha_n n / 8}$$

and for any prime $p \geq e^{d\alpha_n n}$, and any submodule $H \in \mathcal{S}_L$, for any proper subspace H' of \mathbf{Z}_0^n / p containing H/p ,

$$\mathbf{P}(X/p \in H') = O_{d,T} \left(\frac{\sqrt{\log n}}{\alpha_n \sqrt{n}} \right),$$

where X is any column of $L_{n \times n}$.

Lemma 8.17 can be shown exactly the same way Lemma 6.2 was deduced. Indeed, we can use Lemma 6.5 to lift the existence of sparse normal vector on any modulo p with $p \geq e^{d\alpha_n n}$ to the existence of sparse normal vector on characteristic zero, for which we then can use Lemma 8.6 (or Lemmas 8.21 and 8.22) to show that this event is unlikely. We then apply Theorem 8.4 (combined with Claim 8.7) to get the desired probability bound when the normal vectors are non-sparse.

Now similarly to the iid case, Lemma 8.17 allows us to justify Equations (15) and (16) only for $\alpha_n \geq n^{-1/6+\varepsilon}$. To extend to $\alpha_n \geq n^{-1+\varepsilon}$, we will have to need the following analog of Lemma 6.7

Lemma 8.18. *There is an absolute constant $c_2 > 0$ such that the following holds. Suppose that $\alpha_n \geq n^{-1+\varepsilon}$. There is a set \mathcal{S}'_L of submodules of \mathbf{Z}_0^n , such that*

$$\mathbf{P}(W_{n_0} \in \mathcal{S}'_L) \geq 1 - e^{-c_2 \alpha_n n}$$

and for n sufficiently large given d, ε, T , any prime $p \geq e^{d\alpha_n n}$, any submodule $H \in \mathcal{S}'_L$, and any proper subspace H' of $(\mathbf{Z}_0/p)^n$ containing H/p ,

$$\max_i \mathbf{P}_{X \in \mathcal{T}_i} (X/p \in H') \leq n^{-3}.$$

The deduction of Equations (15) and (16) from this lemma is similar to how Proposition 2.2 was deduced from Lemma 6.7. It remains to verify Lemma 8.18. For this we will make use of Theorem 8.4 and the following corollary of Theorem 7.3 for vectors from \mathcal{T}_i .

Theorem 8.19 (inverse Erdős-Littlewood-Offord for the Laplacian). *Let $\varepsilon < 1$ and C be positive constants. Assume that p is a prime that is larger than $C'n^C$ for a sufficiently large constant C' depending on ε and*

C. Let ξ be a random variable taking values in $\mathbf{Z}/p\mathbf{Z}$ which is α_n -balanced with $a_n \geq n^{-1+\varepsilon}$. Assume $w = (w_1, \dots, w_n) \in (\mathbf{Z}/p\mathbf{Z})^n$ such that

$$\rho(w) = \max_{i, (\xi_1, \dots, \xi_n) \in \mathcal{T}_i} \sup_{a \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_1 w_1 + \dots + \xi_n w_n = a) \geq n^{-C},$$

Then for any $n^{\varepsilon/2} \alpha^{-1} \leq n' \leq n$, there exists $1 \leq i \leq n$ and there exists a proper symmetric GAP Q in $\mathbf{Z}/p\mathbf{Z}$ of rank $r = O_{\varepsilon, C}(1)$ which contains all but n' elements of $\{w_1 - w_i, \dots, w_n - w_i\}$ (counting multiplicity), where

$$|Q| \leq \max \left\{ 1, O_{C, \varepsilon}(\rho(w)^{-1}/(\alpha n')^{r/2}) \right\}.$$

8.20. Proof of Lemma 8.18. Our method is similar to Section 7, so we will be brief. First we need an analog of Lemma 7.4 to estimate the probability that for some large prime p the module W_{n_0}/p of \mathbf{Z}_0^n/p accepts an extremely sparse normal vector.

Lemma 8.21. *There are absolute constants c_1, C_0 such that the following holds. Let $\beta_n := 1 - \max_{x \in \mathbf{Z}} \mathbf{P}(\xi_n = x)$, and assume $\beta_n \geq C_0 \log n/n$ and $\alpha_n \geq 6 \log n/n$. For $n \geq 2$, the following happens with probability at most $e^{-c_1 \beta_n n^{1/2}}$: for some prime $p > 2n^T$, the subspace W_{n_0}/p has a non-zero normal vector with at most $144\beta_n^{-1}$ non-zero coordinates.*

We also need an analog of Lemma 7.5 which will allow us to control the event that for some prime p of order $e^{o(n)}$ or $p = 0$ the subspace W_{n_0}/p accepts a normal vector of $o(n)$ non-zero entries.

Lemma 8.22. *There exist absolute constants c_0, C_0 such that the following holds. Let $\beta_n := 1 - \max_{x \in \mathbf{Z}} \mathbf{P}(\xi_n = x)$. Assume $\alpha_n \geq \frac{C_0 \log n}{n}$ and let p be a prime $> 2n^T$. The following happens with probability at most $(2/3)^{n/4}$: the subspace W_{n_0}/p has a non-zero normal vector w with $144\beta_n^{-1} \leq |\text{supp}(w)| \leq c_0 n$.*

Lemmas 8.21 and 8.22 will be verified in Appendix D by following the proofs of Lemmas 7.4 and 7.5.

We next discuss an analog of Lemma 7.6 on the existence of structured but not sparse normal vectors of W_{n_0}/p . Similarly to Section 7, we let $n' = \lceil n^{\varepsilon/2} \alpha_n^{-1} \rceil$ and $m = n - n'$, and we will apply Theorem 8.19 with this choice of n' and $C = 3$. By replacing $w = (w_1, \dots, w_n)$ by $(w_1 - w_i, \dots, w_n - w_i)$ if needed (note that this shifted vector is again a normal vector of W_{n_0}/p because this subspace consists of vectors of zero entry sum modulo p), we can simply say that Theorem 8.19 implies structure for w_1, \dots, w_n . We call a GAP Q well-bounded if it is of rank $\leq C_\varepsilon$ and $|Q| \leq C_\varepsilon n^3$, where C_ε is the maximum of C' and the constants from the $O_{C, \varepsilon}$ bounding the rank of volume of $|Q|$. Motivated by this, and similarly to Section 7, we call a vector w structured if it is non-zero, and there exists a symmetric well-bounded GAP Q such that all but n' coordinates of w belong to Q .

Lemma 8.23. *Let $\alpha_n \geq n^{-1+\varepsilon}$. Let p be a prime with $p \geq C_\varepsilon n^3$. The following event happens with probability $O_\varepsilon(p^{C_\varepsilon n^{-\varepsilon n/5}})$: the space W_{n_0}/p has a structured normal vector w , and W_{n_0}/p does not have a non-zero normal vector w' such that $|\text{supp}(w')| \leq c_0 n$ with c_0 from Lemma 8.22.*

Lemma 8.18 then can be shown by combining Lemmas 8.21, 8.22 and 7.7 the way Lemma 6.7 was concluded in the end of Section 7. Finally, the proof of Lemma 8.23 is almost identical to that of Lemma 7.6, the only difference is that we need to apply Theorems 8.4 and 8.19 instead of Theorems 6.3 and 7.3 in the argument leading to Equations (23) and (24), and thus we again omit the detail.

ACKNOWLEDGEMENTS

We would like to thank Wesley Pegden and Philip Matchett Wood for helpful conversations. We are also grateful to Nathan Kaplan, Lionel Levine and Philip Matchett Wood for valuable comments on an earlier version of this manuscript. Finally, we thank the anonymous referees for their careful reading of the paper and highly valuable suggestions and corrections. The first author is partially supported by National Science Foundation grants DMS-1600782 and NSF CAREER DMS-1752345. The second author is partially supported by a Packard Fellowship for Science and Engineering, a Sloan Research Fellowship, National Science Foundation grants DMS-1301690 and DMS-1652116, an NSF Waterman award, and a Vilas Early Career Investigator Award.

APPENDIX A. INEQUALITY LEMMAS

We have a straightforward inequality (easily checked by considering the cases when $A \leq 1$ and $A > 1$).

Lemma A.1. *Let $0 < d_1 < 1$ and $A > 0$, and $n \geq 1$ be an integer. Then we have*

$$(d_1 + (1 - d_1) \exp(-A))^n \leq \max(\exp(-(1 - d_1)An/2), ((1 + d_1)/2)^n).$$

The following is a standard estimate with binomial coefficients.

Lemma A.2. *Let $D_1 > 0$. For every $f > 0$, for all $\delta > 0$ sufficiently small (given D_1, f), we have that for all sufficiently large n (given D_1, f, δ), that*

$$\sum_{k=0}^{\lfloor \delta n \rfloor} \binom{n}{k} D_1^k \leq \exp(fn).$$

We put these lemmas together to obtain the inequality below.

Lemma A.3. *Let $D_1 > 0$ and $0 < d_1 < 1$. Then for positive γ sufficiently small (given D_1 and d_1), the following holds. Let $\Delta' > 2/(1 - d_1)$. For n sufficiently large (given $D_1, d_1, \gamma, \Delta'$) we have*

$$\sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k (d_1 + (1 - d_1) \exp(-(\Delta' \log n/n)k))^n \leq 3n^{-((1 - d_1)\Delta'/2 - 1)}.$$

Proof. For $n \geq 2$, we have that $(\Delta' \log n/n)k > 0$. So by Lemma A.1 with $A = (\Delta' \log n/n)k$, we have

$$\begin{aligned} & \sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k (d_1 + (1 - d_1) \exp(-(\Delta' \log n/n)k))^n \\ & \leq \sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k \exp(-(1 - d_1)An/2) + \sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k ((1 + d_1)/2)^n \\ & = \sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k \exp(-(1 - d_1)\Delta'(\log n)k/2) + \sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k ((1 + d_1)/2)^n. \end{aligned}$$

We have

$$r \cdot \binom{n}{k} D_1^k \exp(-(1 - d_1)\Delta'(\log n)k/2) \geq \binom{n}{k+1} D_1^{k+1} \exp(-(1 - d_1)\Delta' \log n(k+1)/2)$$

if and only if

$$r \geq \frac{n - k}{k + 1} \frac{D_1}{n^{(1 - d_1)\Delta'/2}}.$$

If $(1 - d_1)\Delta'/2 > 1$, then for n sufficiently large given D_1, d_1 and Δ' , we have

$$\frac{n - k}{k + 1} \frac{D_1}{n^{(1 - d_1)\Delta'/2}} \leq \frac{D_1 n}{2n^{(1 - d_1)\Delta'/2}} = \frac{D_1}{2n^{(1 - d_1)\Delta'/2 - 1}} \leq \frac{1}{2}.$$

So, if $(1 - d_1)\Delta'/2 > 1$, then for n sufficiently large given D_1, d_1 and Δ' , we have

$$\frac{1}{2} \cdot \binom{n}{k} D_1^k \exp(-(1 - d_1)\Delta'(\log n)k/2) \geq \binom{n}{k+1} D_1^{k+1} \exp(-(1 - d_1)\Delta' \log n(k+1)/2).$$

In particular, that implies

$$\sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k \exp(-(1 - d_1)\Delta'(\log n)k/2) \leq 2n^{-((1 - d_1)\Delta'/2 - 1)}.$$

For the second sum, let $f = -\log(1 + d_1/2)/2 > 0$ and note that $f + \log((1 + d_1/2)) < 0$. Then by Lemma A.2, we have that for γ sufficiently small given D_1, d_1 that for all n sufficiently large (given D_1, d_1, γ)

$$\sum_{k=1}^{\lfloor \gamma n \rfloor} \binom{n}{k} D_1^k ((1 + d_1)/2)^n \leq \exp(fn + \log(1 + d_1/2)n) = \exp(\log(1 + d_1/2)n/2).$$

For n sufficiently large given d_1, Δ' , we have

$$\exp(\log(1 + d_1/2)n/2) \leq n^{((1-d_1)\Delta'/2-1)}.$$

The lemma follows. \square

APPENDIX B. APPROXIMATE TRANSITION PROBABILITIES: PROOF OF THEOREM 5.3

First, we prove the following simplified version of the theorem. Using a coupling, we show that if the transition probabilities of two sequences of random variables are close, then the distribution of the random variables must be close.

Lemma B.1. *Let w and w' be sequences of random variables with $w_0 = w'_0 = 0$, for each $i \geq 0$*

$$\begin{aligned} \mathbf{P}(w'_{i+1} = a | w'_i = b) &= \mathbf{P}(w_{i+1} = a | w_i = b) + \delta(i, b, a) \\ \text{for all } a \text{ and } b \text{ such that } \mathbf{P}(w'_i = b) &= \mathbf{P}(w_i = b) \neq 0, \end{aligned}$$

and w_i and w'_i only take on countably many values. Then for any $n \geq 0$ and any set A of values taken by w_n or w'_n , we have

$$|\mathbf{P}(w_n \in A) - \mathbf{P}(w'_n \in A)| \leq \frac{1}{2} \sum_{0 \leq i \leq n-1} \sum_b \sum_c |\delta(i, b, c)| \mathbf{P}(w_i = b) \leq \frac{1}{2} \sum_{0 \leq i \leq n-1} \max_b \sum_c |\delta(i, b, c)|,$$

where b is summed over $\{b \mid \mathbf{P}(w_k = b) \neq 0 \text{ and } \mathbf{P}(w'_k = b) \neq 0\}$ and c is summed over $\{c \mid \mathbf{P}(w_{i+1} = c) \neq 0 \text{ or } \mathbf{P}(w'_{i+1} = c) \neq 0\}$.

Proof. Let S_i be the set of values taken on by w_i and w'_i . Let μ be Lebesgue measure on the interval $[0, 1]$. For each $i \geq 0$ and $b \in S_i$, we can choose measurable functions $\phi_{i,b} : [0, 1] \rightarrow S_{i+1}$ and $\phi'_{i,b} : [0, 1] \rightarrow S_{i+1}$ such that for all $c \in S_i$,

$$\mathbf{P}(w_{i+1} = c | w_i = b) = \mu(\phi_{i,b}^{-1}(c)) \quad \text{and} \quad \mathbf{P}(w'_{i+1} = c | w'_i = b) = \mu((\phi')_{i,b}^{-1}(c)),$$

and $\mu(\{x \in [0, 1] | \phi_{i,b}(x) \neq \phi'_{i,b}(x)\}) = \frac{1}{2} \sum_{c \in S_{i+1}} |\delta(i, b, c)|$. (If b isn't a value taken by one of the variables, we will just take $\phi_{i,b} = \phi'_{i,b}$) Then we construct Markov chains x_i and x'_i , with $x_0 = x'_0 = 0$, and to determine x_{i+1} and x'_{i+1} , we pick a random $x \in [0, 1]$, and then let $x_{i+1} = \phi_{i,x_i}(x)$ and $x'_{i+1} = \phi'_{i,x_i}(x)$. Note that for all $i \geq 0$ and all $b \in S_i$ and $c \in S_{i+1}$, we have

$$\mathbf{P}(x_{i+1} = c | x_i = b) = \mathbf{P}(w_{i+1} = c | w_i = b) \quad \text{and} \quad \mathbf{P}(x'_{i+1} = c | x'_i = b) = \mathbf{P}(w'_{i+1} = c | w'_i = b).$$

Thus, for all $n \geq 0$, we have $\mathbf{P}(x_i = a) = \mathbf{P}(w_i = a)$ and $\mathbf{P}(x'_i = a) = \mathbf{P}(w'_i = a)$. Note that x_n and x'_n are equal, unless for some $0 \leq i \leq n-1$, we have that $x_i = x'_i$, but $x_{i+1} \neq x'_{i+1}$, and in particular, $\phi_{i,x_i}(x) \neq \phi'_{i,x_i}(x)$. To see how likely this is for a given i , we sum over all $b \in S_i$, and have

$$\begin{aligned} \mathbf{P}(x_i = x'_i \wedge x_{i+1} \neq x'_{i+1}) &\leq \sum_{b \in S_i} \mathbf{P}(x_i = b) \mathbf{P}(x_i = x'_i = b \wedge x_{i+1} \neq x'_{i+1} | x_i = b) \\ &\leq \sum_{b \in S_i} \mathbf{P}(x_i = b) \mathbf{P}(\phi_{i,b}(x) \neq \phi'_{i,b}(x) | x_i = b) \\ &\leq \sum_{b \in S_i} \mathbf{P}(x_i = b) \frac{1}{2} \sum_{c \in S_{i+1}} |\delta(i, b, c)|. \end{aligned}$$

So

$$\mathbf{P}(x_n \neq x'_n) \leq \frac{1}{2} \sum_{i=0}^{n-1} \sum_{b \in S_i} \sum_{c \in S_{i+1}} \mathbf{P}(x_i = b) |\delta(i, b, c)|,$$

and from this the result follows, since

$$|\mathbf{P}(w_n \in A) - \mathbf{P}(w'_n \in A)| = |\mathbf{P}(x_n \in A) - \mathbf{P}(x'_n \in A)| \leq \max(\mathbf{P}(x_n \in A \wedge x'_n \notin A), \mathbf{P}(x_n \notin A \wedge x'_n \in A)).$$

\square

Theorem 5.3 will follow by applying Lemma B.1 to a modified sequence.

Proof of Theorem 5.3. We insert half-steps $x_{i+1/2} = (x_i, g_i)$ and $y_{i+1/2} = (y_i, 1)$. We compare the transitional probabilities as follows, first for i integral:

$$\mathbf{P}(y_{i+1/2} = (r, 1)|y_i = r) - \mathbf{P}(x_{i+1/2} = (r, 1)|x_i = r) = \mathbf{P}(g_i \neq 1|x_i = r)$$

and

$$\mathbf{P}(y_{i+1/2} = (r, 0)|y_i = r) - \mathbf{P}(x_{i+1/2} = (r, 1)|x_i = r) = -\mathbf{P}(g_i \neq 1|x_i = r).$$

Also for i integral, we have

$$\mathbf{P}(y_{i+1} = s|y_{i+1/2} = (r, 1)) - \mathbf{P}(x_{i+1} = s|x_{i+1/2} = (r, 1)) = \delta(i, r, s).$$

Then applying Lemma B.1, we have

$$\begin{aligned} & |\mathbf{P}(x_n \in A) - \mathbf{P}(y_n \in A)| \\ & \leq \frac{1}{2} \sum_{i=0}^{n-1} \sum_r 2\mathbf{P}(g_i \neq 1|x_i = r)\mathbf{P}(x_i = r) + \frac{1}{2} \sum_{i=0}^{n-1} \sum_r \sum_s |\delta(i, r, s)|\mathbf{P}(x_i = r) \\ & = \sum_{i=0}^{n-1} \mathbf{P}(g_i \neq 1) + \frac{1}{2} \sum_{i=0}^{n-1} \sum_r \sum_s |\delta(i, r, s)|\mathbf{P}(x_i = r). \end{aligned}$$

□

APPENDIX C. INVERSE THEOREM: PROOF OF THEOREM 7.3

We first introduce a more general structure in finite additive groups.

Definition C.1. A set P in a given finite additive group G is a *coset progression* of rank r if it can be expressed as in the form of

$$H + Q,$$

where H is a finite subgroup of G , and $Q = \{a_0 + x_1 a_1 + \dots + x_r a_r | M_i \leq x_i \leq M'_i \text{ and } x_i \in \mathbf{Z} \text{ for all } 1 \leq i \leq r\}$ is a GAP of rank r .

- We say that P with this presentation (i.e. choice of H , a_i , M_i , M'_i) is *proper* if the sums $h + a_0 + x_1 a_1 + \dots + x_r a_r, h \in H, M_i \leq x_i \leq M'_i$ are all distinct.
- More generally, given a positive integer t we say that P is t -proper with this presentation if $H + tQ$ is proper.
- If $-M_i = M'_i$ for all $i \geq 1$ and $a_0 = 0$, then we say that P with this presentation is *symmetric*.

To prove Theorem 7.3 we will make use of two results from [31] by Tao and Vu. The first result allows one to pass from coset progressions to proper coset progressions without any substantial loss.

Theorem C.2. [31, Corollary 1.18] *There exists a positive integer C_1 such that the following statement holds. Let Q be a symmetric coset progression of rank $d \geq 0$ and let $t \geq 1$ be an integer. Then there exists a t -proper symmetric coset progression P of rank at most d such that we have*

$$Q \subset P \subset Q_{(C_1 d)^{3d/2} t}.$$

We also have the size bound

$$|Q| \leq |P| \leq t^d (C_1 d)^{3d^2/2} |Q|.$$

The second result, which is directly relevant to us, says that as long as $|kX|$ grows slowly compared to $|X|$, then it can be contained in a structure. This is a long-ranged version of the Freiman-Ruzsa theorem.

Theorem C.3. [31, Theorem 1.21] *There exists a positive integer C_2 such that the following statement holds: whenever $d, k \geq 1$ and $X \subset G$ is a non-empty finite set such that*

$$k^d |X| \geq 2^{2^{C_2 d^2 2^{6d}}} |kX|,$$

then there exists a proper symmetric coset progression $H + Q$ of rank $0 \leq d' \leq d - 1$ and size $|H + Q| \geq 2^{-2^{C_2 d^2 2^{6d}}} k^{d'} |X|$ and $x, x' \in G$ such that

$$x + (H + Q) \subset kX \subset x' + 2^{2^{C_2 d^2 2^{6d}}} (H + Q).$$

Note that any GAP $Q = \{a_0 + x_1 a_1 + \cdots + x_r a_r \mid -N_i \leq x_i \leq N_i \text{ for all } 1 \leq i \leq r\}$ is contained in a symmetric GAP $Q' = \{x_0 a_0 + x_1 a_1 + \cdots + x_r a_r \mid -1 \leq x_0 \leq 1, -N_i \leq x_i \leq N_i \text{ for all } 1 \leq i \leq r\}$. Thus, by combining Theorem C.3 with Theorem C.2 we obtain the following

Corollary C.4. *Whenever $d, k \geq 1$ and $X \subset G$ is a non-empty finite set such that*

$$k^d |X| \geq 2^{2^{C_2 d^2 2^{6d}}} |kX|,$$

then there exists a 2-proper symmetric coset progression $H + P$ of rank $0 \leq d' \leq d$ and size $|H + P| \leq 2^d (C_1 d)^{3d^2/2} 2^{d^2 C_2 d^2 6d} |kX|$ such that

$$kX \subset H + P.$$

As for Theorem 7.3, the explicit constants in Corollary C.4 will not be important. (Although a more careful analysis would allow α_n to be as small as $n^{-1+O(\frac{1}{\log \log n})}$ here, and hence in our main theorems. But in order to keep our presentation simple we will not work with this technical assumption, only staying with $\alpha_n \geq n^{-1+\varepsilon}$.) Now we give a detailed proof of Theorem 7.3. In general our method follows that of [22], but the details are more complicated because we have to obtain an actual inverse result in $\mathbf{Z}/p\mathbf{Z}$, as well as we need to take into account the almost sharp sparsity of the randomness.

Proof. (of Theorem 7.3) First, for convenience we will pass to symmetric distributions. Let $\psi = \nu_1 - \nu_2$ be the symmetrization of ν and let ψ' be a lazy version of ψ that

$$\mathbf{P}(\psi' = x) = \begin{cases} \frac{1}{2} \mathbf{P}(\psi = x) & \text{if } x \neq 0 \\ \mathbf{P}(\psi' = x) = \frac{1}{2} \mathbf{P}(\psi = x) + \frac{1}{2}, & \text{if } x = 0. \end{cases}$$

Notice that ψ' is symmetric as ψ is symmetric. Similarly to (21), we can check that $\max_x \mathbf{P}(\psi = x) \leq 1 - \alpha_n$, and so

$$\sup_x \mathbf{P}(\psi' = x) \leq 1 - \alpha_n/2.$$

We assume that $\{0, \pm t_1, \dots, \pm t_l\}$ is the range of ψ' and $\mathbf{P}(\psi' = t_j) = \mathbf{P}(\psi' = -t_j) = \beta_j/2$ for $1 \leq j \leq l$, and that $\mathbf{P}(\psi' = 0) = \beta_0$, where $t_{j_1} \pm t_{j_2} \neq 0 \pmod{p}$ for all $j_1 \neq j_2$.

Consider $a \in \mathbf{Z}/p\mathbf{Z}$ where the maximum is attained, $\rho = \rho(w) = \mathbf{P}(S = a)$, here $S = \nu_1 w_1 + \cdots + \nu_n w_n = a$. Using the standard notation $e_p(x)$ for $\exp(2\pi\sqrt{-1}x/p)$, we have

$$\rho = \mathbf{P}(S = a) = \mathbf{E} \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e_p(x(S - a)) = \mathbf{E} \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e_p(xS) e_p(-xa) \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{E} e_p(xS)|. \quad (37)$$

By independence

$$|\mathbf{E} e_p(xS)| = \prod_{i=1}^n |\mathbf{E} e_p(x\nu_i w_i)| \leq \prod_{i=1}^n \left(\frac{1}{2} (|\mathbf{E} e_p(x\nu_i w_i)|^2 + 1) \right) = \prod_{i=1}^n |\mathbf{E} e_p(x\psi' w_i)| = \prod_{i=1}^n (\beta_0 + \sum_{j=1}^l \beta_j \cos \frac{2\pi x t_j w_i}{p}).$$

It follows that

$$\rho \leq \frac{1}{p} \left| \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \prod_{i=1}^n \left(\beta_0 + \sum_{j=1}^l \beta_j \cos \frac{2\pi x t_j w_i}{p} \right) \right| \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \prod_{i=1}^n \left(\beta_0 + \sum_{j=1}^l \beta_j \left| \cos \frac{\pi x t_j w_i}{p} \right| \right), \quad (38)$$

where we made the change of variable $x \rightarrow x/2$ (in $\mathbf{Z}/p\mathbf{Z}$) and used the triangle inequality.

By convexity, we have that $|\sin \pi z| \geq 2\|z\|$ for any $z \in \mathbf{R}$, where $\|z\| := \|z\|_{\mathbf{R}/\mathbf{Z}}$ is the distance of z to the nearest integer. Thus,

$$\left| \cos \frac{\pi x}{p} \right| \leq 1 - \frac{1}{2} \sin^2 \frac{\pi x}{p} \leq 1 - 2 \left\| \frac{x}{p} \right\|^2. \quad (39)$$

Hence for each w_i

$$\beta_0 + \sum_{j=1}^l \beta_j \left| \cos \frac{\pi x t_j w_i}{p} \right| \leq 1 - 2 \sum_{j=1}^l \beta_j \left\| \frac{x t_j w_i}{p} \right\|^2 \leq \exp(-2 \sum_{j=1}^l \beta_j \left\| \frac{x t_j w_i}{p} \right\|^2).$$

Consequently, we obtain a key inequality

$$\rho \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \prod_{i=1}^n (\beta_0 + \sum_{j=1}^l \beta_j |\cos \frac{\pi x t_j w_i}{p}|) \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \exp(-2 \sum_{i=1}^n \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2). \quad (40)$$

Large level sets. Now we consider the level sets $S_m := \{\xi | \sum_{i=1}^n \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2 \leq m\}$. We have

$$n^{-C} \leq \rho \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \exp(-2 \sum_{i=1}^n \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2) \leq \frac{1}{p} + \frac{1}{p} \sum_{m \geq 1} \exp(-2(m-1)) |S_m|.$$

As p is assumed to be much larger than n^C , and as $\sum_{m \geq 1} \exp(-m) < 1$, there must be a large level set S_m such that

$$|S_m| \exp(-m+2) \geq \rho p. \quad (41)$$

In fact, since $\rho \geq n^{-C}$, we can assume that $m = O(\log n)$.

Double counting and the triangle inequality. By double counting we have

$$\sum_{i=1}^n \sum_{x \in S_m} \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2 = \sum_{x \in S_m} \sum_{i=1}^n \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2 \leq m |S_m|.$$

So, for most w_i

$$\sum_{x \in S_m} \sum_{j=1}^l \beta_j \|\frac{x t_j w_i}{p}\|^2 \leq \frac{m}{n'} |S_m|. \quad (42)$$

More precisely, by averaging, the set of w_i satisfying (42) has size at least $n - n'$. We call this set W' . The set $\{w_1, \dots, w_n\} \setminus W'$ has size at most n' and this is the exceptional set that appears in Theorem 7.3. In the rest of the proof, we are going to show that W' is a dense subset of a proper GAP.

Since $\|\cdot\|$ is a norm, by the triangle inequality, we have for any $a \in kW'$

$$\sum_{x \in S_m} \sum_{j=1}^l \beta_j \|\frac{x t_j a}{p}\|^2 \leq k^2 \frac{m}{n'} |S_m|. \quad (43)$$

More generally, for any $k' \leq k$ and $a \in k'W'$

$$\sum_{x \in S_m} \sum_{j=1}^l \beta_j \|\frac{x t_j a}{p}\|^2 \leq k'^2 \frac{m}{n'} |S_m|. \quad (44)$$

Dual sets. Set

$$\alpha'_n := \sum_{j=1}^l \beta_j = 1 - \beta_0.$$

Then by definition of ψ , we have

$$\alpha'_n \geq \alpha_n/2 \geq n^{-1+\varepsilon}.$$

Define

$$S_m^* := \{a | \sum_{x \in S_m} \sum_{j=1}^l \beta_j \|\frac{x t_j a}{p}\|^2 \leq \frac{\alpha'_n}{200} |S_m|\}$$

where the constant 200 is ad hoc and any sufficiently large constant would do. We have

$$|S_m^*| \leq \frac{8p}{|S_m|}. \quad (45)$$

To see this, define $T_a := \sum_{x \in S_m} \sum_{j=1}^l \beta_j \cos \frac{2\pi a t_j x}{p}$. Using the fact that $\cos 2\pi z \geq 1 - 100\|z\|^2$ for any $z \in \mathbf{R}$, we have, for any $a \in S_m^*$

$$T_a \geq \sum_{x \in S_m} (1 - 100 \sum_{j=1}^l \beta_j \|\frac{x t_j a}{p}\|^2) \geq \frac{\alpha'_n}{2} |S_m|.$$

One the other hand, using the basic identity $\sum_{a \in \mathbf{Z}/p\mathbf{Z}} \cos \frac{2\pi az}{p} = p\mathbf{I}_{z=0}$, we have (taking into account that $t_{j_1} \neq t_{j_2} \pmod{p}$)

$$\sum_{a \in \mathbf{Z}/p\mathbf{Z}} T_a^2 \leq 2p|S_m| \sum_j \beta_j^2 \leq 2p|S_m| \max_{1 \leq j \leq l} \beta_j \left(\sum_{j=1}^l \beta_j \right) \leq 2p|S_m| \alpha'_n{}^2.$$

Equation (45) then follows from the last two estimates and averaging.

Next, for a properly chosen constant c_1 we set

$$k := c_1 \sqrt{\frac{\alpha'_n n'}{m}}.$$

By (44) we have $\cup_{k'=1}^k k'W' \subset S_m^*$. Next, set

$$W'' := W' \cup \{0\}.$$

We have $kW'' \subset S_m^* \cup \{0\}$. This results in the critical bound

$$|kW''| = O\left(\frac{p}{|S_m|}\right) = O(\rho^{-1} \exp(-m+2)). \quad (46)$$

The long range inverse theorem. We are now in the position to apply Corollary C.4 with X as the set of distinct elements of W'' . As $k = \Omega\left(\sqrt{\frac{\alpha'_n n'}{m}}\right) = \Omega\left(\sqrt{\frac{\alpha'_n n'}{\log n}}\right)$,

$$\rho^{-1} \leq n^C \leq k^{4C/\varepsilon+1}. \quad (47)$$

It follows from Corollary C.4 that kX is a subset of a 2-proper symmetric coset progression $H + P$ of rank $r = O_{C,\varepsilon}(1)$ and cardinality

$$|H + P| \leq O_{C,\varepsilon}|kX|.$$

Now we use the special property of $\mathbf{Z}/p\mathbf{Z}$ that it has only trivial proper subgroups. As $|kX| = O(n^C)$, and as $p \gg n^C$, the only possibility that $|kX| \gg |H + P|$ is that $H = \{0\}$. Consequently, kX is now a subset of P , a 2-proper symmetric GAP of rank $r = O_{C,\varepsilon}(1)$ and cardinality

$$|P| \leq O_{C,\varepsilon}|kX|. \quad (48)$$

To this end, we apply the following dividing trick from [22, Lemma A.2].

Lemma C.5. *Assume that $0 \in X$ and that $P = \{\sum_{i=1}^r x_i a_i : |x_i| \leq N_i\}$ is a 2-proper symmetric GAP that contains kX . Then*

$$X \subset \left\{ \sum_{i=1}^r x_i a_i : |x_i| \leq 2N_i/k \right\}.$$

Proof. (of Lemma C.5) Without loss of generality, we can assume that $k = 2^l$. It is enough to show that $2^{l-1}X \subset \{\sum_{i=1}^r x_i a_i : |x_i| \leq N_i/2\}$. Since $0 \in X$, $2^{l-1}X \subset 2^lX \subset P$, any element x of $2^{l-1}X$ can be written as $x = \sum_{i=1}^r x_i a_i$, with $|x_i| \leq N_i$. Now, since $2x \in P \subset 2P$ and $2P$ is proper GAP (as P is 2-proper), we must have $0 \leq |2x_i| \leq N_i$. \square

Combining (48) and Lemma C.5 we thus obtain a GAP Q that contains X and

$$\begin{aligned} |Q| &= O_{C,\varepsilon}(k^{-r}|kX|) = O_{C,\varepsilon}(k^{-r}|kW''|) = O_{C,\varepsilon} \left(\rho^{-1} \exp(-m) \left(\sqrt{\frac{\alpha'_n n'}{m}} \right)^{-r} \right) \\ &= O_{C,\varepsilon}(\rho^{-1}(\alpha'_n n')^{-r}), \end{aligned}$$

concluding the proof. \square

APPENDIX D. SPARSE SUBSPACES FOR THE LAPLACIAN CASE: PROOF OF LEMMAS 8.21 AND 8.22

Our methods are almost identical to those of Lemma 7.4 and 7.5, with a few minor exceptions.

Proof. (of Lemma 8.21) Argue similarly as in the proof of Lemma 7.4, it suffices to show that the following holds with probability at least $1 - e^{-c_1 \beta_n n/2}$. For any $1 \leq t \leq 144\beta_n^{-1}$, and any $\sigma \in \binom{[n]}{t}$, there are at least two columns X_i, X_j with $i, j \notin \sigma$ whose restriction $(X_j - X_i)|_\sigma$ has exactly one non-zero entry.

For a given σ of size t , assume that $\{i_1, \dots, j_{n_0-t}\} \subset [n_0] \setminus \sigma$. For $i \in \{1, 3, \dots, 2\lfloor(n_0-t)/2\rfloor - 1\}$, consider the vectors $Y_i = X_{j_{i+1}}|_\sigma - X_{j_i}|_\sigma$. Note that as $j_i, j_{i+1} \notin \sigma$ and $X_{j_i} \in \mathcal{T}_{j_i}$ and $X_{j_{i+1}} \in \mathcal{T}_{j_{i+1}}$, the entries of Y_i are iid copies of the symmetrized random variable $\psi = \xi - \xi'$, where ξ', ξ are independent and have distribution ξ_n . Recall that with $1 - \beta'_n = \mathbf{P}(\psi = 0)$, then $\beta_n \leq \beta'_n \leq 2\beta_n$. Now let p_σ be the probability that all $Y_i|_\sigma, i \in \{1, 3, \dots, 2\lfloor(n_0-t)/2\rfloor - 1\}$ fail to have exactly one non-zero entry (in \mathbf{Z}), then by independence of the columns and of the entries

$$p_\sigma = (1 - t\beta'_n(1 - \beta'_n)^{t-1})^{\lfloor(n_0-t)/2\rfloor} \leq (1 - t\beta'_n e^{-(t-1)\beta'_n})^{\lfloor(n_0-t)/2\rfloor} \leq e^{-nt\beta'_n e^{-(t-1)\beta'_n}/4},$$

where we used $n_0 - t > n/2$ because $\alpha_n \geq 6 \log n/n$ and $t \leq 144\beta_n^{-1}$. The rest of the proof is similar to that of Lemma 7.4. \square

Proof. (of Lemma 8.22) For $\sigma \subset [n]$ with $144\beta_n^{-1} \leq t = |\sigma| \leq c_0 n$, consider the event that W_{n_0}/p is normal to a vector w with $\text{supp}(w) = \sigma$ but not with any other vector of smaller support size. With a loss of a multiplicative factor $\binom{n}{t}$ in probability, we assume that $\sigma = \{1, \dots, t\}$. Consider the submatrix $L_{t \times n_0}$ of $L_{n \times n}$ consisting of the first t rows and first n_0 columns of $L_{n \times n}$. Since the restriction $w|_\sigma$ of w to the first t coordinates is normal to all the columns of $L_{t \times n_0}/p$, the matrix $L_{t \times n_0}/p$ has rank $t-1$ (if $p=0$, we mean rank over \mathbf{R}). We assume that the column space of $L_{t \times n_0}/p$ is spanned the columns $\{X_{i_1}, \dots, X_{i_{t-1}}\}$ for some $\{i_1, \dots, i_{t-1}\} \subset [n_0]$.

Note that for $p > 2n^T$, the value of ξ_n is determined by its value mod p , and so $\beta_n = 1 - \max_{x \in \mathbf{Z}/p\mathbf{Z}} \mathbf{P}(\xi_n = x)$. If we fix $X_{i_1}|_\sigma, \dots, X_{i_{t-1}}|_\sigma$ such that the subspace $W_{i_1, \dots, i_{t-1}}|_\sigma/p$ generated by these vectors has a normal vector with all t coordinates non-zero, then by Theorem 6.3, the probability that $X_i|_\sigma/p \in W_{i_1, \dots, i_{t-1}}|_\sigma/p$ for all $i \in [n_0] \setminus (\sigma \cup \{i_1, \dots, i_{t-1}\})$ (as for these vectors the entries of \mathcal{T}_i restricted to σ are independent) is at most

$$\left(\frac{1}{p} + \frac{2}{\sqrt{\beta_n t}}\right)^{n_0-2t-1} \leq \left(\frac{1}{p} + \frac{2}{\sqrt{\beta_n t}}\right)^{(1-3c_0)n} \leq \left(\frac{2}{3}\right)^{n/2}$$

as long as C_0 is sufficiently large and c_0 is sufficiently small. Thus the total probability of the event in the lemma is at most

$$\sum_{144\beta_n^{-1} \leq t \leq c_0 n} \binom{n}{t} \binom{n_0}{t-1} \left(\frac{2}{3}\right)^{n/2} \leq \left(\frac{2}{3}\right)^{n/4}.$$

\square

REFERENCES

- [1] A. Basak and M. Rudelson, Invertibility of sparse non-Hermitian matrices, *Advances in Mathematics*, 310 (2017), 426–483.
- [2] M. Bhargava, The geometric sieve and the density of squarefree values of invariant polynomials, preprint, arxiv.org/abs/1402.0031.
- [3] J. Bourgain, V. Vu and P. M. Wood, On the singularity probability of discrete random matrices, *Journal of Functional Analysis* 258 (2010), no. 2, 559–603.
- [4] G. Chinta, N. Kaplan, and S. Koplewitz, The cotype zeta function of \mathbf{Z}^d , preprint, arxiv.org/abs/1708.08547.
- [5] J. Clancy, T. Leake, N. Kaplan, S. Payne, and M. M. Wood, On a Cohen-Lenstra heuristic for Jacobians of random graphs, *Journal of Algebraic Combinatorics* 42 (2015), no. 3, 701–723.
- [6] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983* (Noordwijkerhout, 1983), volume 1068 of *Lecture Notes in Math.*, pages 331–762. Springer, Berlin, 1984.
- [7] T. Ekedahl, An infinite version of the Chinese remainder theorem, *Comment. Math. Univ. St. Paul.*, 40 (1991), no. 1, 53–59.
- [8] M. Farrell and L. Levine, Co-Eulerian graphs, *Proceedings of the American Mathematical Society*, 144 (2016), 2847–2860.
- [9] M. Farrell and L. Levine, Multi-Eulerian tours of directed graphs, *Electron. J. Combin.*, 23 (2016), no. 2, Paper 2.21, 7 pp.

- [10] S. D. Fisher and M. N. Alexander, Classroom Notes: Matrices over a Finite Field, Amer. Math. Monthly, 73 (1966), no. 6, 639–641.
- [11] E. Friedman and L. C. Washington. On the distribution of divisor class groups of curves over a finite field. In Theorie des nombres (Quebec, PQ, 1987), pages 227–239. de Gruyter, Berlin, 1989.
- [12] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, Period. Math. Hungar. 8 (1977), no. 3-4, 197–211.
- [13] A. E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp and D. B. Wilson, Chip-firing and rotor-routing on directed graphs, in *In and out of equilibrium. 2*, 331–364, Progr. Probab., 60, Birkhäuser, Basel.
- [14] J. Kahn and J. Komlós, Singularity probabilities for random matrices over finite fields. Combin. Probab. Comput., 10(2):137–157, 2001.
- [15] J. Kahn, J. Komlós and E. Szemerédi, On the probability that a random ± 1 matrix is singular, J. Amer. Math. Soc. 8 (1995), 223–240.
- [16] J. Komlós, On the determinant of $(0 - 1)$ -matrices, Studia Sci. Math. Hungar. 2 (1967), 7–22.
- [17] S. Koplewitz, Sandpile groups and the coeulerian property for random directed graphs, Advances in Applied Mathematics, Volume 90, September 2017, Pages 145–159.
- [18] S. Koplewitz, Random Graphs, Sandpile Groups, and Surjectivity of Random Matrices, Thesis (Ph.D.)-Yale University. 2017. 58 pp. ISBN: 978-0355-01826-4.
- [19] K. Maples, Singularity of Random Matrices over Finite Fields, preprint, arxiv.org/abs/1012.2372
- [20] H. Nguyen, Inverse Littlewood-Offord problems and the singularity of random symmetric matrices, Duke Math. J., Vol. 161, 4 (2012), 545–586.
- [21] H. Nguyen and E. Paquette, Surjectivity of near square matrices, submitted, Combinatorics, Probability and Computing, Vol. 29, 2 (2020), 267–292.
- [22] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, Advances in Mathematics, Vol. 226 6 (2011), 5298–5319.
- [23] P. Q. Nguyen and I. E. Shparlinski, Counting co-cyclic lattices, SIAM J. Discrete Math., 30 (2016), no. 3, 1358–1370.
- [24] V. M. Petrogradsky, Multiple zeta functions and asymptotic structure of free abelian groups of finite rank, J. Pure Appl. Algebra, 208 (2007), no. 3, 1137–1158.
- [25] T. V. Pham, Orbita of rotor-router operation and stationary distribution of random walks on directed graphs, Adv. in Appl. Math., 70 (2015), 45–53.
- [26] B. Poonen, Squarefree values of multivariable polynomials, Duke Math. J., 118 (2003), no. 2, 353–373.
- [27] M. Rudelson and R. Vershynin, The Littlewood-Offord problem and invertibility of random matrices, Advances in Mathematics 218 (2008), 600–633.
- [28] R. Stanley, Smith normal form in Combinatorics, Journal of Combinatorial Theory, Series A Volume 144, November 2016, Pages 476–495.
- [29] R. Stanley and Y. Wang, The Smith normal form distribution of a random integer matrix, SIAM J. Discrete Math., 31 (3), 2017, 2247–2268.
- [30] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, Journal of the A. M. S 20 (2007), 603–673.
- [31] T. Tao and V. Vu, John-type theorems for generalized arithmetic progressions and iterated sumsets, Adv. Math. 219 (2008), no. 2, 428–449.
- [32] T. Tao and V. Vu, Inverse Littlewood-Offord theorems and the condition number of random matrices, Annals of Mathematics (2), 169 (2009), no 2, 595–632.
- [33] M. M. Wood, The distribution of sandpile groups of random graphs, Journal of the A. M. S, 30 (2017), pp. 915–958.
- [34] M. M. Wood, Random integral matrices and the Cohen-Lenstra Heuristics, Amer. J. Math., Vol. 141, 2 (2019), pp. 383–398.

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 231 W 18TH AVE, COLUMBUS, OH 43210 USA
Email address: nguyen.1261@math.osu.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, 480 LINCOLN DRIVE, MADISON, WI 53705 USA
Email address: mmwood@math.wisc.edu