On Some Universality Problems in Combinatorial Random Matrix Theory

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Sean Meehan, B.S., M.S.

Graduate Program in Department of Mathematics

The Ohio State University

2019

Dissertation Committee:

Hoi Nguyen, Ph.D., Advisor

David Sivakoff, Ph.D.

Elliot Paquette, Ph.D.

Abstract

This dissertation will exhibit some universal behavior of random matrices in two settings. First, we will study the eigenvectors of random symmetric matrices M_n whose entries are sampled from symmetric distributions. We will then shift our study from characteristic zero to matrices over $\mathbf{F}_p$, instead studying the random normal vector, a (not necessarily unique) random vector orthogonal to each column. We will see that both of these respective vectors, the eigenvectors of M_n and the normal vector over $\mathbf{F}_p$, behave like the uniform model.

Dedicated to my wife Amelia, whose love and support know no bounds.

Acknowledgments

This dissertation would not have been possible without the help of my advisor, Hoi Nguyen. He was the one constant throughout this entire process, and I would not have been able to complete this without his guidance. Thank you so much, Hoi.

Many Ohio State colleagues have also been instrumental in this process, both in academic support as well as non-academic. To Kevin, Osama, Evan, Katie, Samir, Hanbaek, and many others, thank you. Also thank you to Ohio State University for providing me the ability to attend this program and learn from your incredible staff.

A million thank you's to my wife, Amy. With all that you do for me on a day-to-day basis, this Ph.D. is yours as much as it is mine, if not more.

I would also like to thank my first teachers - my parents. Your parenting has brought me here and I owe a lot to it. To my brother, whose late-night conversations kept me going, thank you.

Finally, to my new-found friends, who helped me carry through in more ways they know, thank you. Ben, Stacee, Jon, and Tyler, we only crossed paths just a year or two ago, but it feels like we've known each other forever.

Vita

May 20, 1991	Born - Secaucus, NJ
2013	B.S. Mathematics, University of Notre Dame
2016	M.S. Mathematics, The Ohio State University
2013 - present	Graduate Teaching Associate, Department of Mathematics, The Ohio State University.

Publications

Johnson, C., Lins, B., Luo, V., and **Meehan, S.** “Ordering graphs in a normalized singular value measure” *Involve.*, Vol. 8 No. 2: 263-273, 2015.

Meehan, S., Tefera, A., Weselcouch, M., and Zeleke, A. “Proofs of Ruehr’s identities.” *Integers.*, Vol. 14 A10: 1-6, 2014.

Meehan, S., and Nguyen, H. “Eigenvectors of Wigner matrices of symmetric entry distributions.” Proceedings of the American Mathematical Society, (2019) Volume 147, Number 2, 835-847.

Fields of Study

Major Field: Department of Mathematics

Table of Contents

List of Figures

Figure **Page**

Chapter 1

Introduction to Random Matrix Theory

This dissertation seeks to extend some known results in the field of random matrix theory. To be precise, a random matrix is a matrix whose entries are all random variables. As a natural, probabilistic extension from, say, random variables that take values over a $\mathbf{R}-$valued sample space to that of a matrix-valued sample space, random matrix theory is a blossoming area that is now widely studied.

There are no limitations on the scope of random matrices that can be studied. The random variables that comprise the entries can be real or complex, or the matrix can be symmetric (Hermitian) or non-symmetric (non-Hermitian), square or rectangular, orthogonal or unitary, etc. Relevant statistics that give us info and tell us the story of our matrix distribution involve but are not limited to the eigenvalues, eigenvectors, and even rank in some combinatorial models in the discrete setting.

In a multitude of disciplines, it is often convenient to collect large amounts of data and then analyze this data. This is one of many ways in which random matrix theory can serve a purpose. As a result, the study of random matrices has applications in mathematical physics, data science, number theory, and statistics to name a few.

1.1 Introductory Random Matrix Models

To get a sense of the subject, we begin by examining a few models of random matrices. Once these models are in place, we will be able to assess the characteristics of some of these naturally-occurring random matrices later in this work.

To begin, we define the Erdős-Rényi graph on n vertices with probability p. Suppose that we construct the empty graph on n vertices E_n, a collection of n nodes with no current edge set. Counting the number of possible edges that can be constructed, we arrive at $\binom{n}{2}$ possible edge pairings. The Erdős-Rényi graph $G(n, p)$ is simply chosen uniformly among all of the possible edge drawings:

Definition 1.1.1. Let n be a positive integer and $0 \leq p \leq 1$. The Erdős-Rényi graph $G(n, p)$ is defined to be the graph on n vertices where each potential edge is drawn independently with probability p.

With the selection of a Erdős-Rényi graph $G(n, p)$ being a random process, we can look to the adjacency matrix of $G(n, p)$ as our first random matrix model:

Definition 1.1.2 (Adjacency Matrix of Random Graph Model)**.** Let $G(n, p)$ be the Erdős-Rényi graph on n vertices with probability p. We denote $A_n(p)$ as the zero-one adjacency matrix of $G(n, p)$.

Fixing the diagonal entries to be zero and fixing each strictly lower-diagonal entry $\{a_{ij} : i < j\}$ to be identical to its corresponding strictly upper-diagonal entry $\{a_{ji} : i < j\}$, the resulting matrix $A_n(p)$ is a symmetric random matrix whose strictly upper-diagonal entries are independent and identically distributed Bernoulli(p) random variables.

Next, we look to a more general random matrix model. The Wigner matrix model with atom variables ψ_1, ψ_2 is beneficial to consider because its entries can be modeled with any distribution, which contrasts directly with the previous random adjacency matrix model, whose entries have to be either deterministically zero or modeled with the Bernoulli(p) distribution:

Definition 1.1.3 (Random Symmetric Wigner Matrix Model)**.** Let ψ_1, ψ_2 be $\mathbf{R}-$valued random variables with mean zero. Let W be a $n \times n$ random real symmetric matrix. We will say that W is a symmetric Wigner matrix with atom variables ψ_1, ψ_2 if:

(1) (independence) The upper diagonal entries $\{w_{ij} : 1 \leq i \leq j \leq n\}$ are all independent.

(2) (off-diagonal entries) The off-diagonal entries $\{w_{ij} : 1 \leq i < j \leq n\}$ are independent and identically distributed copies of ψ_1.

(3) (diagonal entries) The diagonal entries $\{w_{ii} : 1 \leq i \leq n\}$ are independent and identically distributed copies of ψ_2.

Now that we have defined symmetric Wigner matrices, we can look at a specific example of a symmetric Wigner matrix. The GOE (Gaussian Orthogonal Ensemble) is a naturally-occurring symmetric Wigner random matrix which is of interest:

Definition 1.1.4 (Gaussian Orthogonal Ensemble)**.** We say that W is a member of the Gaussian Orthogonal Ensemble (GOE) if W is a symmetric Wigner random matrix with atom variables ψ_1, ψ_2, where ψ_1 is the standard normal distribution and ψ_2 is a normal random variable with mean zero and variance two.

It's worth noting that while all of the previous models were symmetric in nature, there are plenty of random matrix models that are not required to be symmetric. This non-symmetric setting will often also be of interest to this paper.

Notation. Throughout this document, we regard n as an asymptotic parameter going to infinity (in particular, we will implicitly assume that n is larger than any fixed constant, as our claims are all trivial for fixed n), and allow all mathematical objects in the paper to implicitly depend on n unless they are explicitly declared to be "fixed" or "constant". We write $X = O(Y)$, $X \ll Y$, or $Y \gg X$ to denote the claim that $|X| \leq CY$ for some fixed C; this fixed quantity C is allowed to depend on other fixed quantities such as K_1, K_2 of ξ unless explicitly declared otherwise. We also use $o(Y)$ to denote any quantity bounded in magnitude by $c(n)Y$ for some $c(n)$ that goes to zero as $n \to \infty$. For a square matrix M_n and a number λ, for short we will write $M_n - \lambda$ instead of $M_n - \lambda I_n$. At times, we will use $\mathbf{r}_i(M_n)$ and $\mathbf{c}_i(M_n)$ to denote the i^{th} row and i^{th} column of M_n, respectively. All the norms in this note, if not specified, will be the usual ℓ_2-norm. $\mathbf{P}$ will denote probability and $\mathbf{E}$ will denote expectation. When working with an event $\mathcal{E}$, we will often write $\overline{\mathcal{E}}$ as the complement event. We will write $\exp(x)$ for the exponential function e^x. $[n]$ will denote the set of positive integers $\{1, \cdots, n\}$. Given a set $I \subset [n]$ and a vector $\mathbf{x} = (x_1, \cdots x_n)$, we will when necessary use $\mathbf{x}_I$ to refer to the truncated vector with components indexed by I. We will write $\mathbf{supp}(\mathbf{x})$ to be the support of the vector $\mathbf{x}$, i.e. $\mathbf{supp}(\mathbf{x}) := \{i \in [n] : x_i \neq 0\}$. The quantity $\mathbf{x} \cdot \mathbf{w}$ will refer to the standard Euclidean dot product $\sum_{i=1}^n x_i w_i$, which we will also denote as $\langle \mathbf{x}, \mathbf{w} \rangle$ when convenient. We will say $\mathbf{w}$ is a normal vector for a subspace H if $\mathbf{x} \cdot \mathbf{w} = 0$ for all $\mathbf{x} \in H$. We will use $|| \cdot ||_{\mathbf{R}/\mathbf{Z}}$ to denote the distance to the nearest integer.

1.2 Random Unit Eigenvectors in the GOE Regime

As a specific example of some properties that may result in a random matrix model, we will look at the random unit eigenvectors in our Gaussian Orthogonal Ensemble. The following results are mainly expository and can be found in [21] by O'Rourke et al.

For this section, W will be a member of the Gaussian Orthogonal Ensemble. Via the spectral theorem, we can decompose W as $W = UDU^T$, where U is an orthogonal matrix with columns consisting of the eigenvectors of W and D is a diagonal matrix whose diagonal entries consist of the eigenvalues of W. But W, as a member of the GOE, is invariant under orthogonal transformations. This means that U and D are independent. Thus we see that the eigenvectors of W are uniformly distributed on the unit half-sphere

$$S_+^{n-1} := \{(x_1, \cdots, x_n) \in S^{n-1} : x_1 > 0\}.$$

So to study the properties of unit random eigenvectors of W, where W is a member of the Gaussian Orthogonal Ensemble, we can simply study the properties of random unit vectors pulled from the half-sphere with uniform distribution. Note that $-W$ and W observe the same law; thus, our study reduces to the study of random unit vectors uniformly distributed on the unit sphere S^{n-1}. In what follows, we will list a few interesting properties of uniform vectors sampled over S^{n-1}. Although we won't address the universality aspect of these properties for other models, we insert these here for completeness.

1.2.1 Properties of the Extreme Coordinates and Norm

A natural starting point is the size of the components of the random unit eigenvector **v**; more specifically, the magnitude of the smallest component and the magnitude of the largest component. Let

$$||\mathbf{v}||_{l^\infty} := \max_i |v_i|$$

be the $l^\infty-$norm of **v** and

$$||\mathbf{v}||_{min} := \min_i |v_i|$$

be the minimal coordinate of **v** in absolute value. Then we have:

Theorem 1.2.1 (Order of Largest Component)**.** *For any fixed constant $C > 1$, we have*

$$||\mathbf{v}||_{l^\infty} \leq \sqrt{\frac{2C^3 \log n}{n}}$$

with probability at least

$$1 - 2n^{1-C} - e^{-\frac{(C-1)^2}{4C^2}n}.$$

Theorem 1.2.2 (Order of Smallest Component)**.** *Let $n \geq 2$. For any fixed constants $0 \leq c < 1$ and $a > 1$, we have*

$$||\mathbf{v}||_{min} \geq \frac{c}{an^{3/2}}$$

with probability at least

$$e^{-2c} - e^{-\frac{a^2-\sqrt{2a^2-1}}{2}n}.$$

Before we discuss our next result, we quickly recall the definition of the $p-$norm of a fixed vector.

Definition 1.2.3. Let $p \geq 1$ be a fixed real number. The l^p-norm of the fixed vector $\mathbf{v} := (v_1, \cdots, v_n)$ is the quantity

$$||\mathbf{v}||_{l^p} := (\sum_{i=1}^{n} |v_i|^p)^{1/p}.$$

Thus in the instance that $p = 2$, we have the standard Euclidean norm. While these vectors $\mathbf{v}$ have unit Euclidean norm, their l^p-norms vary as $\mathbf{v}$ varies over the unit sphere. For instance, Theorem 1.2.1 looks at how $\mathbf{v}$ behaves under the l^∞ norm. We can quantify how $\mathbf{v}$ behaves under the l^p-norm almost surely:

Theorem 1.2.4 (Order of l^p-norm)**.** *Let $p \geq 1$ be given. Then there exists $c_p > 0$ such that*

$$||\mathbf{v}||_{l^p}^p = n^{1-p/2}c_p + o(n^{1-p/2})$$

almost surely.

For our next result, we first quickly recall the beta distribution: the beta distribution on $0 \leq x \leq 1$ with parameters $\alpha, \beta > 0$ has probability density function

$$f(x) := \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & 0 \leq x \leq 1 \\ 0 & x \in \mathbf{R} \setminus [0,1] \end{cases}$$

where $B(\alpha, \beta)$ is the normalization constant

$$B(\alpha, \beta) := \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

When a random variable ξ is beta-distributed with parameters α and β, we will write $\xi \sim \text{Beta}(\alpha, \beta)$.

We can consider how the Euclidean norm behaves over a subset of the coordinates $S \subset \{1, \cdots, n\}$; that is to say, we define

$$||\mathbf{v}_S|| := \sqrt{\sum_{i \in S} |v_i|^2}.$$

It turns out that for any subset S, the norm $||\mathbf{v}||_S^2$ is beta-distributed with parameters $|S|/2$ and $n/2 - |S|/2$. As a corollary, we can quantify the expected value and variance of $||\mathbf{v}||_S^2$ through our knowledge of the beta distribution.

Theorem 1.2.5 ($||\mathbf{v}||_S^2$ is beta-distributed)**.** *Let S be a proper, nonempty subset of $\{1, \cdots, n\}$. Then $||\mathbf{v}||_S^2$ is distributed according to the beta distribution*

$$||\mathbf{v}||_S^2 \sim Beta(\frac{|S|}{2}, \frac{n-|S|}{2}).$$

Corollary 1.2.6. *Let S be a proper, nonempty subset of $\{1, \cdots, n\}$. Then $||\mathbf{v}||_S^2$ has mean $\frac{|S|}{n}$ and variance $\frac{|S|(n-|S|)}{n^2(n/2+1)}$.*

It's often natural to ask how concentrated a random variable is around its mean. As an additional corollary of Theorem 1.2.5, we can achieve such a central limit theorem type result for the norm of the truncated vector $\mathbf{v}_S$:

Corollary 1.2.7. *Let $\delta \in (0,1)$ be fixed, and let $S_n \subset \{1, \cdots, n\}$ with $|S_n| = \lfloor \delta n \rfloor$. Then we have*

$$\sqrt{\frac{n^3}{2|S_n|(n-|S_n|)}}(||\mathbf{v}||_{S_n}^2 - \frac{|S_n|}{n}) \to \mathbf{N}(0,1)$$

in distribution as $n \to \infty$.

We can go further and form a concentration inequality between $||\mathbf{v}||_S^2$ and its mean $|S|/2$:

Theorem 1.2.8. *With $S \subset \{1, \cdots, n\}$ as before, then for any $t > 0$, we have*

$$\big| ||\mathbf{v}||_S^2 - \frac{|S|}{n} \big| \leq \frac{8}{n}(\sqrt{nt} + t)$$

with probability at least

$$1 - e^{-cn} - 4e^{-t},$$

where $c > 0$ is an absolute constant.

1.2.2 Properties of the Extreme Order Statistics

Consider the max and min order statistics

$$\max_{S \subset [n]:|S| = \lfloor \delta n \rfloor} ||\mathbf{v}||_S$$

and

$$\min_{S\subset[n]:|S|=\lfloor\delta n\rfloor} ||\mathbf{v}||_S.$$

We will exhibit some properties of these order statistics, but first, we need to recall some characteristics of the χ^2 distribution. The χ^2 distribution with k degrees of freedom is defined to be the distribution of $X = \sum_{i=1}^{k} Z_i^2$ where $Z_1, \cdots, Z_k$ are independent standard normal random variables. Let F be the CDF of the χ^2 distribution with $k = 1$ degree of freedom. Furthermore, we define the quantile function of F as follows:

Definition 1.2.9. The quantile function $Q(s)$ of F is the piecewise function

$$Q(s) := \begin{cases} \inf\{x \in \mathbf{R} : F(x) \geq s\} & 0 < s \leq 1 \\ \lim_{s\to 0^+} Q(s) & s = 0 \end{cases}$$

Let $H(s) := -Q(1-s)$. We can now study the properties of the maximum and minimum order statistics as n grows large:

Theorem 1.2.10 (Extreme Order Statistics)**.** *For any fixed $0 < \delta < 1$, we have as n grows large*

$$\max_{S\subset[n]:|S|=\lfloor\delta n\rfloor} ||\mathbf{v}||_S^2 \to -\int_0^\delta H(u)du$$

and

$$\min_{S\subset[n]:|S|=\lfloor\delta n\rfloor} ||\mathbf{v}||_S^2 \to -\int_{1-\delta}^1 H(u)du.$$

We can also show the extent to which the maximum and minimum order statistics are concentrated around their mean:

Theorem 1.2.11 (Concentration Result for Extreme Order Statistics). *For any* $1 \leq m \leq n$ *and* $t \geq 0$*, we have*

$$\mathbf{P}(\big|\max_{S \subset [n]:|S|=m} ||\mathbf{v}||_S - \mathbf{E} \max_{S \subset [n]:|S|=m} ||\mathbf{v}||_S\big| > t) \leq Ce^{-cnt^2}$$

and

$$\mathbf{P}(\big|\min_{S \subset [n]:|S|=m} ||\mathbf{v}||_S - \mathbf{E} \min_{S \subset [n]:|S|=m} ||\mathbf{v}||_S\big| > t) \leq Ce^{-cnt^2}$$

where $C, c > 0$ *are absolute constants.*

1.2.3 Reduction to Gaussian Vectors

Here we arrive at a nice result that will motivate our later work. A useful property when working with the GOE ensemble is the rotational invariance that results. In using this, we can show that it is possible to model the distribution of uniform vectors on the unit sphere via the multivariate distribution of standard Gaussian random variables:

Theorem 1.2.12. *Let* $\mathbf{x}$ *be a random vector uniformly distributed on the unit sphere* S^{n-1}*, where* $n \to \infty$*. Then* $\mathbf{x}$ *can be represented as*

$$\mathbf{v} := (\frac{\xi_1}{S}, \dots, \frac{\xi_n}{S}),$$

where ξ_i are iid standard Gaussian and $S = \sqrt{\sum_{i=1}^{n} |\xi_i|^2}$

Proof. Let $\mathbf{X} = (\xi_1, \cdots, \xi_n)$ be a vector consisting of iid standard Gaussian random variables, and let Q be an orthogonal matrix. It is clear that $Q\mathbf{X}$ has the same distribution as $\mathbf{X}$ and $\mathbf{X}$ is rotationally invariant. Consider $\mathbf{Y} := \frac{\mathbf{X}}{||\mathbf{X}||}$. Then by rotational invariance of $\mathbf{X}$, $\mathbf{Y}$ is rotational invariant as well, and $||\mathbf{Y}|| = 1$. So $\mathbf{Y}$ is uniformly distributed on the sphere. □

Chapter 2

Eigenvectors in a Symmetric Regime

Some content from this chapter, as well as from Chapter 3 and Chapter 4, has been published in [13] but will motivate our future work.

2.1 Motivation

Let $\mathbf{x}$ be a random vector uniformly distributed on the unit sphere S^{n-1}, where $n \to \infty$. As we have seen, $\mathbf{x}$ can be represented as

$$\mathbf{v} := (\frac{\xi_1}{S}, \dots, \frac{\xi_n}{S})$$

where ξ_i are iid standard Gaussian and $S = \sqrt{\sum_{i=1}^n |\xi_i|^2}$. One then can deduce that for any deterministic vector $\mathbf{f} = (f_1, \dots, f_n) \in \mathbf{R}^n$ with $\sum_i f_i^2 = n$,

$$\mathbf{f}^T \mathbf{v} \xrightarrow{d} \mathbf{N}(0, 1). \tag{2.1}$$

Let M_n be a random symmetric matrix of size $n \times n$ of real-valued entries. When M_n is GOE, we have seen that the individual eigenvectors of M_n have the same distribution as $\mathbf{x}$ above. Motivated by the *universality phenomenon*, it is natural to ask if the properties listed in the previous section are universal.

Question 2.1.1. *Is it true that the eigenvectors of M_n are "asymptotically uniformly distributed" for more general random ensembles M_n?*

We assume for the moment that M_n has simple spectrum. Let $\lambda_1 < \cdots < \lambda_n$ be the real eigenvalues of M_n, and $\mathbf{u}_1, \dots, \mathbf{u}_n$ be the corresponding unit eigenvectors (which are unique up to a sign). Among many nice results, the following can be read from [29, Theorem 13] and [1, Theorem 1.2] regarding Question 2.1.1 as it pertains to 2.1.

Theorem 2.1.2. *Let M_n be a random symmetric matrix where $m_{ij}, 1 \leq i \leq j \leq n$ are iid copies of a random variable ξ. Let $\mathbf{f} = (f_1, \dots, f_n) \in \mathbf{R}^n$ be any deterministic vector with $\sum_i f_i^2 = n$.*

- *[29] Assume that ξ is symmetric, $\xi \overset{d}{=} -\xi$, and ξ has moment matching up to the fourth order with $\mathbf{N}(0,1)$. Then for any $1 \leq i \leq n$,*

$$\mathbf{f}^T \mathbf{u}_i \xrightarrow{d} \mathbf{N}(0,1).$$

More precisely, there exists a positive constant c such that for any $x > 0$,

$$\mathbf{P}(|\mathbf{f}^T \mathbf{u}_i| \leq x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt + O(n^{-c}). \tag{2.2}$$

- *[1] Assume that ξ has mean zero, variance one, and ξ has finite moment of all orders. Then (2.2) holds for any eigenvector $\mathbf{u}_i$ with $i \in [1, n^{1/4}] \cup [n^{1-\delta}, n - n^{1-\delta}] \cup [n - n^{1/4}, n]$, with possibly different c.*

We also refer the readers to [29] and [1] for further beautiful results such as the joint independence and gaussianity of the eigenvectors.

Note that the constants c above can be made explicit but are rather small in both results. Thus, assume that if we are interested in the tail bound estimates $|\mathbf{f}^T \mathbf{u}_i| \le x$, then the above results are less effective when $x \ll n^{-c}$. In fact, it was not even known whether asymptotically almost surely $\mathbf{f}^T \mathbf{u}_i \neq 0$. This question was raised by Dekel, Lee and Linial in [2] for $\mathbf{f} = (1, 0, \dots, 0)$ in connection to the notion of strong and weak nodal domains in the random graph $G(n,p)$. This question has been confirmed in [15] in the following form.

Theorem 2.1.3. *Assume that F_n is a symmetric matrix with $\|F_n\|_2 \le n^\gamma$ for some constant $\gamma > 0$. Consider the matrix $M_n + F_n$ with the random symmetric matrix M_n of entries $m_{ij}, 1 \le i < j \le n$, being iid copies of a random variable ξ of mean zero, variance one, and bounded $(2+\varepsilon)$-moment for given $\varepsilon > 0$. Then for any A, there exists B depending on A and γ, ε such that*

$$\mathbf{P}\Big(\exists \text{ a unit eigenvector } \mathbf{u} = (u_1, \dots, u_n) \text{ of } M_n \text{ with } |u_i| \le n^{-B} \text{ for some } i\Big) = O(n^{-A}).$$

Although the above result holds for very general matrices, the approach does not seem to extend to the case that $\mathbf{f}$ has many non-zero entries, which is the main focus of this section.

Condition 2.1.1. *Let c, K_1, K_2 be positive parameters.*

- *(assumption for $\mathbf{f}$) We assume that the following holds for all but cn indices $1 \leq i \leq n$*
$$n^{-c} \leq |f_i| \leq n^c.$$

- *(assumption for M_n) We assume that the entries of $m_{ij}, 1 \leq i \leq j \leq n$, are iid copies of a random variable ξ of mean zero, variance one, and so that*
  - *For every $t > 0$,*
$$\mathbf{P}(|\xi| \geq t) \leq K_1 \exp(-t^2/K_2) \tag{2.3}$$
  - *ξ is symmetric.*

For the rest of this section we will be conditioning on the following result.

Theorem 2.1.4. *[28, 15] With M_n as above, there exists a constant $c > 0$ such that with probability at least $1 - \exp(-n^c)$, M_n has simple spectrum.*

In the above setting, we are able to prove the following, which is one of the main results of this dissertation:

Theorem 2.1.5 (Main result - Asymptotic inner product)**.** *Let M_n and $\mathbf{f}$ be as in Condition 2.1.1 for some positive constants K_1, K_2, and for some sufficiently small constant c. Conditioning on the event of Theorem 2.1.4, let $\lambda_1 < \cdots < \lambda_n$ be the eigenvalues of M_n and $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be the associated eigenvectors. Then the following* *holds for any* $\delta \geq \exp(-n^c)$

$$\mathbf{P}\Big(\sup_i |\langle \mathbf{u}_i, \mathbf{f}\rangle| \leq \delta\Big) \leq n^c\delta.$$

We will prove this Theorem in Chapter 4. In what follows, we discuss an application.

2.2 Matrix Controllability

Here we connect our result to the study of controllability of matrices. Consider the discrete-time linear state-space system whose state equation is

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{u}(k).$$

In the above, A and B are $n \times n$ and $n \times r$ matrices, respectively, while each $\mathbf{u}(k)$ is an $r \times 1$ vector that we wish to solve for based on the state values $\mathbf{x}(k)$ of size $n \times 1$.

We say that our system is controllable if we can always find the control values $\mathbf{u}(n-1), \mathbf{u}(n-2), \cdots, \mathbf{u}(0)$ based on the state values $\mathbf{x}(\cdot)$. Note that

$$\begin{aligned}
\mathbf{x}(1) &= A\mathbf{x}(0) + B\mathbf{u}(0)\\
\mathbf{x}(2) &= A\mathbf{x}(1) + B\mathbf{u}(1) = A^2\mathbf{x}(0) + AB\mathbf{u}(0) + B\mathbf{u}(1)\\
&\vdots\\
\mathbf{x}(n) &= A^n\mathbf{x}(0) + A^{n-1}B\mathbf{u}(0) + A^{n-2}B\mathbf{u}(1) + \cdots + AB\mathbf{u}(n-2) + B\mathbf{u}(n-1).
\end{aligned}$$

That is,

$$\mathbf{x}(n) - A^n\mathbf{x}(0) = (A^{n-1}B \; A^{n-2}B \; \cdots \; AB \; B)(\mathbf{u}^T(0) \; \mathbf{u}^T(1) \; \cdots \; \mathbf{u}^T(n-1))^T.$$

From here it is easy to see that we can always find the control values $\mathbf{u}(\cdot)$ if and only if the left matrix has full rank. Restricting to the case where $r = 1$ and switching around columns to remain consistent with other literature, this motivates the following definition of controllability:

Definition 2.2.1. Let A be an $n \times n$ matrix and let $\mathbf{b}$ be a vector in $\mathbf{R}^n$. We say that the pair $(A, \mathbf{b})$ is controllable if the $n \times n$ column matrix

$$(\mathbf{b} \quad A\mathbf{b} \quad \cdots \quad A^{n-1}\mathbf{b})$$

has full rank.

As it turns out, the notion of controllability is related to the existence of eigenvectors orthogonal to $\mathbf{b}$ via the Popov-Belevitch-Hautus test [19]:

Theorem 2.2.2. *With A and $\mathbf{b}$ as above, $(A, \mathbf{b})$ is uncontrollable if and only if there exists an eigenvector $\mathbf{v}$ of A such that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$.*

This is [20, Lemma 1], which we prove in Appendix A.

Recent developments in the area of matrix controllability have come from imposing randomness on the matrix A and imposing varying rigidity on the deterministic vector $\mathbf{b}$. For example, O'Rourke and Touri in [20] were able to prove the following conjecture of Godsil.

Conjecture 2.2.3. *Let $\mathbf{1}_n$ be the vector in $\mathbf{R}^n$ consisting of all 1's and A_n be the adjacency matrix of $G(n, 1/2)$. Then as n approaches infinity, $(A_n, \mathbf{1}_n)$ is controllable asymptotically almost surely.*

This has been verified recently by O'Rourke and Touri in stronger form through their focus on $(K, \delta)-$delocalized vectors; we say that $\mathbf{v}$ is $(K, \delta)-$delocalized if most of the entries of $\mathbf{v}$ are non-zero rational numbers of bounded height (a precise definition is given in [20]).

Through this notion, the authors of [19, 20] were able to prove Godsil's conjecture by the following theorem.

Theorem 2.2.4. *[19, Theorem 3.4] Assume that M_n is a random symmetric matrix where the off-diagonal entries $m_{ij}, 1 \leq i < j \leq n$, are iid copies of ξ as in Theorem 2.1.5, while the diagonal entries are iid copies of a possibly different subgaussian random variable ζ. Fix $K \geq 1$ and $\alpha > 0$. Then there exist constants $C > 0$ and $\delta \in (0, 1)$ (depending on K, α, ξ, and ζ) such that the following holds. Let $\mathbf{b}$ be a vector in $\mathbf{R}^n$ which is $(K, \delta)-$delocalized. Then $(M_n, \mathbf{b})$ is controllable with probability at least $1 - Cn^{-\alpha}$.*

Our result, Theorem 2.1.5, can be seen as a near optimal generalization of Theorem 2.2.4 (in the case that m_{ii} and m_{ij} have the same distribution) where the entries of $\mathbf{f}$ are not necessarily rational.

Chapter 3

Small Ball Probability

In this chapter, we introduce a few ingredients to prove Theorem 2.1.5. As we can see, we must deal with $\mathbf{P}(|\langle \mathbf{u}, \mathbf{f} \rangle| < \delta)$, where $\mathbf{u} = (u_1, \cdots, u_n)$ is an eigenvector and $\mathbf{f} = (f_1, \cdots, f_n)$ is a fixed, deterministic vector. To start, we will focus on a more elementary problem, the concentration probability of random walks with discrete random variables. This concentration probability will also be useful when we work with matrices over $\mathbf{F}_p$ in subsequent chapters. Broadly speaking, we will ultimately have to deal with the small ball probability (defined explicitly in the next section) where $u_1, \cdots, u_n$ are random but not necessarily independent.

3.1 Littlewood-Offord Theory

3.1.1 Intro

In this section, we will deal with the case where the u_i are independent and identically distributed. This falls into the concept of Littlewood-Offord theory.

Let ζ be a real random variable with mean zero and variance one. Also, let $A = \{a_1, \cdots, a_n\}$ be a deterministic multi-set in $\mathbf{R}^d$. We will isolate our study to the instance where our dimension d is simply 1. In this section, we analyze the random weighted sum

$$S_A := a_1\zeta_1 + \cdots a_n\zeta_n,$$

where the ζ_i are independent and identically distributed copies of ζ. Note that our A plays the role of $\mathbf{f}$ from before.

In the instance that all of the a_i are the same (and nonzero), we can look at the Lyapunov Central Limit Theorem to analyze S_A. Treating S_A as a sum of independent random variables $a_i\zeta_i$, we have that

$$\frac{S_A}{\sqrt{\sum |a_i|^2}} \to \mathbf{N}(0,1).$$

If the random variable ζ has bounded third moment, then under the above condition on the a_i, the Berry-Esseen theorem tells us that the rate of convergence is $O(n^{-1/2})$, and this is in fact true for any small open interval I; that is to say,

$$\mathbf{P}(S_A \in I) = O(|I|/n^{1/2}).$$

Definition 3.1.1. Given our distribution ζ, multiset A, and open interval I, we will refer to $\mathbf{P}(S_A \in I)$ as the small ball probability.

Definition 3.1.2. We will refer to $\rho(A) := \sup_{x\in\mathbf{R}} \mathbf{P}(S_A = x)$ as the concentration probability.

It is natural to now look at the more complicated scenario where the a_i are not identical. In [9], Littlewood and Offord, under the popular assumptions that ζ is Bernoulli ± 1 and the a_i have magnitude at least 1, were able to show the following:

Theorem 3.1.3. *Suppose A is a multiset with each $|a_i| \geq 1$. Then for any open interval I of length 2, we have*

$$\mathbf{P}(S_A \in I) = O(\frac{\log n}{n^{1/2}}).$$

The above was one of the first results where the a_i were allowed to be arbitrary. After coming across the previous Littlewood-Offord statement, Erdős [3] was able to provide a beautiful proof of the following refinement:

Theorem 3.1.4. *Under the above assumptions, we have*

$$\mathbf{P}(S_A \in I) \leq \frac{\binom{n}{\lfloor n/2 \rfloor}}{2^n} = O(\frac{1}{n^{1/2}}).$$

Proof. Let x be a fixed real number. Without loss of generality, we assume that each $a_i \geq 1$. Consider $\mathcal{F}$, the set of all subsets X of $\{1, 2, \cdots, n\}$ such that

$$(\sum_{i \in X} a_i - \sum_{j \in [n] \setminus X} a_j) \in (x-1, x+1).$$

Clearly, $\mathcal{F}$ is an antichain. Thus by Sperner's Lemma, $|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$ and

$$\mathbf{P}(S_A \in I) \leq \frac{\binom{n}{\lfloor n/2 \rfloor}}{2^n}.$$

Since

$$\binom{n}{\lfloor n/2 \rfloor} \sim O(\frac{n^n}{(n/2)^n \sqrt{n}})$$

via Stirling's approximation, the result follows. □

This beautiful combinatorial result by Erdős is typically regarded as the forefront of Littlewood-Offord theory.

3.1.2 A Nice Hierarchy of Bounds

We can improve the result of Theorem 3.1.4 by imposing additional conditions on the members of the set A. Erdős and Moser [4] were able to show that if we assume that all of the a_i are distinct, we arrive at a significantly better bound:

Theorem 3.1.5. *Suppose each a_i is distinct. Then*

$$\sup_{x \in \mathbf{R}} \mathbf{P}(S_A = x) = O(n^{-3/2} \log n).$$

Mirroring the previous result for the bound of $O(n^{-1/2} \log n)$, it's natural to wonder if the $\log n$ term is necessary in this instance. Sarkozy and Szemerédi [24] were able to prove that it is not:

Theorem 3.1.6. *Suppose each a_i is distinct. Then*

$$\sup_{x \in \mathbf{R}} \mathbf{P}(S_A = x) = O(n^{-3/2}).$$

Theorem 3.1.6 shows us that we can achieve a better bound on Theorem 3.1.4 by forbidding the criterion that gave us sharpness in the bound of $O(n^{-1/2})$. Since the bound $O(n^{-3/2})$ of Theorem 3.1.6 is sharp when the a_i form an arithmetic progression, we can then ask what would happen if we were to forbid this from occurring. Using a more general result proved by Halász, it is possible to show the following (see [26]):

Theorem 3.1.7. *Let A be a multiset in $\mathbf{R}_{>0}$. Let l be a fixed integer and let R_l be the number of solutions of*

$$a_{i_1} + \cdots a_{i_l} = a_{j_i} + \cdots + a_{j_l}.$$

Then

$$\sup_{x \in \mathbf{R}} \mathbf{P}(S_A = x) = O(n^{-2l-1/2} R_l).$$

We can rephrase the results of this section more intuitively as a function of the structure of the set A. In the instance of each a_i being identical and the sharpness of the bound $O(n^{-1/2})$ being achieved, we view this as the set A having structure - lots of it, in fact. When the a_i form an arithmetic progression, then the sharpness of our bound lessens to $O(n^{-3/2})$; the set A has slightly less structure now. We can generalize the notion of an arithmetic progression and the structure that results, something we will do in a later section.

3.2 Inverse Littlewood-Offord Theory

What if, instead of bounding the small ball probability, we were to find the reason as to why the small ball probability might be large? This is the methodology of inverse Littlewood-Offord theory, a notion developed by Van Vu and Terence Tao. More specifically, let's suppose that there exists a constant C such that $\rho(V) \geq n^{-C}$ for each sufficiently large n. From the previous section, we expect the elements $v_1, \cdots, v_n$ of V to have some sort of additive structure, depending on C.

To see this at work, write

$$S_V = (v_1 + \cdots + v_n) - 2 \sum_{i:\epsilon_i=-1} v_i.$$

Then if

$$\rho(V) = \sup_{x\in\mathbf{R}} \mathbf{P}(S_V = x) = \sup_{x\in\mathbf{R}} \mathbf{P}(\frac{v_1 + \cdots + v_n - x}{2} = \sum_{i:\epsilon_i=-1} v_i) \geq n^{-C},$$

we see that at least $2^n/n^C$ of the subsets of $\{1, \cdots, n\}$ attain the value $\frac{v_1+\cdots+v_n-x}{2}$, which certainly implies a strong additive structure.

We seek to in some sense quantify the additive structure that V may contain. Prior to that, we will introduce Freiman's inverse theorem. Freiman [7] considered the size of sets of the form $V + V := \{v_i + v_j : v_i, v_j \in V\}$. This set will have at most $|V|^2$ elements. Freiman showed that if $|V + V|$ is of size $O(|V|)$, then there is a lot of collision and V resembles an arithmetic progression. If V is an arithmetic progression, then note that $|V + V|$ is roughly $2V$. Before stating Freiman's theorem, we will state the definition of generalized arithmetic progressions (GAPs):

Definition 3.2.1. Let Q be a subset of an abelian group Z. Then Q is a generalized arithmetic progression (GAP) of rank r if it can be expressed as

$$Q = \{g_0 + m_1 g_1 + \cdots + m_r g_r : M_i \leq m_i \leq M_i', m_i \in \mathbf{Z}\}$$

for some $g_0, \cdots, g_r \in Z$ and $M_1, M_1', \cdots, M_r, M_r' \in \mathbf{R}$.

We will often call $g_0, \cdots, g_r$ the generators of Q, M_i and M_i' the dimensions of Q. Thinking of Q as the image of $B := \{(m_1, \cdots, m_r) \in \mathbf{Z}^r : M_i \leq m_i \leq M_i'\}$ under the map $\phi : (m_1, \cdots, m_r) \to g_0 + m_1 g_1 + \cdots + m_r g_r$, we will define the volume of Q to be Vol$(Q) := |B|$. Q is said to be proper if $|Q| =$Vol(Q), i.e. if our map is injective. We will refer to Q as symmetric if $M_i = -M_i'$ for each i and $g_0 = 0$. Finally, given a symmetric GAP Q, we define the dilate tQ, where $t > 0$, to be the set

$$tQ := \{g_0 + m_1 g_1 + \cdots + m_r g_r : -tM_i' \leq m_i \leq tM_i\}.$$

Note that if Q is a proper GAP of rank r, then $|Q + Q| \leq 2^r |Q|$. So if A if a subset of a proper rank r GAP Q with density $\delta = \Theta(1)$, we have that $|A + A| \leq |Q + Q| \leq 2^r |Q| \leq 2^r (\frac{|A|}{\delta}) = O(|A|)$, i.e. dense subsets of a proper GAP satisfy $|A+A| = O(|A|)$. In fact, Freiman [7] showed that this is the only way that the bound $|A+A| = O(|A|)$ occurs:

Theorem 3.2.2 (Freiman's inverse theorem in $\mathbf{Z}$). *Let $C > 0$ be a constant, and suppose X is a subset of $\mathbf{Z}$ such that $|X + X| \leq C|X|$. Then we can find a proper GAP Q of rank $O_C(1)$ and cardinality $O_C(|X|)$ such that Q contains X.*

Let Q be a proper symmetric GAP of rank r and volume N, and let $v_1, \cdots, v_n$ be elements of Q. By the central limit theorem, the random Bernoulli sum $S_V = v_1\epsilon_1 + \cdots + v_n\epsilon_n$ takes values in $10n^{1/2}Q$ with probability at least 2/3. Note that $|tQ| \leq t^r N$; hence, we can find a point x where

$$\mathbf{P}(S_V = x) = \Omega(\frac{1}{n^{r/2}N}).$$

So if $|Q| = N = O(n^{C-r/2})$ for some $C \geq r/2$, we have

$$\rho(V) \geq \mathbf{P}(S_V = x) = \Omega(\frac{1}{n^C}).$$

What this tells us is if V's members are elements of a symmetric proper GAP with small rank and small cardinality, then $\rho(V)$ is large. Turns out, we also have the converse [30]:

Theorem 3.2.3. *Let C, ϵ be given. Then there are constants r, B such that the following holds: Let V be a multi-set of n real numbers such that $p(V) \geq n^{-C}$. Then there is a GAP Q of rank r and volume n^B such that all but n^ϵ elements of V belong to Q.*

While Tao and Vu achieved an almost sharp dependence of B on r, C in the previous theorem, it was later improved, leading to the optimal inverse Littlewood-Offord theorem [16]:

Theorem 3.2.4 (Optimal inverse Littlewood-Offord theorem, discrete case)**.** *Let $\epsilon < 1$ and C be positive constants. Assume that*

$$\rho(V) \geq n^{-C}.$$

Then for any $n^{\epsilon} \leq n' \leq n$, we can find a proper symmetric GAP Q of rank $r = O_{C,\epsilon}(1)$ which contains all but at most n' elements of V (counting multiplicities), where

$$|Q| = O_{C,\epsilon}(\rho(V)^{-1} n'^{-\frac{r}{2}}).$$

In particular, there exists a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and cardinality $O_{C,\epsilon}(p(V)^{-1} n^{-\frac{r}{2}})$ which contains all but at most ϵn elements of V (counting multiplicities).

As an application of the optimal inverse Littlewood-Offord theorem, one can prove a counting theorem (see [17]) that serves to be very helpful in our subject:

Theorem 3.2.5 (Counting theorem, discrete case)**.** *Let C_1, C_2 be given. The number N of multisets A of integers satisfying $\max |a_i| \leq n^{C_1}$ and $\rho(A) \geq n^{-C_2}$ is bounded by*

$$N = (O_{C_1,C_2,\epsilon}(1))^n n^{O_\epsilon(1)n} (\rho(A)^{-1} n^{-1/2})^n,$$

where ϵ is some constant $0 < \epsilon < 1$.

While the inverse theorems are hard to apply directly, using the counting theorem allows us to bound the probability of certain bad events occurring via a union bound.

The heuristic remains as follows: the larger the small ball probability, the larger the structure of the set A.

As it will be relevant to our interests in future sections, we also have the inverse Littlewood-Offord theorem in a mod p setting. Given a random variable μ, prime integer p and $\alpha > 0$, we will say that μ is $\alpha-$balanced if we have

$$\max_{a \in \mathbf{F}_p} \mathbf{P}(\mu = a \mod p) \leq 1 - \alpha.$$

This allows us to limit the types of entries that our random matrix may take. In this setting, we can show the following:

Theorem 3.2.6 (Inverse Littlewood Offord mod p)**.** *Let $\epsilon < 1$ and C be positive constants, and let p be a prime integer. Suppose μ is a random variable that is α_n-balanced taking values in $\mathbf{Z}/p\mathbf{Z}$. Also, assume that $\mathbf{w} = (w_1, \cdots, w_n) \in (\mathbf{Z}/p\mathbf{Z})^n$ is such that*

$$|\rho(\mathbf{w}) - \frac{1}{p}| = \sup_{a \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{P}(\mu_1 w_1 + \cdots + \mu_n w_n = a) - \frac{1}{p}| \geq n^{-C},$$

where $\mu_1, \cdots, \mu_n$ are independent and identically distributed copies of μ. Then for any $n^{\epsilon/2} \leq n' \leq n$, there is a set W' of $n - n'$ components w_i such that one of the following holds.

- *If $p \ll n^C$, then there exists a GAP of rank one Q that contains W', where*

$$|Q| \leq O_{C,\varepsilon}(p\sqrt{(\log n)/n'}).$$

- *If $p \gg n^C$, there exists a proper symmetric GAP Q of rank $r = O_{C,\epsilon}(1)$ that contains W', where*

$$|Q| \leq \max\{1, O_{C,\epsilon}(\rho^{-1}/n'^{r/2})\}.$$

At this point, we refer the reader to [18], in which Nguyen and Wood proved a very similar theorem. We note that their statement differed in two key areas. First, they worked under the assumption that $\rho(\mathbf{w}) \geq n^{-C}$, whereas we have $|\rho(\mathbf{w}) - 1/p| \geq n^{-C}$. Second, their proof was for prime integers larger than $C'n^C$ for sufficiently large $C'(C, \varepsilon)$. In Appendix B we replicate their proof almost verbatim, making the required modifications to prove our form of the statement. This will appear in [10].

3.3 Rudelson-Vershynin LCD and Applications

3.3.1 LCD as a Measure of Structure

Building off of the previous section, we can define a new notion of structure among vectors, inspired by Rudelson and Vershynin. The Rudelson-Vershynin *least common denominator* (see [22]) is a quantity that will be large when our vector $\mathbf{v}$ has small amounts of structure, and small when $\mathbf{v}$ has a lot of structure. In fact, we will be able to bound the previously discussed concentration probability by the inverse of the least common denominator.

Fix parameters κ and γ (which may depend on n) with $\gamma \in (0, 1)$.

Definition 3.3.1 (Rudelson-Vershynin **LCD**)**.** Let $\mathbf{x}$ be a nonzero vector. We define the Rudelson-Vershynin **LCD** of $\mathbf{x}$ to be the quantity

$$\mathbf{LCD}_{\kappa,\gamma}(\mathbf{x}) := \inf \Big\{ \theta > 0 : \operatorname{dist}(\theta\mathbf{x}, \mathbf{Z}^n) < \min(\gamma\|\theta\mathbf{x}\|, \kappa) \Big\}.$$

Because the distance is smaller than $\gamma\|\theta\mathbf{x}\|$, we will be considering non-trivial integer points near $\theta\mathbf{x}$. The inequality $\operatorname{dist}(\theta\mathbf{x}, \mathbf{Z}^n) < \kappa$ then allows us to assume that most coordinates of $\theta\mathbf{x}$ are close to non-trivial integers.

We can use the **LCD** to provide an upper bound on the small ball probability.

Theorem 3.3.2 (Small ball probability via **LCD**)**.** *Let ξ be a sub-gaussian random variable of mean zero and variance one, and let $\xi_1, \dots, \xi_n$ be iid copies of ξ. Consider a vector $\mathbf{x} \in \mathbf{R}^n$. Then, for every $\kappa > 0$ and $\gamma \in (0,1)$, and for*

$$\varepsilon \geq \frac{1}{\mathbf{LCD}_{\kappa,\gamma}(\mathbf{x}/\|\mathbf{x}\|)},$$

we have

$$\rho_\varepsilon(\mathbf{x}) = O\left(\frac{\varepsilon}{\gamma\|\mathbf{x}\|} + e^{-\Theta(\kappa^2)} \right),$$

where the implied constants depend on ξ.

We will sketch the proof, as this is an important statement but the proof is largely analytical. The full details can be found in [22].

Sketch of proof of Theorem 3.3.2. We can first appeal to Esseen's Inequality, which bounds the small ball probability of a random variable S by the L_1-norm of its chracteristic function $\phi(t) = \phi_S(t) = \mathbf{E}[\exp{(iSt)}]$:

$$\sup_{v \in \mathbf{R}} \mathbf{P}(|S - v| \leq \epsilon) \leq C \int_{-\pi/2}^{\pi/2} |\phi(t/\epsilon)| dt.$$

We can apply Esseen's Inequality to the random sum S whose characteristic function is

$$\phi(t) = \prod_{k=1}^{n} \phi_k(t),$$

where each $\phi_k(t) = \mathbf{E}[\exp{(ix_k \xi t)}]$. Using some Taylor Series estimates, we can estimate the integral in Esseen's Lemma as

$$\begin{aligned} p_\epsilon(\mathbf{x}) &\leq C \int_{-\pi/2}^{\pi/2} |\phi(t/\epsilon)| dt \\ &\leq C \sup_{z \geq 1} \int_{-\pi/2}^{\pi/2} \exp{(-\beta f(zt/\epsilon))} \end{aligned}$$

where

$$f(t) := \sum_{k=1}^{n} \sin^2{(\frac{1}{2} x_k t)}$$

and β is a fixed positive constant.

Consider the maximum $M := \max_{|t| \leq \pi/2} f(zt/\epsilon)$. One can show, via an averaging argument, that

$$\frac{n}{4} \leq M \leq n.$$

Define the following level sets of f for $m, r \geq 0$:

$$T(m,r) := \{t : |t| \leq r, f(zt/\epsilon) \leq m\}.$$

We can appeal to a lemma of Halász to show that the Lebesgue measure of these level sets behaves in a regular way:

Lemma 3.3.3. *Let l be such that $l^2 m \leq M$, where M is the max from before. Then*

$$T(m, \pi/2) \leq \frac{2}{l}|T(l^2 m, \pi)|.$$

It thus makes sense to decompose our integral into two classes, one where m is small and the other where m is large. We have

$$\begin{aligned}
p_\epsilon(a) &\leq C \sup_{z\geq 1} \int_{-\pi/2}^{\pi/2} \exp\left(-\beta f(zt/\epsilon)\right) \\
&\leq C \int_0^\infty |T(m,\pi/2)|\beta e^{-\beta m} dm \\
&\leq C \int_0^{m_1 M} 4\sqrt{\frac{m}{m_1 M}}|T(m_1 M, \pi)|\beta e^{-\beta m} dm + C \int_{m_1 M}^\infty \pi\beta e^{-\beta m} dm \\
&\leq \frac{C_2 B}{\sqrt{m_1 n}}|T(m_1 n, \pi)| + C\pi e^{-c_2 m_1 n/B^2}.
\end{aligned}$$

We now seek to bound the measure of the level sets $|T(m_1 n, \pi)|$. To do so, we will look at the recurrence set of $\mathbf{x}$. Let $t \in T(m_1 n, \pi)$ and consider $y := z/2\epsilon$. Then

$$f(zt/\epsilon) = \sum_{k=1}^n \sin^2(x_k y t) \leq m_1 n.$$

Fix $m_1 := \frac{\alpha^2 \kappa}{4n}$. Then at least $n - \kappa$ terms in the sum satisfy

$$\sin^2(x_k y t) \leq \frac{m_1 n}{\kappa} = \frac{\alpha^2}{4},$$

in which case those terms follow $\text{dist}(x_y t, \pi \mathbf{Z}) \leq \alpha$. This motivates the following definition:

Definition 3.3.4. Let $\alpha \in (0, 1)$ and $\kappa \geq 0$. We define the recurrence set $I(\mathbf{x})$ of a vector $\mathbf{x} \in \mathbf{R}^n$ to be the set of all $t \in \mathbf{R}$ such that all except κ coordinates of $t\mathbf{x}$ are within distance at most α of $\mathbf{Z}$.

From above, we see that yt/π is in the recurrence set of $\mathbf{x}$ and so $T(m_1 n, \pi) \subset \frac{\pi}{y} I(\mathbf{x})$. We can also define the density of the set I as follows:

Definition 3.3.5. The density of the set I relative to y is $\text{dens}(I, y) := \frac{1}{2y}|I \cap [-y, y].|$

We have actually shown that $|T(m_1 n, \pi)| \leq 2\pi \ \text{dens}(I(\mathbf{x}), y)$. So we now consider the density of $I(\mathbf{x})$ relative to y, and it suffices to show that this density is bounded above by the reciprocal of the **LCD**. We can show that the recurrence set has a lot of gaps, and each gap is bounded below by the **LCD** of $\mathbf{x}$:

Lemma 3.3.6. *Let $t_0 \in I(\mathbf{x})$. Then $t_0 + 3\alpha \notin I(\mathbf{x})$. Furthermore, if $t_1 \in I(\mathbf{x})$ such that $t_1 > t_1 + 3\alpha$, then we have $t_1 - t_0 \geq D_{2\alpha, 2\kappa}(\mathbf{x})$.*

Using this lemma, it is possible to bound the density of the recurrence set:

Lemma 3.3.7. *For every $y > 0$, we have dens$(I(\mathbf{x}), y) \leq 3\alpha(\frac{1}{2y} + \frac{2}{D_{2\alpha, 2\kappa}(\mathbf{x})})$.*

This is enough to complete the proof. □

As an application of the Rudelson-Vershynin least common denominator, we can look to approximate eigenvectors in a specific setting of symmetric Wigner matrices. For the remainder of this chapter, we will assume that M_n is a symmetric matrix whose entries are independent and identically distributed subgaussian random variables with mean zero and variance one. We also assume that these random variables are symmetric. Note that we are now working in Condition 2.1.1, which will eventually help us prove Theorem 2.1.5 in Chapter 4. First, we can show that approximate eigenvectors $\mathbf{v}$ are rarely sparse. Then, under the condition that $\mathbf{v}$ is not sparse, we will use the **LCD** to show that $\mathbf{v}$ does not contain structure.

3.3.2 Approximate Eigenvectors are not Asymptotically Sparse

Let us condition on the following event, which (see [25], for instance) is known to hold with probability $1 - \exp(-\Theta(n))$:

$$\|M_n\| \leq 10\sqrt{n}. \tag{3.1}$$

We begin by introducing the definition of compressible and incompressible vectors.

Definition 3.3.8. Let $c_0, c_1 \in (0, 1)$ be two numbers (chosen depending on the parameters K_1, K_2, K_1', K_2' of ξ, ζ) A vector $\mathbf{x} \in \mathbf{R}^n$ is called *sparse* if $|\mathbf{supp}(\mathbf{x})| \leq c_0 n$. A vector $\mathbf{x} \in S^{n-1}$ is called *compressible* if $\mathbf{x}$ is within Euclidean distance c_1 from the set of all sparse vectors. A vector $\mathbf{x} \in S^{n-1}$ is called *incompressible* if it is not compressible.

The sets of compressible and incompressible vectors in S^{n-1} will be denoted by $\mathbf{Comp}(c_0, c_1)$ and $\mathbf{Incomp}(c_0, c_1)$ respectively.

Regarding the behavior of $M_n\mathbf{x}$ for compressible vectors, the following was proved in [31]:

Lemma 3.3.9. *[31, Proposition 4.2] There exist positive constants c_0, c_1 and α_0 (depending on K_1, K_2 of ξ) such that the following holds for any λ_0 of order $O(\sqrt{n})$. For any fixed $\mathbf{u} \in \mathbf{R}^n$ one has*

$$\mathbf{P}(\inf_{\mathbf{x}\in\mathbf{Comp}(c_0,c_1)} \|(M_n - \lambda_0)\mathbf{x} - \mathbf{u}\| \ll \sqrt{n}) = O(\exp(-\alpha_0 n)).$$

We deduce the following immediate consequence:

Lemma 3.3.10 (Approximate eigenvectors are not asymptotically sparse)**.** *There exist positive constants c_0, c_1 and α_0 (depending on K_1, K_2 of ξ) such that*

$$\mathbf{P}\left(\exists \text{ a unit vector } \mathbf{v} \in \mathbf{Comp}(c_0, c_1) \text{ and } \lambda = O(\sqrt{n}) : \|(M_n - \lambda)\mathbf{v}\| \ll \sqrt{n}\right)$$

is $O(\exp(-\alpha_0 n))$.

Proof. (of Lemma 3.3.10) Assuming (3.1), we can find λ_0 as a multiple of n^{-2} inside $[-10\sqrt{n}, 10\sqrt{n}]$ such that $|\lambda - \lambda_0| \le n^{-2}$. Hence

$$\|(M_n - \lambda_0)\mathbf{v}\| = \|(\lambda - \lambda_0)\mathbf{v}\| \le n^{-2}.$$

On the other hand, for each fixed λ_0, by Lemma 3.3.9,

$$\mathbf{P}(\exists \mathbf{v} \in \mathbf{Comp}(c_0, c_1) : \|(M_n - \lambda_0)\mathbf{v}\| \leq n^{-2}) = O(\exp(-\alpha_0 n)).$$

The claim follows by a union bound with respect to λ_0. □

3.3.3 Approximate Eigenvectors Cannot have Structure

One of the key properties of vectors of small **LCD** is that they accept a fine net of small cardinality (see [23, Lemma 4.7] and also [15, Lemma B6] for the current form).

Lemma 3.3.11. *Let $D_0 \geq c\sqrt{n}$. Then the set*

$$\{\mathbf{x} \in \mathbf{R}^n, \|\mathbf{x}\| \leq 1, c\sqrt{m} \leq \mathbf{LCD}_{\kappa,\gamma}(\mathbf{x}/\|\mathbf{x}\|) \leq D_0\}$$

has a $(2\kappa/D_0)$-net of cardinality at most $(C_0 D_0/\sqrt{m})^m D_0^2$ for some absolute constant C_0.

Unless otherwise noted, we will take $\gamma = 1/2$ and $\kappa = n^{2c}$ for some constant c chosen sufficiently small (compared to all other parameters).

To deal with symmetric or Hermitian Wigner matrices, it is more convenient to work with the so-called *regularized least common denominator.* Let $\mathbf{x} = (x_1, \ldots, x_n) \in S^{n-1}$. Let $c_0, c_1 \in (0, 1)$ be given constants, and assume $\mathbf{x} \in \mathbf{Incomp}(c_0, c_1)$. It is not hard to see that (see for instance [22, Section 3]) there are at least $c_0 c_1^2 n/2$ coordinates x_k of $\mathbf{x}$ which satisfy

$$\frac{c_1}{\sqrt{2n}} \leq |x_k| \leq \frac{1}{\sqrt{c_0 n}}. \tag{3.2}$$

Thus for every $\mathbf{x} \in \mathbf{Incomp}(c_0, c_1)$ we can assign a subset $\mathbf{spread}(\mathbf{x}) \subset [n]$ such that (3.2) holds for all $k \in \mathbf{spread}(\mathbf{x})$ and

$$|\mathbf{spread}(\mathbf{x})| = \lceil c'n \rceil,$$

where we set

$$c' := c_0 c_1^2/4. \tag{3.3}$$

Definition 3.3.12 (Regularized LCD, see also [31])**.** Let $\alpha \in (0, c'/4)$. We define the *regularized LCD* of a vector $\mathbf{x} \in \mathbf{Incomp}(c_0, c_1)$ as

$$\widehat{\mathbf{LCD}}_{\kappa,\gamma}(\mathbf{x}, \alpha) = \max\Big\{\mathbf{LCD}_{\kappa,\gamma}\big(\mathbf{x}_I/\|\mathbf{x}_I\|\big) : I \subseteq \mathbf{spread}(\mathbf{x}),\, |I| = \lceil \alpha n \rceil\Big\}.$$

Roughly speaking, the reason we choose to work with $\widehat{\mathbf{LCD}}$ is that we want to detect structure of $\mathbf{x}$ in sufficiently small segments. From the definition, it is clear that if $\mathbf{LCD}(\mathbf{x})$ is small (i.e. when $\mathbf{x}$ has strong structure), then so is $\widehat{\mathbf{LCD}}(\mathbf{x}, \alpha)$.

For given D, κ, γ and α, we denote the set of vectors of norm $1+o(1)$ with bounded regularized LCD by

$$T_{D,\kappa,\gamma,\alpha} := \{\mathbf{x} \in \mathbf{Incomp}(c_0, c_1) : \widehat{\mathbf{LCD}}_{\kappa,\gamma}(\mathbf{x}, \alpha) \leq D\}.$$

The following is [15, Lemma 5.9].

Lemma 3.3.13. *Assume that M_n is a random Wigner matrix with subgaussian entries. Then there exist $c > 0, \alpha_0 > 0$ depending on c_0, c_1 from Lemma 3.3.10 such* *that the following holds with $\kappa = n^{2c}$ and $\gamma = 1/2$. Let α, D be such that*

$$n^{-c} \le \alpha \le c'/4 \text{ and } 1 \le D \le n^{c/\alpha}.$$

Then for any fixed $\mathbf{u} \in \mathbf{R}^n$ and any real number λ_0 of order $O(\sqrt{n})$, with $\beta = \frac{\kappa}{\sqrt{\alpha D}}$ we have

$$\mathbf{P}\left(\exists \mathbf{x} \in T_{D,\kappa,\gamma,\alpha} : \|(M_n - \lambda_0)\mathbf{x} - \mathbf{u}\| = o(\beta\sqrt{n})\right) = O(\exp(-\alpha_0 n)).$$

We remark that, while Lemma 3.3.10 and Lemma 3.3.13 were proved for unit vectors $\mathbf{x}$, the proofs automatically extend to vectors of norm $1 \pm n^{-2c}$. For instance Lemma 3.3.13 can be extended to show that

$$\mathbf{P}\left(\exists \mathbf{x} : 1 - n^{-2c} \le \|\mathbf{x}\| \le 1 + n^{-2c} \wedge \mathbf{x}/\|\mathbf{x}\| \in T_{D,\kappa,\gamma,\alpha} : \|(M_n - \lambda_0)\mathbf{x} - \mathbf{u}\| = o(\beta\sqrt{n})\right)$$

is $O(\exp(-\alpha_0 n))$. Indeed, the event $\|(M_n - \lambda_0)\mathbf{x} - \mathbf{u}\| = o(\beta\sqrt{n})$ implies $\|(M_n - \lambda_0)\mathbf{x}/\|\mathbf{x}\| - \mathbf{u}/\|\mathbf{x}\|\| = o(\beta\sqrt{n})$, and the later implies that $\|(M_n - \lambda_0)\mathbf{x}/\|\mathbf{x}\| - \mathbf{u}_i\| = o(\beta\sqrt{n})$ for some deterministic $\mathbf{u}_i$ appropriately chosen to approximate $\mathbf{u}/\|\mathbf{x}\|$ with an error, say, at most β. As one can easily construct a set of size $n^{O(1)}/\beta$ for the $\mathbf{u}_i$'s, taking union bound over these approximating points will not dramatically change the exponential bound $O(\exp(-\alpha_0 n))$ of the right hand side of Lemma 3.3.13 as $\beta \ge \exp(-n^c)$.

We deduce the following crucial consequence from Lemma 3.3.10 and Lemma 3.3.13.

Corollary 3.3.14. *Let* $\mathbf{u} \in \mathbf{R}^n, \lambda_0$ *be fixed, and* D, β *be as above. Let* $\mathcal{E}_{\mathbf{u},\lambda_0}$ *be the event that for any* $\mathbf{x}$ *with* $1 - n^{-2c} \leq \|\mathbf{x}\| \leq 1 + n^{-2c}$ *, if* $\|(M_n - \lambda_0)\mathbf{x} - \mathbf{u}\| = o(\beta\sqrt{n})$ *then* $\mathbf{x}/\|\mathbf{x}\| \notin T_{D,\kappa,\gamma,\alpha}$ *and* $\mathbf{x}/\|\mathbf{x}\| \in \mathbf{Incomp}(c_0, c_1)$*. We then have the bound*

$$\mathbf{P}(\mathcal{E}_{\mathbf{u},\lambda_0}) \geq 1 - O(\exp(-\alpha_0 n)).$$

Together with the structural results above, we will also later utilize the following result (see [22, Lemma 2.2]) to pass from small ball bounds to a total bound.

Theorem 3.3.15 (Tensorization Lemma)**.** *Let* $\zeta_1, \cdots, \zeta_n$ *be independent nonnegative random variables, and let* $K, t_0 > 0$*. If one has*

$$P(\zeta_k < t) \leq Kt$$

for all $k = 1, \cdots, n$ *and all* $t \geq t_0$*, then for all* $t \geq t_0$

$$P(\sum_{k=1}^{n} \zeta_k^2 < t^2 n) \leq O((Kt)^n).$$

Chapter 4

Proof of Theorem 2.1.5

Now we can use the tools developed in Chapter 2 and Chapter 3 to prove Theorem 2.1.5. This process will occur in two stages. First, we will invoke some extra randomness on the eigenvector $\mathbf{u}$ by noticing that it observes the same law as a random vector $\mathbf{u}'$. We will then utilize this extra randomness to study our main problem.

4.1 Extra Randomness

A key observation, by using the fact that ξ is symmetric, is that if $\varepsilon_1, \ldots, \varepsilon_n$ are iid Bernoulli random variables independent of M_n, then M_n and $M_n' = (\varepsilon_i \varepsilon_j m_{ij})$ have the same matrix distribution. Furthermore, a quick calculation shows that $M_n \mathbf{u} = \lambda \mathbf{u}$ if and only if $M_n' \mathbf{u}' = \lambda \mathbf{u}'$, where $\mathbf{u}' = (\varepsilon_1 u_1, \ldots, \varepsilon_n u_n)$. So the eigenvalues of M_n and M_n' are identical, and the spectrum of M_n is simple if and only if the spectrum of M_n' is simple.

Lemma 4.1.1. *[20, Lemma 10.2] Conditioning on the event $\mathcal{E}$ that the spectrum of M_n is simple, for any $\delta > 0$ and any deterministic vector $\mathbf{f}$ we have*

$$\mathbf{P}\big(|\langle \mathbf{u}, \mathbf{f}\rangle| \leq \delta|\mathcal{E}) = \mathbf{P}(|\langle \mathbf{u}', \mathbf{f}\rangle| \leq \delta|\mathcal{E}\big).$$

Consequently, by Theorem 2.1.4,

$$\mathbf{P}\big(\sup_i |\langle \mathbf{u}_i, \mathbf{f}\rangle| \leq \delta\big) \leq \mathbf{P}\big(\sup_i |\langle \mathbf{u}_i', \mathbf{f}\rangle| \leq \delta\big) + \exp(-n^c).$$

As the proof of this lemma is short but crucial, we insert it here for the reader's convenience.

Proof. (of Lemma 4.1.1) Let λ be the eigenvector associated to both $\mathbf{u}$ and $\mathbf{u}'$. Let P_λ denote the orthogonal projection of M_n onto the eigenspace associated with λ, and let P_λ' denote the orthogonal projection of M_n' onto the eigenspace associated with λ. From the fact that M_n and M_n' have the same distribution, P_λ and P_λ' also have the same distribution. Also, when our spectrum is simple, we have that $P_\lambda(\cdot) = \langle \mathbf{u}, \cdot\rangle\mathbf{u}$ and $P_\lambda'(\cdot) = \langle \mathbf{u}', \cdot\rangle\mathbf{u}'$. It thus follows that

$$\begin{aligned}\mathbf{P}(|\langle \mathbf{u}, \mathbf{f}\rangle| \leq \delta \,|\, \mathcal{E}) &= \mathbf{P}(|\langle \mathbf{u}, \mathbf{f}\rangle||\mathbf{u}| \leq \delta \,|\, \mathcal{E}) \\ &= \mathbf{P}(|P_\lambda(\mathbf{f})| \leq \delta \,|\, \mathcal{E}) \\ &= \mathbf{P}(|P_\lambda'(\mathbf{f})| \leq \delta \,|\, \mathcal{E}) \\ &= \mathbf{P}(|\langle \mathbf{u}', \mathbf{f}\rangle| \leq \delta \,|\, \mathcal{E}),\end{aligned}$$

i.e. $\mathbf{P}(|\langle \mathbf{u}, \mathbf{f}\rangle| \le \delta \cap \mathcal{E}) = \mathbf{P}(|\langle \mathbf{u}', \mathbf{f}\rangle| \le \delta \cap \mathcal{E})$. Hence

$$\mathbf{P}(|\langle \mathbf{u}, \mathbf{f}\rangle| \le \delta) \le \mathbf{P}(|\langle \mathbf{u}', \mathbf{f}\rangle| \le \delta \cap \mathcal{E}) + \exp(-n^c) \le \mathbf{P}(|\langle \mathbf{u}', \mathbf{f}\rangle| \le \delta) + \exp(-n^c),$$

as desired. □

It is remarked that one can deduce from here an almost optimal analog of (2.2) of Theorem 2.1.2, say, for the sequence $\mathbf{f} = (1, \dots, 1)$. Indeed, by Lemma 4.1.1 it suffices to show the comparison for $\mathbf{u}' = (\varepsilon_1 u_1, \dots, \varepsilon_n u_n)$. To this end, by the classical Berry-Esseen bound, as $\sum_i (f_i u_i)^2 = \sum_i u_i^2 = 1$ and $\max_i |u_i| \le n^{-1/2+o(1)}$ (see for instance [5, 6, 32]),

$$\begin{aligned}\mathbf{P}_{\varepsilon_1,\dots,\varepsilon_n}(\sum_i \varepsilon_i u_i f_i \le x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt + \sup_i |f_i u_i| \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt + O(n^{-1/2+o(1)}).\end{aligned}$$

4.2 Starting from Controlled Sets

Now suppose $|\langle \mathbf{u}, \mathbf{f}\rangle| = |u_1 f_1 + \cdots + u_n f_n| \le \delta$ for some unit eigenvector $\mathbf{u}$ of M_n. By Lemma 4.1.1, the probability of this event is bounded above by the probability of the event $|\varepsilon_1 u_1 f_1 + \cdots + \varepsilon_n u_n f_n| \le \delta$ for some unit eigenvector $\mathbf{u}$ of M_n and for some Bernoulli vector $(\varepsilon_1, \dots, \varepsilon_n)$. This extra randomness allows us to study our main problem as follows:

(1) (Randomness on M_n) show that with respect to M_n, the eigenvectors $\mathbf{u} = (u_1, \dots, u_n)$ of M_n do not have structure.

(2) (Randomness on $\varepsilon_1, \dots, \varepsilon_n$) conditioned on the event above, the proof is concluded by applying Theorem 3.3.2.

Now we look at the first step more closely. Without loss of generality we assume that $n^{-c} \le |f_1|, \dots, |f_{n_0}| \le n^c$ for $n_0 = (1-c)n$. For now we fix a parameter i and let $\mathbf{u}$ be the i-th eigenvector. Assume otherwise that

$$\mathbf{P}_{\varepsilon_1,\dots,\varepsilon_n}(|\sum_i \varepsilon_i f_i u_i| \le \delta) \ge n^{2c}\delta.$$

We are not ready to apply Theorem 3.3.2 yet as $\sum_i (u_i f_i)^2$ is not necessarily 1. However, by Condition 2.1.1 and by Lemma 3.3.9, provided that c is sufficiently small, it suffices to consider the case

$$n^{-c} \ll \sqrt{\sum_{i=1}^{n_0} (u_i f_i)^2} \ll n^c.$$

Approximating $\sqrt{\sum_{i=1}^{n_0}(u_i f_i)^2}$ by $\sqrt{p_j}$ where $p_j \in [n^{-c}, n^c]$ is an integral multiple of n^{-5c},

$$1 - n^{-4c} \le \sum_i (\frac{1}{\sqrt{p_j}} f_i u_i)^2 \le 1 + n^{-4c}. \tag{4.1}$$

Thus the event $|\sum_i \varepsilon_i f_i u_i - u| \le \delta$ implies that

$$|\sum_{i=1}^{n_0} \varepsilon_i \frac{1}{\sqrt{p_j}} f_i u_i - \frac{u}{\sqrt{p_j}}| \le n^c \delta.$$

In other words, there exists some p_j such that, with $\delta' = n^c\delta$,

$$\sup_u \mathbf{P}(|\sum_i \varepsilon_i(\frac{f_i u_i}{\sqrt{p_j}} - u)| \le \delta') \ge n^c\delta'.$$

Let $\mathbf{x} = (\frac{f_1 u_1}{\sqrt{p_j}}, \dots, \frac{f_{n_0} u_{n_0}}{\sqrt{p_j}})$. By Theorem 3.3.2, the above implies that

$$D = \mathbf{LCD}_{\gamma,\kappa}(\frac{\mathbf{x}}{\|\mathbf{x}\|}) \le \delta'^{-1}.$$

Notice that as there are many non-zero $u_i, 1 \le i \le n_0$, by Lemma 3.3.9 and by the assumption $\delta \ge \exp(-n^c)$,

$$\sqrt{n} \ll D \ll \exp(n^c) := D_0.$$

By Lemma 3.3.11, there is a set $\mathcal{S}_{j,D_0}$ (corresponding to p_j) of cardinality at most $(CD_0/\sqrt{n})^{n_0}$ which is a $(2\kappa/D_0)$-net for the set of $\mathbf{x}$ above.

For each $\mathcal{S}_{j,D_0}$, we consider the scaling map from $\mathbf{x} = (x_1, \dots, x_{n_0})$ to $\mathbf{v}' = (v_1, \dots, v_{n_0})$:

$$\mathbf{v}' := \frac{(\sqrt{p_j}x_1/f_1, \dots, \sqrt{p_j}x_{n_0}/f_{n_0})}{||\mathbf{x}||}.$$

This creates a new set $\mathcal{V}_{j,D_0}$ of vectors $\mathbf{v}'$ which well approximates the truncated vectors $\mathbf{u}' = (u_1, \ldots, u_{n_0})$ of our eigenvector $\mathbf{u}$,

$$\begin{aligned}\|\mathbf{u}' - \mathbf{v}'\| &= \sqrt{\sum_{i=1}^{n_0}(u_i - \frac{\sqrt{p_j}x_i}{||\mathbf{x}||f_i})^2} \\ &\leq \sqrt{n^{2c}||\mathbf{x}||^2\sum_{i=1}^{n_0}(\frac{||\mathbf{x}||f_iu_i}{\sqrt{p_j}} - x_i)^2} \\ &\leq n^c(1+n^{-4c})\frac{2\kappa}{D_0} \\ &\leq n^c\frac{4\kappa}{D_0}\end{aligned}$$

We can also κ/D_0-approximate the remaining $n - n_0$ coordinates trivially by a set of size $(D_0/\kappa)^{n-n_0} = (D_0/\kappa)^{cn}$. Appending this to $\mathcal{V}_{j,D_0}$ above, and taking the union over p_j, we obtain the following:

Theorem 4.2.1. *There exists a deterministic set $\mathcal{V}$ of size $n^{O(1)}(CD_0/\sqrt{n})^n(\sqrt{n}/\kappa)^{cn}$ such that for any unit vector $\mathbf{u} \in S^{n-1}$ with $\sup_u \mathbf{P}(|\sum_i \varepsilon_i f_i u_i - u| \leq \delta') \geq n^c\delta'$, there exists $\mathbf{v} \in \mathcal{V}$ such that*

$$\|\mathbf{u} - \mathbf{v}\| \ll n^c\kappa/D_0.$$

Notice that by the approximation, for any $\mathbf{v} \in \mathcal{V}$

$$1 - O(n^c\kappa/D_0) \leq \|\mathbf{v}\| \leq 1 + O(n^c\kappa/D_0).$$

Using this approximation, if $(M_n - \lambda)\mathbf{u} = 0$ then by (3.1), with $\beta_0 = \kappa n^c/D_0$,

$$\|(M_n - \lambda)\mathbf{v}\| \leq \sqrt{n}\beta_0.$$

From now on, let $t_i := i/D_0$. We say that $\mathbf{v}$ is an *approximate vector* of M_n if there exists i such that

$$\|(M_n - t_i)\mathbf{v}\| = O(\sqrt{n}\beta_0).$$

4.3 Concluding the Proof of Theorem 2.1.5

In what follows we will choose $\alpha = n^{-6c}$, for a constant c to be chosen sufficiently small. Our main goal is to show the following.

Theorem 4.3.1. *With $\mathcal{V}$ from Theorem 4.2.1,*

$$\mathbf{P}\Big(\exists i, \exists \mathbf{v} \in \mathcal{V}, \|(M_n - t_i)\mathbf{v}\| \le \beta_0 n^{1/2}\Big) \le \exp(-\alpha_0 n).$$

It is clear that Theorem 2.1.5 follows from Theorem 4.3.1. It remains to prove Theorem 4.3.1 for a fixed t_i, and then take the union bound over t_i (the factor of D_0 will be absorbed by $\exp(-c_0 n)$). Recall that $\beta_0 = \kappa n^c/D_0$ and $\alpha = n^{-6c}$. We now condition on the event $\mathcal{E}_{\mathbf{0},t_i}$ of Corollary 3.3.14 with $D = D_0$ and $\beta_1 = \kappa/\sqrt{\alpha}D_0$. On this event, if $\|(M_n - t_i)\mathbf{v}\| \le \beta_0 n^{1/2} = o(\beta_1 n^{1/2}), \mathbf{v} \in S^{n-1}$, then

$$\mathbf{v}/\|\mathbf{v}\| \in \mathbf{Incomp}(c_0, c_1) \text{ and } \widehat{\mathbf{LCD}}_{\kappa,\gamma}(\mathbf{v}/\|\mathbf{v}\|, \alpha) \ge D_0. \tag{4.2}$$

Consequently, on $\mathcal{E}_{\mathbf{0},t_i}$, for any $\mathbf{v} \in \mathcal{V}$ we either have $\|(M_n - t_i)\mathbf{v}\| > \beta_0 n^{1/2}$ or (4.2) holds for $\mathbf{v}$. So to prove Theorem 4.3.1 for t_i one just need to focus on these vectors $\mathbf{v}$.

Set $n' = \alpha n$. For $\mathbf{v} = (v_1, \dots, v_n)$, let $p_{\alpha,\beta}(\mathbf{v})$ be as below

$$p_{\alpha,\beta}(\mathbf{v}) = \inf_{i_1,\dots,i_{n'}} \sup_x \mathbf{P}(|\xi_{i_1} v_{i_1} + \cdots + \xi_{i_{n'}} v_{i_{n'}} - x| \le \beta).$$

By splitting M_n accordingly,

$$M_n = \begin{pmatrix} M_{n-n'} & B \\ B^* & M_{n'} \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} \mathbf{v}' \\ \mathbf{v}'' \end{pmatrix},$$

where $M_{n'}$ is the $n' \times n'$ principle minor of M_n with indices $i_1, \dots, i_{n'}$ and $M_{n-n'}$ is the remaining principle minor. Here $\mathbf{v}' \in \mathbf{R}^{n-n'}$ and $\mathbf{v}'' \in \mathbf{R}^{n'}$.

So $\|(M_n - t_i)\mathbf{v}\| \le \beta_0\sqrt{n}$ implies that

$$\|B\mathbf{v}'' - (M_{n-n'} - t_i)\mathbf{v}'\| \le \beta_0\sqrt{n}.$$

We will condition on the matrix $M_{n-n'}$. Using Theorem 3.3.15, we thus have

$$\mathbf{P}(\|(M_n - t_i)\mathbf{v}\| \le n^{1/2}\beta_0) \le (2\rho_{\alpha,\beta_0}(\mathbf{v}))^{n-n'}.$$

Indeed, we will consider $\mathbf{P}(\sum r_i^2 \le \beta_0^2 n)$, where

$$\mathbf{r}_i = b_{i,1}v_{n-n'+1} + \cdots + b_{i,n'}v_n - (m_{i,1}v_1 + \cdots + (m_{i,i} - t_i)v_i + \cdots + m_{i,n-n'}v_{n-n'})$$

denotes the i^{th} row of $B\mathbf{v}'' - (M_{n-n'} - t_i)\mathbf{v}'$. Conditioning on B, we have that $\mathbf{P}(|\mathbf{r}_i| \le \beta_0) \le \rho_{\alpha,\beta_0}$ by the definition of $\rho_{\alpha,\beta}$. We claim that $\mathbf{P}(|\mathbf{r}_i| \le t)$ is true for every $t \ge t_0$ with $t_0 = \beta_0$ and $K = \rho_{\alpha,\beta_0}/\beta_0$. Indeed, breaking the interval $[0, t)$ into $\lceil t/\beta_0 \rceil$

intervals each of length at most β_0, we have that

$$\mathbf{P}(|\mathbf{r}_i| \le t) \le (t/\beta_0 + 1)\rho_{\alpha,\beta_0} \le 2Kt$$

and we are done via Theorem 3.3.15.

Now we estimate the event considered in Theorem 4.3.1 for a fixed t_i conditioning on $\mathcal{E}_{\mathbf{0},t_i}$

$$\mathbf{P}\Big(\exists \mathbf{v} \in \mathcal{V}, \mathbf{v} \text{ satisfies } (4.2), \|(M_n - t_0)\mathbf{v}\| \le \beta_0 n^{1/2}\Big) \le \sum_{\mathbf{v}\in\mathcal{V},\mathbf{v}\in(4.2)} (2\rho_{\alpha,\beta_0}(\mathbf{v}))^{n-n'}.$$

To this end, as $\mathbf{v}$ satisfies (4.2)

$$\widehat{\mathbf{LCD}}_{\kappa,\gamma}(\mathbf{v}/\|\mathbf{v}\|, \alpha) \ge D_0.$$

By definition, there exists $I \subseteq \mathbf{spread}(\mathbf{v})$, $|I| = \lceil \alpha n \rceil$ such that

$$\mathbf{LCD}_{\kappa,\gamma}\big(\mathbf{v}_I/\|\mathbf{v}_I\|\big) \ge D_0 = n^c \beta_0^{-1}.$$

Thus

$$\rho_{\alpha,\beta_0}(\mathbf{v}) \le \rho_{\beta_0/\sqrt{\alpha}}(\mathbf{v}_I/\|\mathbf{v}_I\|) = O(\beta_0 n^{4c}),$$

where in the last estimate we apply Theorem 3.3.2 as $\beta_0 n^{4c} > 1/D_0$.

So

$$\begin{aligned}\sum_{\mathbf{v}\in\mathcal{V}}(2\rho_{\alpha,\beta_0}(\mathbf{v}))^{n-n'} &\le (\beta_0 n^{4c})^{(1-\alpha)n}|\mathcal{V}|\\ &\le (C'\beta_0 n^{4c})^{(1-\alpha)n} n^{O(1)}(CD_0/\sqrt{n})^n(\sqrt{n}/\kappa)^{cn}\\ &\le (C'\beta_0 n^{4c})^{(1-\alpha)n} n^{O(1)}(Cn^c\beta_0^{-1}/\sqrt{n})^n(\sqrt{n}/n^{2c})^{cn}\\ &\le \beta_0^{-\alpha n} n^{-(1/2-6c)n}\\ &\le e^{n^c n^{-6c} n} n^{-(1/2-6c)n}\\ &\le n^{-(1/2-6c)n},\end{aligned}$$

provided that n is sufficiently large, where we noted that $\beta_0 > 1/D_0 = \exp(-n^c)$ and c is sufficiently small.

The proof of Theorem 4.3.1 is then complete where the bound $\exp(-\alpha_0 n)$ comes from the complement of the event of Corollary 3.3.14 we conditioned on.

Chapter 5

Extending the LCD to ULCD

In this chapter, we seek to generalize our previous results to that of characteristic p. We will be able to take the previous notions of small ball probability, **LCD**, distance and others, in order to find equivalent statements over $\mathbf{F}_p$. This is a very nontrivial process, as there is no obvious geometric generalization of vectors when switching to $\mathbf{F}_p^n$. The content from following chapters will appear in [10].

5.1 Generalization of LCD from R to $\mathbf{F}_p^n$

Let $\mathbf{w} = (w_1, \dots, w_n)$ be a non-zero vector in $\mathbf{F}_p^n$, where p is a prime. We start with an analogue of Theorem 3.1.4 in $\mathbf{F}_p$. This Erdős-Littlewood-Offord type theorem [11] shows that the concentration probability $\rho(\mathbf{w})$ is actually uniformly centered around $1/p$:

Theorem 5.1.1. *Let $c_{nsp} > 0$ be a constant, and assume that*

$$|\mathbf{supp}(\mathbf{w})| \geq c_{nsp} n. \tag{5.1}$$

Then

$$\rho(\mathbf{w}) := \sup_r |\mathbf{P}(X \cdot \mathbf{w} = r(\mod p)) - \frac{1}{p}| = O(\frac{1}{\sqrt{n}})$$

where the implied constant depends on c_{nsp}*, and where* $X = (x_1, \dots, x_n)$ *and* x_i *are iid Bernoulli.*

We will still refer to $\rho(\mathbf{w})$ as the concentration probability of $\mathbf{w}$ (with respect to $\mathbf{F}_p$.) We will see that in many situations, our deterministic vector $\mathbf{w}$ satisfies the non-sparsity property (5.1), so we will assume this to be true for now. Note that if $\mathbf{w}$ satisfies (5.1), then so do the dilations $t\mathbf{w}$ of $\mathbf{w}$ in $\mathbf{F}_p^n$ for non-zero t.

Motivated by the work of [22] by Rudelson and Vershynin on structure of vectors in characteristic zero, we will develop a similar method with respect to $\mathbf{F}_p$. There are many differences between the two settings, such as there is no straightforward geometric extension of compressible and incompressible vector in $\mathbf{F}_p^n$.

In some situations, if $\mathbf{w} = (w_1, \dots, w_n)$ is a vector in $\mathbf{F}_p^n$, then by viewing $\mathbf{F}_p$ as the interval $\mathbf{I}_p = [-(p-1)/2, (p-1)/2]$ in $\mathbf{Z}$, we will consider the components w_i as integers from this interval. We then write $\mathbf{w}'$ as the vector in $\mathbf{R}^n$

$$\mathbf{w}' = (w_1', \dots, w_n') = \frac{1}{p}\mathbf{w} = (w_1/p, \dots, w_n/p).$$

Note that we have associated our original vector $\mathbf{w}$ in $\mathbf{F}_p^n$ with a new vector $\mathbf{w}'$ in $\mathbf{R}^n$. Now that we have switched to $\mathbf{R}^n$, we can appeal to our usual notions of geometry in order to define a new LCD:

Definition 5.1.2. Let $0 < \gamma < 1$ and κ be given. Let $\mathbf{w}' = (w_1', \dots, w_n') \in \mathbf{R}^n$ be a non-zero vector in $\frac{1}{p}\mathbf{Z}^n$ where $\|\mathbf{w}'\|_\infty \leq 1/2$. We denote by $\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}')$ to be the smallest (infimum) positive integer L such that

$$\mathrm{dist}(L\mathbf{w}', \mathbf{Z}^n) \leq \min\{\gamma \|L\mathbf{w}'\|_2, \kappa\}.$$

Throughout this and later chapters, $\gamma < 1$ is an absolute constant (such as $\gamma = 1/8$), and $\kappa \leq n^c$, for some positive constant $c \leq 1/2$ to be chosen.

This definition is in characteristic zero. Here we used the notion of **ULCD** (compared to the notion of **LCD** from [22]) to emphasize that $(w_1', \dots, w_n')$ is not normalized (i.e. its ℓ_2-norm might not be unit). Notice that

$$2 \leq \mathbf{ULCD}_{\gamma,\kappa} \leq p.$$

Furthermore, if $\mathbf{ULCD}_{\gamma,\kappa} = p$ then by definition we would have for all $1 \leq L \leq p-1$ that

$$\mathrm{dist}(L\mathbf{w}', \mathbf{Z}^n) \geq \min\{\gamma \|L\mathbf{w}'\|_2, \kappa\}. \tag{5.2}$$

Remark 5.1.3. *Note that if for some $T > 1$ we have $|w_i'| \leq 1/2T$ for all i, then*

$$\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \geq T.$$

This is because otherwise, then $\|Lw_i'\|_{\mathbf{R}/\mathbf{Z}} = |Lw_i'|$ and hence $\sum_i \|Lw_i'\|_{\mathbf{R}/\mathbf{Z}}^2 = \sum_i \|Lw_i'\|_2^2$, which cannot be smaller than $\gamma^2 \|L\mathbf{w}'\|_2^2$ by definition as $\gamma < 1$.

Our result below says that if $\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}')$ is large then the concentration probability is small. In our notation $t\mathbf{w}$ is another vector in $\mathbf{F}_p^n$, which again can be viewed as a vector in $\mathbf{I}_p^n = [-(p-1)/2, (p-1)/2]^n$, we then define $(t\mathbf{w})'$ as $\frac{1}{p}t\mathbf{w}$ accordingly in this projection to characteristic zero.

With the definition of $\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}')$ now in place, we can look to prove a generalization of Theorem 3.3.2:

Theorem 5.1.4 (anti-concentration modulo a prime)**.** *Let $p \geq 3$ be a prime, and let $C > 0$ be an arbitrary constant. Let $\mathbf{w} = (w_1, \dots, w_n)$ be a non-zero vector in $\mathbf{Z}^n$, where $|w_i| \leq p/2$, and let $\mathbf{w}' = \frac{1}{p}\mathbf{w} = (\frac{w_1}{p}, \dots, \frac{w_n}{p})$. Then*

1. *If there is no non-zero $t \in \mathbf{F}_p$ such that $\|(t\mathbf{w})'\|_2 < \kappa$ then we have*

$$\rho(\mathbf{w}) \leq \exp(-\kappa^2/2).$$

2. *Otherwise, assume that $1 \leq \|\mathbf{w}'\|_2 \leq \kappa$ and $\mathbf{w}$ satisfies (5.1), with $\kappa \leq C\sqrt{n}$ and*

$$\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \geq L.$$

Then

$$\rho(\mathbf{w}) \leq O\Big(\frac{1}{L\|\mathbf{w}'\|_2} + \exp(-\Theta(\kappa^2))\Big),$$

where the implied constants depend on C, γ, c_{nsp}, and where $\mathbf{X} = (x_1, \dots, x_n)$ and x_i are iid Bernoulli in the concentration definition $\rho(\mathbf{w})$.

Corollary 5.1.5. *Assume that there exists a quantity $\rho \gg \exp(-\Theta(\kappa^2))$ such that $\rho(\mathbf{w}) \geq \rho$, then there exists a dilation $t\mathbf{w}$ of $\mathbf{w}$, where $t \in \mathbf{F}_p$ non-zero, so that with*

$\mathbf{w} = t\mathbf{w}$ *we have* $\|\mathbf{w}'\|_2 < \kappa$ *and there exists* $L = L(\mathbf{w}) \geq 1$ *such that*

$$L = O\Big(\frac{\rho^{-1}}{\|\mathbf{w}'\|_2}\Big)$$

and

$$\mathrm{dist}(L\mathbf{w}', \mathbf{Z}^n) \leq \kappa.$$

We next deduce another elementary but useful result, which will be used later on in the applications.

Corollary 5.1.6. *Assume that* $\mathbf{w}$ *has at least* m *non-zero coordinates, and* $p < \sqrt{m}$. *We then have*

$$\rho(\mathbf{w}) \leq \exp(-cm/p^2).$$

Proof. As $t\mathbf{w}'$ has at least m non-zero coordinates for any non-zero t, we have that

$$\|t\mathbf{w}'\|_2 \geq \sqrt{m(1/p)^2} = \sqrt{m}/p,$$

as desired. □

We now present a proof of Theorem 5.1.4, which we prove in full. Readers may note the similarities between this proof and the sketched proof of 3.3.2.

Proof. (of Theorem 5.1.4) Write $e_p(x) = e^{2\pi i x/p}$, then for any $r \in \mathbf{F}_p$ we have

$$\mathbf{P}(X \cdot \mathbf{w} = r) - \frac{1}{p} = \frac{1}{p} \sum_{t \in \mathbf{F}_p, t \neq 0} \prod_{l=1}^{n} \mathbf{E} e_p(x_l w_l t) e_p(-tr).$$

So

$$\begin{aligned}|\mathbf{P}(X\cdot\mathbf{w}=r)-\frac{1}{p}| &\le \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0}|\prod_{l=1}^{n}\mathbf{E}e_p(x_l w_l t)|\\ &= \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0}\prod_{l=1}^{n}|\cos(2\pi w_l t/p)|\\ &= \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0}\prod_{l=1}^{n}|\cos(\pi w_l t/p)|\\ &\le \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0}e^{-2\sum_{l=1}^{n}\|\frac{w_l t}{p}\|^2}\\ &\le \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0}e^{-2\sum_{l=1}^{n}\|tw_l'\|_{\mathbf{R}/\mathbf{Z}}^2},\end{aligned}$$

where we used the fact that $|\sin\pi z|\ge 2\|z\|_{\mathbf{R}/\mathbf{Z}}$ for any $z\in\mathbf{R}$, where $\|z\|_{\mathbf{R}/\mathbf{Z}}$ is the distance of z to the nearest integer, and that

$$|\cos\frac{2\pi x}{p}|\le 1-\frac{1}{2}\sin^2\frac{2\pi x}{p}\le 1-2\|\frac{2x}{p}\|^2.$$

From here, (1) follows as $\|tw_l'\|_{\mathbf{R}/\mathbf{Z}}=\|(t\mathbf{w})_l'\|_{\mathbf{R}/\mathbf{Z}}$.

We are now in the assumption of (2). For each integer m, let $T(m,p/2)$ be the (level) set of $t\in\mathbf{F}_p$ corresponding to m,

$$T(m,p/2)=\Big\{t\in\mathbf{F}_p:\|t\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}}^2:=\sum_{l=1}^{n}\|tw_l'\|_{\mathbf{R}/Z}^2\le m\Big\}=\Big\{t\in\mathbf{I}_p:\|t\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}}^2\le m\Big\}.$$

By the non-sparsity of $\mathbf{w}$, we can show that $T(c_{nsp}n/64,p/2)$ is not the whole $\mathbf{F}_p$ (we can show this by using the fact that if $w_i\neq 0$ then $\sum_{t\in\mathbf{F}_p}\|\frac{tw_i}{p}\|_{\mathbf{R}/\mathbf{Z}}=$

$\sum_{t\in\mathbf{F}_p} \|\frac{t}{p}\|_{\mathbf{R}/\mathbf{Z}})$. Thus

$$|T(c_{nsp}n/64, p/2)| \leq p-1.$$

Our next claim shows that the level sets consist of well separated intervals.

Claim 5.1.7 (spacing of the level sets)**.** *Assume that* $m < \kappa^2/4 < c_{nsp}n/64$ *and* $t_1 < t_2 \in T(m, p/2)$ *and*

$$|t_2 - t_1| \geq \gamma^{-1}\kappa/\|w'\|_2.$$

Then

$$t_2 - t_1 \geq L.$$

Consequently we have

$$T(m, p/2) \leq \frac{p\lceil \gamma^{-1}\kappa/\|w'\|_2\rceil}{L}.$$

Proof. (of Claim 5.1.7) Assume that $t_1, t_2 \in T(m, p/2)$, then by the triangle inequality

$$\|(t_1 - t_2)\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}} \leq \|t_1\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}} + \|t_2\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}} \leq 2\sqrt{m} < \kappa.$$

Thus

$$\operatorname{dist}((t_2 - t_1)\mathbf{w}', \mathbf{Z}^n) = \|(t_2 - t_1)\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}} < \kappa \leq \gamma(t_2 - t_1)\|\mathbf{w}'\|_2,$$

where in the last estimate we used $|t_2 - t_1| > \gamma^{-1}\kappa/\|\mathbf{w}'\|_2$. Thus by the definition of $\mathbf{ULCD}_{\gamma,\kappa}$ we must have

$$t_2 - t_1 \geq L,$$

completing our proof □

Next, we will need a Cauchy-Davenport-type bound on size of sumsets in $\mathbf{F}_p$. Observe from the Cauchy-Schwarz inequality, $k(\|x_1\|^2_{\mathbf{R}/\mathbf{Z}} + \cdots + \|x_k\|^2_{\mathbf{R}/\mathbf{Z}}) \geq \|x_1 + \cdots + x_k\|^2_{\mathbf{R}/\mathbf{Z}}$, and so

$$kT(m, p/2) \leq T(k^2 m, p/2),$$

where we view these sets as subsets of $\mathbf{F}_p$. Hence, by Cauchy-Davenport's inequality in $\mathbf{F}_p$ (see Appendix C) we have that

$$|T(k^2 m)| \geq |kT(m)| \geq k|T(m)| - (k-1).$$

Thus for all $m \leq \min\{c_{nsp}\kappa^2/64, c_{nsp}n/64\}$, by choosing

$$k = \lfloor \sqrt{\min\{c_{nsp}\kappa^2/64, c_{nsp}n/64\}/m} \rfloor$$

we have

$$|T(m, p/2)| - 1 \leq k^{-1}(|T(c_{nsp}n/64, p/2)| - 1) \leq \sqrt{\frac{m}{O_{c_{nsp},C}(\kappa^2)}}p,$$

where we used $|T(c_{nsp}n/64, p/2)| - 1 < p$.

We thus deduce

$$
\begin{aligned}
|\mathbf{P}(X\cdot\mathbf{w}=r)-\frac{1}{p}| &\le \frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0} e^{-2\sum_{l=1}^n \|t\mathbf{w}'\|^2}\\
&\le \frac{1}{p}\Big(\frac{p\lceil\gamma^{-1}\kappa/\|\mathbf{w}'\|_2\rceil}{L\sqrt{O_{c_{nsp},C}(\kappa^2)}}\sum_{m\le\min\{c_{nsp}\kappa^2/64,c_{nsp}n/64\}}\sqrt{m}e^{-m/2}\\
&+p\sum_{m>\min\{c_{nsp}\kappa^2/64,c_{nsp}n/64\}} e^{-m/2}\Big)\\
&= O_{c_{nsp},C,\gamma}\Big(\frac{1}{L\|\mathbf{w}'\|_2}+\exp(-\Theta_{C,c_{nsp}}(\kappa^2))\Big),
\end{aligned}
$$

completing the theorem proof. □

Remark 5.1.8. *Under the assumption of (2) of Theorem 5.1.4, we have actually shown a stronger estimate that*

$$
\frac{1}{p}\sum_{t\in\mathbf{F}_p,t\neq 0} e^{-2\sum_{l=1}^n \|t\mathbf{w}'\|^2} \le O\Big(\frac{\lceil c_{nsp}^{-1/2}\gamma^{-1}\kappa/\|\mathbf{w}'\|_2\rceil}{\kappa L}+\exp(-c_{nsp}\kappa^2/64)\Big).
$$

We note that Theorem 5.1.1 can be deduced from Theorem 5.1.4 by setting $\kappa = C\sqrt{n}$ with sufficiently large C; for this there is a dilation $\mathbf{w}=t\mathbf{w}$ with $\|\mathbf{w}'\|_2$ of order $\sqrt{n}$ but $\|\mathbf{w}'\|_2<\kappa$. We then just apply (2) of Theorem 5.1.4, noting that $L\ge 1$.

Roughly speaking, our next result is similar to Theorem 5.1.4, but instead of working with the concentration event $X\cdot\mathbf{w}=r$, we are working with a coarser event that $X\cdot\mathbf{w}$ belongs to an arc in $\mathbf{F}_p$. We find it more convenient to write in mod 1 as follows.

Theorem 5.1.9 (anti-concentration modulo one). *Assume that* $0 < a_1, \dots, a_n < 1$. *Assume that*

$$\mathbf{ULCD}_{\gamma,\kappa}((a_1, \dots, a_n)) = L \geq 1.$$

Then for any

$$\varepsilon > 2/L$$

we have

$$\mathbf{P}(\| \sum_i \xi_i a_i \|_{\mathbf{R}/\mathbf{Z}} \leq \varepsilon) \leq O\Big(\varepsilon + \log(1/\varepsilon)[\exp(-\kappa^2) + \exp(-4\gamma^2 \sum_i a_i^2)]\Big).$$

A key difference of this bound compared to the classical small ball estimate (say studied in [22]) is that we are looking at the balls modulo one, rather than with respect to the whole real line.

Proof. (of Theorem 5.1.9) Let μ be the distribution of $\sum_i a_i \xi_i$ modulo one, where we can write $\mu = \mu_1 * \cdots * \mu_n$, and where $\mu_i(a_i) = \mu(-a_i) = 1/2$. Let $L_0 = \lfloor 1/\varepsilon \rfloor$. We use the Erdős-Turán inequality,

$$\mu([-\varepsilon, \varepsilon]) = \mathbf{P}(\| \sum_i a_i \xi_i \|_{\mathbf{R}/\mathbf{Z}} \leq \varepsilon) \leq \varepsilon + \frac{1}{L_0} + \sum_{k=1}^{L_0} \frac{|\widehat{\mu}(k)|}{k},$$

As ξ_i are iid Bernoulli, bounding the cosine as in the proof of Theorem 5.1.4 we have

$$|\widehat{\mu}(k)| = | \int_0^1 e^{2\pi\sqrt{-1}k\theta} d\mu(\theta)| = | \prod_i \int_0^1 e^{2\pi\sqrt{-1}k\theta} d\mu_i(\theta)| \leq \exp(-\sum_i \|2a_i k\|_{\mathbf{R}/\mathbf{Z}}^2).$$

Now by definition of $\mathbf{ULCD}_{\gamma,\kappa}$, as $L_0 \le 1/\varepsilon < \mathbf{ULCD}_{\gamma,\kappa}((a_1,\dots,a_n))/2$, for any $k \le L_0$ we have

$$\sum_i \|2a_i k\|^2_{\mathbf{R}/\mathbf{Z}} \ge \min\{\kappa^2, 4\gamma^2 k^2 \sum_i a_i^2\},$$

and so

$$\exp(-\sum_i \|2a_i k\|^2_{\mathbf{R}/\mathbf{Z}}) \le \exp(-\kappa^2)+\exp(-4\gamma^2 k^2 \sum_i a_i^2) \le \exp(-\kappa^2)+\exp(-4\gamma^2 \sum_i a_i^2).$$

Summing over all $k \le L_0$ we have

$$\mathbf{P}(\|\sum_i \xi_i a_i\|_{\mathbf{R}/\mathbf{Z}} \le \varepsilon) \le O(\varepsilon) + \log(1/\varepsilon)[\exp(-\kappa^2) + \exp(-4\gamma^2 \sum_i a_i^2)]$$

and the result follows. □

It is remarked that the bound above depends on $\|\mathbf{a}\|_2$, which becomes almost meaningless if $\|\mathbf{a}\|_2$ is small, say of order $O(1)$. To avoid this situation, we will need to consider vectors $\mathbf{w}'$ that have large size and large $\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}')$ at the same time.

Our next result roughly says that non-sparse vectors cannot have very small LCD, at least with respect to $\mathbf{F}_p$ with p not too large but also not too small. To be more precise, we have the following.

Remark 5.1.10. *As we will be working with vectors* $\mathbf{w}$ *satisfying* (5.1)*, we easily see that for any* $t \neq 0$ *in* $\mathbf{F}_p$

$$\|(t\mathbf{w})'\|^2 \ge cn(\frac{1}{p})^2.$$

Notice that this quantity is larger than κ^2 if $p \le n^{1/2}/\kappa$, and in this case the first part of 5.1.4 holds, and hence automatically

$$\rho(\mathbf{w}) \le \exp(-c\kappa^2).$$

As such, in what follows we will be working with

$$p \gg n^{1/2}/\kappa.$$

Lemma 5.1.11 (LCD and size in fields of small order)**.** *Assume that $\kappa = n^c$ for a positive constant $c < 1/16$. Assume that p is a prime smaller than* $\exp(c\kappa^2)$*, and $\mathbf{w} \in \mathbf{F}_p^n$ is a vector satisfying* (5.1) *and such that $\rho(\mathbf{w}) \ge 2\exp(-\kappa^2/2)$. Then there exists $t \in \mathbf{F}_0$ so that with $\mathbf{w} = t\mathbf{w}$ we have $\|\mathbf{w}'\|_2$ has order κ and either $\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') = p$ (in which case we can apply* (5.2)*) or else*

$$\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \ge \kappa^{5/4-c}.$$

Before proving this lemma, we first need the following simple statement.

Claim 5.1.12. *Assume that $\mathbf{w} \in \mathbf{F}_p^n$ is a non-zero vector satisfying* (5.1) *and such that $\rho(\mathbf{w}) \ge 2\exp(-\kappa^2/2)$, with $\kappa = o(\sqrt{n})$. Then there exists $t \in \mathbf{F}_0$ so that with $\mathbf{w} = t\mathbf{w}$ we have*

$$\kappa/2 \le \|\mathbf{w}'\|_2 < \kappa.$$

Proof. (of Claim 5.1.12) As $\mathbf{w}$ satisfies (5.1) and $\rho(\mathbf{w}) \geq 2\exp(-\kappa^2/2)$, (1) of Theorem 5.1.4 does not apply, and so there is a fiber $\mathbf{w} = t\mathbf{w}$ such that $\|\mathbf{w}'\|_2 < \kappa$. If $\|\mathbf{w}'\|_2 \geq \kappa/2$ then we would be done. Otherwise, we just consider the sequence $\mathbf{w}, 2\mathbf{w}, 3\mathbf{w}$, etc. By the triangle inequality (where we recall that $(t\mathbf{w})' = \frac{1}{p}(tw_1(\mod p), \dots, tw_n(\mod p)))$ we have

$$\|((k+1)\mathbf{w})'\|_2 \leq \|(k\mathbf{w})'\|_2 + \|\mathbf{w}'\|_2.$$

On the other hand, by (5.1) $\sum_{k\in\mathbf{F}_p} \|(k\mathbf{w})'\|_2$ has order $\sqrt{n}p$, so there must exist a smallest $k_0 \geq 2$ such that $\|((k_0+1)\mathbf{w})'\|_2 \geq \kappa$. It then follows that $\kappa/2 \leq \|(k_0\mathbf{w})'\|_2 < \kappa$. □

With this in place, we can now construct a proof of Lemma 5.1.11.

Proof. (of Lemma 5.1.11) Assume that we are not in the first case, and also assume otherwise that we are not in the second case either. We will iterate the following process, which will then result in a contradiction. Set

$$\beta = 1/4 - c.$$

We start from any $\mathbf{u}_1 = \mathbf{w}' = (w_1/p, \dots, w_n/p)$ in the fiber $t\mathbf{w}$ of $\mathbf{w}$ with $\kappa/2 \leq \|\mathbf{u}_1\|_2 < \kappa$.

Step 1: Let $D_1 = \mathbf{ULCD}(\mathbf{u}_1)$, then $2 \le D_1 \le \min\{\kappa^{1+\varepsilon}, p-1\}$. Let $\mathbf{u}_1' = D_1\mathbf{u}_1 (= (D_1\mathbf{w})')$, then we have

$$\|\mathbf{u}_1'\|_2 \le \kappa.$$

Step 2: If this vector has norm smaller than $\kappa/2$, then we use Claim 5.1.12 to dilate appropriately by $C_1 \ge 2$ so that $\kappa/2 \le \|C_1\mathbf{u}_1'\|_2 \le \kappa$, and set

$$\mathbf{u}_2 = C_1\mathbf{u}_1' = (C_1D_1\mathbf{w})'.$$

We then return to Step 1 and iterate the process, noting that while the D_i are bounded by $\kappa^{1+\varepsilon}$, we don't have such a bound for the C_i.

Now for each $1 \le t \le p-1$ we can always write

$$t = D_1(C_1(D_2(\dots) + r_2) + s_1) + r_1,$$

where $r_1 < D_1, s_1 < C_1, r_2 < D_2, s_2 < C_2, \dots$, and so on. Indeed, to verify this we first divide t by D_1 and get a remainder r_1; we then divide the quotient by C_1 to get a remainder s_1, and then divide the new quotient by D_2, and so on until the last step. Now as $t \le p-1 \le 2^{\kappa^2}$ (this is where we require p to be small), and as $C_i, D_i \ge 2$, we must stop the division process after κ^2 steps.

Next we we analyze the norm of $\|t\mathbf{u}_1\|_2 = \|(t\mathbf{w})'\|_2$. We write, with $t_1 = t$:

$$\begin{aligned} t_1\mathbf{u}_1 &= D_1(C_1(D_2(\dots) + r_2) + s_1)\mathbf{u}_1 + r_1\mathbf{u}_1 \\ &= (C_1(D_2(\dots) + r_2) + s_1)D_1\mathbf{u}_1 + r_1\mathbf{u}_1 \\ &:= t_1'\mathbf{u}_1' + r_1\mathbf{u}_1. \end{aligned}$$

Thus by the triangle inequality, and as $r_1 \le D_1 - 1 < \kappa^{1+\varepsilon}$ we have

$$\|t_1\mathbf{u}_1\|_2 = \|t\mathbf{u}_1\|_2 \le \|t_1'\mathbf{u}_1'\|_2 + \kappa^{2+\varepsilon}.$$

We next consider

$$\begin{aligned} t_1'\mathbf{u}_1' &= (C_1(D_2(\dots) + r_2) + s_1)\mathbf{u}_1' \\ &= C_1(D_2(\dots) + r_2)\mathbf{u}_1' + s_1\mathbf{u}_1' \\ &= (D_2(\dots) + r_2)\mathbf{u}_2 + s_1\mathbf{u}_1' \\ &:= t_2\mathbf{u}_2 + s_1\mathbf{u}_1'. \end{aligned}$$

By the triangle inequality

$$\|t_1'\mathbf{u}_1'\|_2 \le \|t_2\mathbf{u}_2\|_2 + \|s_1\mathbf{u}_1'\|_2 \le \|t_2\mathbf{u}_2\|_2 + \kappa,$$

where in the last estimate we used the fact that $0 \le s_1 \le C_1 - 1$ and C_1 is the largest integer so that $\kappa/2 \le \|C_1 \mathbf{u}_1'\|_2 \le \kappa$ (where we recall that $\|\mathbf{u}_1'\|_2 \le \kappa$, and the role of C_1 was only to dilate this vector if its norm was much smaller than this, as in the proof of Claim 5.1.12). The analysis for $\|t_2 \mathbf{u}_2\|_2$ and other terms can be done similarly.

Adding all the bounds, we hence obtain

$$\|t\mathbf{u}_1\|_2 \le \kappa^2 \times \kappa^{2+\beta} = \kappa^{4+\beta}.$$

Now as this is true for all $t \in \mathbf{F}_p$, we thus have

$$\sum_{t \in \mathbf{F}_p} \|t\mathbf{w}'\|_2^2 = O(p\kappa^{4+\beta}).$$

On the hand, by (5.1), as $\mathbf{w}'$ has at least $c_{nsp}n$ non-zero entries, the left hand side can be shown to be at least $c_{nsp}np/64$, thus a contradiction if $\kappa \le n^c = n^{1/4-\beta}$. □

With the same proof, we record the following corollary which will be used later.

Corollary 5.1.13 (LCD cannot be small). *Assume that $p \le \exp(c\kappa^2)$ and that $\kappa = n^c$ for $c < 1/16$. Assume that $\mathbf{w} \in \mathbf{F}_p^m$, for $m \ge \kappa^{4+2\varepsilon}$, and $\mathbf{w}$ has at least $\kappa^{4+2(1/4-c)}$ non-zero components. Then we either have either $\|(t\mathbf{w})'\|_2 > \kappa$ for all t, or there exists such $\mathbf{w}' = (t\mathbf{w})'$ such that $\kappa/2 \le \|\mathbf{w}'\|_2 < \kappa$ and that*

$$\mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \ge \kappa^{5/4-c}.$$

Chapter 6

The Random Normal Vector in $\mathbf{F}_p^n$

6.1 Distribution of Rank

The goal of this section is to study $\rho(\mathbf{w})$, where $\mathbf{w}$ is a non-zero vector that is orthogonal to $X_1, \cdots, X_n$, where the X_i are iid Bernoulli. We will approach this using two methods. First, we will apply the inverse Littlewood-Offord machinery from Chapter 3 to arrive at a polynomial bound. Then we will improve this bound by utilizing the **ULCD** structure from Chapter 5.

Proposition 6.1.1 (Normal vectors do not have additive structure, iid matrices)**.** *Let $C > 1/2$ and $0 < \varepsilon < 1$ be given. Assume that $X_1, \ldots, X_{n-d}$ are independent Bernoulli vectors, where $d \leq cn$ for some sufficiently small $c > 0$, and assume that $p \leq \exp(n^{1-\varepsilon})$. Let $\mathbf{w}$ be any non-zero vector that is orthogonal to $X_1, \ldots, X_{n-d}$. Then with probability at least $1 - \exp(-\Theta(n))$, we have*

$$\rho(\mathbf{w}) \leq O(n^{-C}),$$

where the implied constant depends on C and ϵ.

We will then improve the previous polynomial bound using a result that says the random normal vector will not have small **ULCD** :

Proposition 6.1.2 (Normal vectors have large ULCD, iid matrices)**.** *Assume that* $p \le \exp(c\kappa^2)$ *and that* $\kappa = n^c$ *for* $c < 1/16$*. Assume that* $X_1, \dots, X_{n-1}$ *are independent Bernoulli vectors. Let* $\mathbf{w}$ *be any non-zero vector that is orthogonal to* $X_1, \dots, X_{n-1}$*. Then with probability at least* $1 - \exp(-\Theta(n))$*, we have*

$$\rho(\mathbf{w}) \le \exp(-c'\kappa^2)$$

with some c' *depending on* c *and* γ*.*

As a corollary of Propositions 6.1.1 and 6.1.2, we then deduce the following, which is a weak form of [11, 14].

Theorem 6.1.3. *Assume that* $p \le \exp(c\kappa^2)$ *and that* $\kappa = n^c$ *for* $c < 1/16$*. We have*

$$\Big|\mathbf{P}\big(X_n \in W_{n-1} | \text{rank}(W_{n-1}) = n-1\big) - 1/p\Big| \le \exp(-n^c),$$

where W_{n-1} *is the subspace generated by* $X_1, \dots, X_{n-1}$*.*

To prove Propositions 6.1.1 and 6.1.2, we first need a lemma saying that all such normal vectors cannot be too sparse.

Lemma 6.1.4. *There exists an absolute constant* $c_{nsp} > 0$ *such that with probability* $1 - \exp(-\Theta(n))$ *any normal vector* $\mathbf{w}$ *of span*$(X_1, \dots, X_{n-d})$*, where* $d \le cn$ *for some sufficiently small* $c > 0$*, satisfies* (5.1).

In order to prove this, we need a quick lemma about the probability of linear dependence among columns of a random matrix:

Lemma 6.1.5 (Odylzko's Lemma)**.** *Let H be a deterministic subspace of $\mathbf{F}_p^n$ of dimension d, and let X be a random Bernoulli ± 1 vector in $\mathbf{F}_p^n$. Then*

$$\mathbf{P}(X \in H) \leq (1/2)^{n-d}$$

Proof. We view the event that $X = (x_1, \cdots, x_d, x_{d+1}, \cdots, x_n)$ belongs to H. Without loss of generality, we may write H as the span of $h_1, \cdots, h_d$ with each $h_i \in \mathbf{F}_p^n$ and such that the first d rows achieve the full rank d. So we can find $c_1, \cdots, c_d$ such that

$$(x_1, \cdots, x_d) = \sum_{i=1}^{d} c_i (h_{i1}, \cdots, h_{id}).$$

Now when we look to the event that our full vector $X = (x_1, \cdots, x_d, x_{d+1}, \cdots, x_n)$ belongs to H, our coefficients are determined and for each $d+1 \leq j \leq n$ we have

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

where the c_i are as before. The probability that each x_j meets this value is at most 1/2, and so the result follows. □

Proof of Lemma 6.1.4. Suppose $\mathbf{w}$ is a member of $\operatorname{span}(X_1, \ldots, X_{n-d})$ with $\operatorname{support}(\mathbf{w}) = k$, where $k = c_1 n$ with c_1 is to be chosen sufficiently small later. At a loss of $\binom{n}{k}$, we may assume the first k entries are nonzero. This means we can restrict the matrix M_n to the first k entries of each column, a new matrix of size $k \times (n-d)$. Since $\mathbf{w}$ is in the span of these columns, we know the rank of our new matrix is at most $k-1$. At a loss of k, we assume the first row is in the span of the remaining $k-1$ rows. But this subspace has codimension at least $(n-d)-(k-1) = n-d-k-1$. Accounting for our loss and applying Odylzko's lemma, the probability of this event is bounded above by

$$\begin{aligned}\binom{n}{k}(k)(1/2)^{n-d-k+1} &\leq \binom{n}{k}(k)(1/2)^{(1-c-c_1)n+1} \\ &\leq (\frac{ne}{c_1 n})^{c_1 n}(c_1 n)(1/2)^{(1-c-c_1)n+1} \\ &= (\frac{e}{c_1})^{c_1 n}(c_1 n)(1/2)^{(1-c-c_1)n+1} \\ &= (\frac{c_1 n}{2})[(\frac{e}{c_1})^{c_1}(1/2)^{1-c-c_1}]^n \\ &\ll \exp{(-\Theta(n))},\end{aligned}$$

where in the last step we use the fact that $c_1^{c_1}$ tends to one as c_1 tends to zero, in order to choose c_1 sufficiently small, which is enough to complete the proof. □

Proof of Proposition 6.1.1. Using the result of the previous lemma, we may assume that our vector $\mathbf{w}$ is non-sparse at a loss of $\exp{(-\Theta(n))}$ in probability. Suppose that

$$\rho := \rho(\mathbf{w}) \geq n^{-C}$$

for $C \geq 1/2$. Then by the inverse Littlewood-Offord result mod p, Theorem 3.2.6, we know that we can find a generalized arithmetic progression P of rank $O(1)$ in $\mathbf{F}_p$ such that $|P| \leq \min\{\rho^{-1}/n^{\varepsilon}, p/n^{\varepsilon}\}$ and P contains all but $n^{2\varepsilon}$ entries of $\mathbf{w}$. Since each generalized arithmetic progression is determined by its generators and dimensions, the number of ways to choose such a P can be bounded above by

$$p^{O(1)}\rho^{-O(1)}.$$

Given a P, the number of vectors $\mathbf{w}$ whose $n - n^{2\varepsilon}$ components are from P is at most

$$\binom{n}{n^{2\varepsilon}}|P|^{n-n^{2\varepsilon}}p^{n^{2\varepsilon}} \leq 2^n(\rho^{-1}/n^{\varepsilon})^{n-n^{2\varepsilon}}p^{n^{2\varepsilon}} \leq (\rho^{-1}/n^{\varepsilon})^{n-1},$$

where in the final inequality we used the fact that $p \leq \exp(n^{1-\varepsilon})$. Given $\mathbf{w}$ for which $\rho(\mathbf{w}) \geq n^{-C}$, the probability that $\mathbf{w}$ is orthogonal to $X_1, \dots, X_{n-d}$ is bounded by

$$(\rho^{-1}/n^{\varepsilon})^{n-1}\rho^{n-1} \leq n^{-\varepsilon n/2},$$

as long as $d \leq cn$ for small positive constant c. The result follows by taking a union bound over choices for P. □

Remark 6.1.6. *It's important to note that while all of the previous results of Chapter 5 and Chapter 6 were stated for Bernoulli random variables, they can also be proven for $\alpha-$balanced ensembles. The general $\alpha-$balanced case in Chapter 5 can be treated* *almost identically, by using the symmetrization trick in the proof of Theorem 3.2.6. The statement of Theorem 6.1.2, however, does not solely extend to the $\alpha-$balanced case, as we will see; we eventually pass to a net of vectors in $\mathbf{R}^n$, and this net will be too large if our matrix has norm larger than $O(\sqrt{n})$.*

We will now prove Proposition 6.1.2. We will show that with high probability, there does not exist $\mathbf{w}$ in the fiber of $t\mathbf{w}$ such that $\kappa/2 \leq \|\mathbf{w}'\|_2 < \kappa$ and that

$$\kappa^{1+(1/4-c)} < \mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \leq \exp(c'\kappa^2/2)/\kappa.$$

To do this, we divide this range into $O(\kappa^2)$ dyadic intervals $(D_i, D_{i+1} = 2D_i)$. For $D = D_i$, let

$$S_D = \Big\{\mathbf{w}' = (w_1/p, \ldots, w_n/p) : \kappa/2 \leq \|\mathbf{w}'\|_2 < \kappa \wedge D \leq \mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \leq 2D\Big\}.$$

Lemma 6.1.7 (Size of the approximating net)**.** *Let $c_0 > 0$ be given sufficiently small compared to c (where $\kappa = n^c$). S_D accepts a $O(\kappa/D)$-net $\mathcal{N}$ of size $D(C\kappa D/\sqrt{n})^n$ if $\kappa D \geq c_0\sqrt{n}$ and of size $Dn^{c_0 n}$ if $\kappa D < c_0\sqrt{n}$ and such that $\mathcal{N} \subset S_D$.*

Before proving this result by following [22], let us introduce a fact that will be useful to our nets.

Fact 6.1.8. *Assume that $\mathcal{S}$ accepts an δ-net $\mathcal{U}$ of size $|\mathcal{N}|$, then S also accepts a 2δ-net $\mathcal{U}'$ such that $\mathcal{U}' \subset \mathcal{S}$ and which has size at most $|\mathcal{N}|$.*

Proof. By throwing away vectors from $\mathcal{U}$ if needed, we assume that each $\mathbf{u} \in \mathcal{U}$ δ-approximates at least one vector $\mathbf{s}'$ from $\mathcal{S}$. Let $\mathcal{N}'$ be a collection of such $\mathbf{s}'$

(we choose an arbitrary $\mathbf{s}'$ from $\mathcal{S}$ that is δ-approximated by any $\mathbf{u}$.) Thus $\mathcal{N}' \subset \mathcal{S}$ and $|\mathcal{N}'| \leq |\mathcal{N}|$. Now for any $\mathbf{s} \in \mathcal{S}$, there exists $\mathbf{u} \in \mathcal{U}$ such that $\|\mathbf{u} - \mathbf{s}\|_2 \leq \delta$, and also by definition there also exists $\mathbf{s}' \in \mathcal{U}'$ such that $\|\mathbf{u} - \mathbf{s}'\|_2 \leq \delta$. Thus we have $\|\mathbf{s} - \mathbf{s}'\|_2 \leq 2\delta$, so $\mathcal{U}'$ is a 2δ-net of $\mathcal{S}$. □

Proof of Lemma 6.1.7. By taking union bound over a small number of choices (at most $O(\kappa \times (D/\kappa)) = O(D)$ choices) we assume that for some $T \in \kappa/D \cdot \mathbf{Z}$ we have

$$T - \kappa/D \leq \|\mathbf{w}'\|_2 \leq T + \kappa/D.$$

By definition, as $\|L\mathbf{w}'\|_{\mathbf{R}/\mathbf{Z}} \leq \kappa$ and $D \leq L \leq 2D$, there exists $\mathbf{p} \in \mathbf{Z}^n$ such that

$$\|L\mathbf{w}' - \mathbf{p}\|_2 \leq \kappa.$$

This implies that

$$\left\|\mathbf{w}' - \frac{\mathbf{p}}{L}\right\|_2 \leq \frac{\kappa}{L} \leq \frac{\kappa}{D},$$

and hence

$$T - 2\frac{\kappa}{D} \leq \frac{\|\mathbf{p}\|_2}{L} \leq T + 2\frac{\kappa}{D}.$$

Thus

$$\begin{aligned}\|\mathbf{w}' - T\frac{\mathbf{p}}{\|\mathbf{p}\|_2}\|_2 &\le \|\mathbf{w}' - \frac{\mathbf{p}}{L}\|_2 + \|T\frac{\mathbf{p}}{\|\mathbf{p}\|_2} - \frac{\mathbf{p}}{L}\|_2 \\ &\le \frac{\kappa}{D} + \|\mathbf{p}\|_2 |\frac{T}{\|\mathbf{p}\|_2} - \frac{1}{L}| \\ &\le 3\frac{\kappa}{D}.\end{aligned}$$

Now as $\|\mathbf{w}'\|_2 < T + \kappa/L$, we also have $\|\mathbf{p}/L\|_2 \le T + 2\kappa/L$ and so

$$\|\mathbf{p}\|_2 \le 2DT + 2\kappa \le 2D(\kappa + \kappa/D) + 2\kappa < 3D\kappa.$$

Let $\mathcal{N}$ be the collection of vectors $T\frac{\mathbf{p}}{\|\mathbf{p}\|_2}$, where T ranges over $O(D)$ choices in the set $\kappa/D \cdot \mathbf{Z}$, and $\mathbf{p}$ ranges over all integer vectors in $\mathbf{Z}^n$ satisfying $\|\mathbf{p}\|_2 \le 3D\kappa$.

Now we bound the size of $\mathcal{N}$ based on the magnitude of κD. **Case 1.** If $\kappa D \ge c_0\sqrt{n}$, then the number of integral vectors $\mathbf{p}$ of norm at most $3\kappa D$ is known to be bounded by $(C\kappa D/\sqrt{n})^n$, and so

$$|\mathcal{N}| \le D(C\kappa D/\sqrt{n})^n.$$

Case 2. If $\kappa D \le c_0\sqrt{n}$, where c_0 is sufficiently small, then all but $O((\kappa D)^2)$ entries of $\mathbf{p}$ are zero. So the number of such vectors $\mathbf{p}$ is bounded by $\binom{n}{(\kappa D)^2}(O(1))^{(\kappa D)^2}$, and so

$$|\mathcal{N}| \le D\binom{n}{(\kappa D)^2}C^{(\kappa D)^2} \le Dn^{c_0 n}.$$

Finally, we can always assume $\mathcal{N}$ to consist of vectors from S_D by using Fact 6.1.8. $\square$

Now we use the obtained net to show that normal vectors in iid matrices cannot have small **ULCD**.

Proof. (of Proposition 6.1.2) Assume otherwise, then by the argument above, by passing to an appropriate $t\mathbf{w}$, we can assume that $\kappa/2 \le \|\mathbf{w}'\|_2 < \kappa$, and that $\mathbf{w}' \in S_D$ for some D_i from $O(\kappa^2)$ dyadic intervals. As $\mathbf{w}$ is orthogonal to $X_1, \dots, X_{n-1}$ in $\mathbf{F}_p$, we then have the following key property for $\mathbf{w}' = \frac{1}{p}w$

$$M\mathbf{w}' \in \mathbf{Z}^{n-1},$$

where M is the $n \times (n-1)$ matrix formed by $X_1, \dots, X_{n-1}$.

By Lemma 6.1.7, there exists $\mathbf{u}' \in \mathcal{N}$ such that

$$\|\mathbf{w}' - \mathbf{u}'\|_2 = O(\kappa/D).$$

It is well known $\|M\| = O(\sqrt{n})$ with probability at least $1 - \exp(-\Theta(n))$, so

$$\|M^T\mathbf{w}' - M^T\mathbf{u}'\|_2 \le O(\sqrt{n}\kappa/D).$$

So

$$\mathrm{dist}(M^T\mathbf{u}', \mathbf{Z}^{n-1}) \le O(\sqrt{n}\kappa/D).$$

Let $\mathcal{E}$ be this event, which will be bounded shortly. By Theorem 5.1.9, as obviously $\kappa/D > 1/D$, we have

$$\begin{aligned}\mathbf{P}(\|X_i \cdot \mathbf{u}'\|_{\mathbf{R}/\mathbf{Z}} = O(\kappa/D)) &= O(\kappa/D + (\log(D/\kappa))(\exp(-\kappa^2) + \exp(-4\gamma^2\|u'\|_2^2)) \\ &= O(\kappa/D), \qquad (6.1)\end{aligned}$$

where in the last estimate we used the fact that $D \leq \exp(c'\kappa^2)$ with sufficiently small c'.

By our tensorization lemma, Lemma 3.3.15, we thus have for some absolute positive constant C'

$$\mathbf{P}(\mathcal{E}) \leq (C'\kappa/D)^{n-1}.$$

Putting this together using union bound over all $\mathbf{u}'$ from the net, as $\kappa = n^c$, we obtain in the case $\kappa D \geq c_0\sqrt{n}$ a bound

$$\begin{aligned}\mathbf{P}(\exists \mathbf{u}' \in \mathcal{N}, \|M\mathbf{u}'\|_{\mathbf{R}/\mathbf{Z}} = O(\sqrt{n}\kappa/D)) &\leq D(C\kappa D/\sqrt{n})^n \times (C'\kappa/D)^{n-1} \\ &\leq D^2(CC'\kappa^2/\sqrt{n})^n \\ &< \exp(-\Theta(n)).\end{aligned}$$

Also, in the second case that $\kappa D < c_0\sqrt{n}$, noting that $D \geq k^{5/4-c}$

$$\begin{aligned}\mathbf{P}(\exists \mathbf{u'} \in \mathcal{N}, \|M\mathbf{u'}\|_{\mathbf{R}/\mathbf{Z}} = O(\sqrt{n}\kappa/D)) &\leq n^{c_0 n} \times (C'\kappa/D)^{n-1}\\ &\leq n^{c_0 n} \times (C'/\kappa)^{(1/4-c)n}\\ &\leq C'^n n^{c_0 n} \times n^{-c(1/4-c)n}\\ &\leq n^{-cn/8},\end{aligned}$$

assuming that c_0 is sufficiently large compared to c, and that $c \leq 1/16$. □

Theorem 6.1.9. *Let $d \geq 1$ be a fixed integer. Assume that $p \leq \exp(c\kappa^2)$ and that $\kappa = n^c$ for $c < 1/16$. We have*

$$\left|\mathbf{P}\big(X_n \in W_{n-d} | \mathrm{rank}(W_{n-d}) = n-d\big) - 1/p^d\right| \leq \exp(-n^c),$$

where W_{n-d} is the subspace generated by $X_1, \dots, X_{n-d}$.

Proof. Suppose that $X_n \in W_{n-d}$. Conditioning on the event that $\mathrm{rank}(W_{n-d}) = n-d$, consider $S := W_{n-d}^{\perp}$, the subspace of vectors orthogonal to W_{n-d}. We will denote the basis of S as $\{\mathbf{v}_1, \cdots, \mathbf{v}_d\}$. Then for each $\mathbf{v}_i$, we have $X_n \cdot \mathbf{v}_i = 0$. Thus we have:

$$\begin{aligned}\mathbf{P}\big(X_n \in W_{n-d}\big) &= \mathbf{P}\big(\wedge_{i=1}^d (X_n \cdot \mathbf{v}_i = 0)\big)\\ &= p^{-d} \sum_{(t_1,\cdots,t_d)\in\mathbf{F}_p^d} \frac{1}{p}\sum_{t\in\mathbf{F}_p} \mathbf{E}[e_p(t[t_1(X_n\cdot\mathbf{v}_1) + \cdots + t_d(X_n\cdot\mathbf{v}_d)])]\\ &= p^{-d} + p^{-d} \sum_{(t_1,\cdots,t_d)\in\mathbf{F}_p^d,\text{ not all zero}} \frac{1}{p}\sum_{t\in\mathbf{F}_p} \mathbf{E}[e_p(t[X_n \cdot \sum_i t_i\mathbf{v}_i])].\end{aligned}$$

Note that the sum

$$\frac{1}{p}\sum_{t\in\mathbf{F}_p}\mathbf{E}[e_p(t[X_n\cdot\sum_i t_i\mathbf{v}_i])]$$

is the concentration probability $\mathbf{P}(X_n\cdot\sum_i t_i\mathbf{v}_i=0)$. We claim that this probability is very small for each $d-$tuple $(t_1,\cdots,t_d)\in\mathbf{F}_p^d$ with not all of the t_i zero. Indeed, by Lemma 6.1.4, we may assume non-sparsity of each normal vector $\sum_i t_i\mathbf{v}_i$. But then, under the condition that each $\sum_i t_i\mathbf{v}_i$ is non-sparse, we know the concentration probability is at most $\exp(-c'\kappa^2)$ by the arguments following our main result Proposition 6.1.2: after applying the tensorization lemma we now replace $(C'\kappa)^{n-1}$ with $(C'\kappa)^{n-d}$, and the result still follows. □

We note that, in finishing the proof of Theorem 6.1.9, we have proven a generalization of Theorem 6.1.2:

Theorem 6.1.10. *Assume that $p\le\exp(c\kappa^2)$ and that $\kappa=n^c$ for $c<1/16$. Let d be fixed (although d can vary slowly with n). Assume that $X_1,\dots,X_{n-d}$ are independent Bernoulli vectors. Let $\mathbf{w}$ be any non-zero vector that is orthogonal to $X_1,\dots,X_{n-d}$. Then with probability at least $1-\exp(-\Theta(n))$, we have*

$$\rho(\mathbf{w})\le\exp(-c'\kappa^2)$$

with some c' depending on c and γ.

Chapter 7

Universality of the Random Normal

Let M_n be a random $n \times n$ matrix with entries that are independent and identically distributed Bernoulli $(0,1)-$random variables, and let p be a prime. Motivated by Chapter 6, it is natural to study the random normal vector over $\mathbf{F}_p$. We now know that the random normal vector does not have structure with high probability. Consider H_{n-1}, the random subspace that is the hyperplane spanned by the first $n-1$ columns, as well as the random normal vector $\mathbf{v}$ that results. When H_{n-1} does not achieve full rank, there are several possible directions that can be chosen for $\mathbf{v}$. So we will instead condition on the event that H_{n-1} has rank $n-1$, which occurs with high probability. Indeed, the **ULCD** result from the previous chapter implies that we can find a $c > 0$ such that $\mathbf{P}(H_{n-1} \text{ has rank } n-1) \geq 1 - \exp(-n^c)$. We note that this can be improved: in [14], Nguyen and Paquette found a $c > 0$ such that $\mathbf{P}(H_{n-1} \text{ has rank } n-1) \geq 1 - \exp(-cn)$. As $\mathbf{v}$ is not unique, since we can alter $\mathbf{v}$ via $t-$dilations for any $t \in \mathbf{F}_p$, we will without loss of generality assume that the last nonzero coordinate of $\mathbf{v}$ is 1 to fix a scaling. We will also assume that $\mathbf{v}$ is non-sparse, which we now know happens with probability at least $1 - \exp(-\Theta(n))$ by Lemma 6.1.4.

For the moment, we will instead assume that the entries of M_n are independent and identically distributed Bernoulli $(-1, 1)$, though it should be noted that this symmetry is not needed until the final section of this chapter.

As before, in this chapter we will assume that our prime $p \leq \exp(c\kappa^2)$ is not too large and $\kappa = n^c$ for $c < 1/16$. We seek to show the following properties involving $\mathbf{v} := (v_1, \cdots, v_n)$ as n grows large:

- For each $i \in \{1, \cdots, n\}$, we can find a constant $c > 0$ such that

$$|\mathbf{P}(v_i = a) - 1/p| \leq O(\exp(-n^c)). \tag{7.1}$$

- For each $i \neq j$ and $a \in \mathbf{F}_p$, we can find a constant $c > 0$ such that

$$|\mathbf{P}(v_i = a \cap v_j = 1) - 1/p| \leq O(\exp(-n^c)). \tag{7.2}$$

- Let n_a denote the number of v_i such that $v_i = a$ and assume that $p \ll n/\log n$. Then for any $0 < \delta < 1$ (which may depend on n), we have

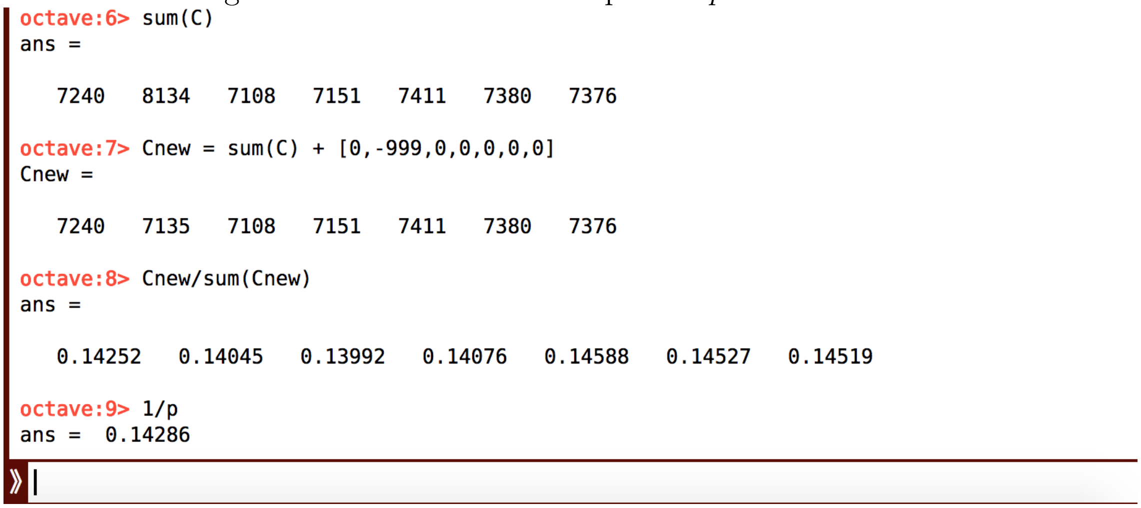
$$\mathbf{P}\big(\cap_{a=0}^{p-1} (|n_a/n - 1/p| \leq \delta/p)\big) \geq 1 - e^{-\delta^2 n/6p} \tag{7.3}$$

- For $\mathbf{f}$, a deterministic binary vector with at least $n - n^{\varepsilon}$ nonzero components, we can find some constant $c > 0$ such that

$$|\mathbf{P}(\langle \mathbf{f}, \mathbf{v} \rangle = a) - 1/p| \leq O(\exp(-n^c)). \tag{7.4}$$

To test the first of the above conjectures, a program was written in OctaveOnline. Each time, the program generates a random $n \times (n-1)$ Bernoulli matrix M_n, calculates a vector $\mathbf{v}$ orthogonal to the $n-1$ columns (with fixed first nonzero coordinate 1), and counts the total instances of each element a mod p among the entries of $\mathbf{v}$. What follows is the output when the program was run 999 times with $n = 50$ and $p = 7$ (here each C is the counting vector, which counts the number of each element of $\mathbf{F}_p$ contained in the normal vector, so that sum(C) collects the number of occurrences of each element in $\mathbf{F}_p$):

Figure 7.1: OctaveOnline Output for $p = 7$ and $n = 50$



```
octave:6> sum(C)
ans =

   7240   8134   7108   7151   7411   7380   7376

octave:7> Cnew = sum(C) + [0,-999,0,0,0,0,0]
Cnew =

   7240   7135   7108   7151   7411   7380   7376

octave:8> Cnew/sum(Cnew)
ans =

   0.14252   0.14045   0.13992   0.14076   0.14588   0.14527   0.14519

octave:9> 1/p
ans =  0.14286
```

We add the vector $(0, -999, 0, 0, 0, 0, 0)$ because of the additional 999 occurrences of the entry "1" based on our assumption. As we see, each component of the new counting vector contains roughly the same count of each element, each being roughly n/p. The code of the preceding program can be found in Appendix D.

In the following sections, we approach the previous four items one at a time.

7.1 Uniformity Among One Entry Meeting a Value

In this section we approach proofs of (7.1) and (7.2). First, we assess the probability that for i fixed, v_i is zero. This quickly results in some linear dependence, the probability of which can be bounded above using previous analysis. But first, we will need a quick lemma about the likelihood of full rank of a special class of rectangular matrices:

Lemma 7.1.1. *Let $0 < \epsilon < 1$ be a given constant and $M_{n\times(n-\epsilon n)}$ be a random matrix as above. Then there exists a positive constant c_ϵ such that with probability at least $1 - O(e^{-n^{c_\epsilon}})$, $M_{n\times(n-\epsilon n)}$ achieves full rank.*

Proof. We apply union bound on the union of the events where $M_{n\times(n-\epsilon n)}$ achieves rank strictly less than $n - \epsilon n$. Decompose this event as $\cup_{i=1}^{n-\epsilon n-1}\mathcal{E}_i$, where $\mathcal{E}_i$ is the event that column c_{i+1} is in the span of the previous i columns and the previous i columns achieve full rank. Through repeated iterations of Odylzko's lemma, we can upper bound $\mathbf{P}(M_{n\times(n-\epsilon n)}$ doesn't achieve full rank) by

$$\sum_{i=1}^{n-\epsilon n-1}(1/2)^{n-i} = 2^{1-n}(2^{n-\varepsilon n-1} - 1) \leq 2^{-\varepsilon n}$$

and the result follows. □

Proposition 7.1.2. *For each $i \in \{1, \cdots, n\}$, we can find a constant $c > 0$ such that $|\mathbf{P}(v_i = 0) - 1/p| \leq O(\exp(-n^c))$.*

Proof. Fix $i = 1$. We seek to bound the probability of the event that our normal vector $\mathbf{v}$ has $v_1 = 0$ under the condition that our first $n-1$ columns achieve full rank, i.e. $\operatorname{rank}(M_{n\times(n-1)}) = n-1$. Suppose we are given such a normal vector. Then restricting to the bottom $n-1$ rows, we see that this is equivalent to the submatrix $M_{(n-1)\times(n-1)}$ having a nontrivial nullspace. So we rewrite $\mathbf{P}(v_1 = 0 \,|\, \operatorname{rank}(M_{n\times(n-1)}) = n-1)$ as $\mathbf{P}(M_{(n-1)\times(n-1)}$ is singular $|\operatorname{rank}(M_{n\times(n-1)}) = n-1)$. We will simply view this as $\mathbf{P}(\operatorname{rank}(M_{(n-1)\times(n-1)}) = n-2 \,|\, \operatorname{rank}(M_{n\times(n-1)}) = n-1) := \mathbf{P}(A|B) = \mathbf{P}(A\cap B)/P(B)$. By conditional expectation and repeated iteration of Theorem 6.1.9 with Lemma 7.1.1, we know that

$$\mathbf{P}(B) = \prod_{i=2}^{\infty}(1-p^{-i}) + O(\exp(-n^{c_1})).$$

Indeed, we choose ϵ sufficiently small and break up our $(n-1)\times(n-1)$ submatrix into its first $n-\epsilon n$ columns and its remaining $\epsilon n-1$ columns. For the $M_{(n-1)\times(n-\epsilon n)}$ matrix, we appeal to Lemma 7.1.1 to deduce that it has full rank with high probability. So it suffices to consider the event that, as we reveal our remaining submatrix column by column, it increases in rank each time. As n tends to infinity, we arrive at the claimed probability via Theorem 6.1.9.

Now consider the event $A \cap B$. This is the event that rows $\mathbf{r}_2, \cdots, \mathbf{r}_n$ span a subspace H of dimension $n-2$ and $\mathbf{r}_1$ is not in the span of H, say $A' \cap B' = \mathbf{P}(A')\mathbf{P}(B'|A')$. For $\mathbf{P}(A')$, we can do as we did in the previous paragraph to show that

$$\mathbf{P}(A') = \prod_{i=2}^{\infty}(1-p^{-i})(1/p)(\frac{1}{1-(1/p)}) + O(\exp(-n^{c_2})).$$

Indeed, we choose ϵ sufficiently small and break up our $(n-1)\times(n-1)$ submatrix into its first $n-\epsilon n$ columns and its remaining $\epsilon n-1$ columns. For the $M_{(n-1)\times(n-\epsilon n)}$ matrix, we appeal to Lemma 7.1.1 to deduce that it has full rank with high probability. So it suffices to consider the event that, as we reveal our remaining submatrix column by column, it increases in rank each time except for one. Through repeated iterations of Theorem 6.1.9, we know this probability is centered around

$$(\frac{1}{p})^{\epsilon n-1}(1-(\frac{1}{p})^{\epsilon n-1})\cdots(1-(\frac{1}{p})^2)+\cdots+(1-(\frac{1}{p})^{\epsilon n-1})(1-(\frac{1}{p})^{\epsilon n-2})\cdots(1-(\frac{1}{p})^2)(\frac{1}{p}),$$

where the i^{th} respective summand is the event that the reveal of the $(i+1)^{st}$ column does not increase the rank of our full matrix. Factoring out common terms, we can rewrite the above as follows:

$$(1-(\frac{1}{p})^{\epsilon n-1})\cdots(1-(\frac{1}{p})^2)[(\frac{1}{p})^{\epsilon n-1}+\cdots+\frac{1}{p}].$$

The desired result for $\mathbf{P}(A')$ quickly follows by factoring out $1/p$ and simplifying the product and geometric sum as n tends to infinity.

For $\mathbf{P}(B'|A')$, our previous section tells us that if we condition on rows $\mathbf{r}_2,\cdots,\mathbf{r}_n$ having rank equal to $n-2$, then our normal vector $\mathbf{v}'$ exists and has large **ULCD**, i.e. $\rho(\mathbf{v}') \leq \exp(-c_3\kappa^2)$. So the probability that $\mathbf{r}_1$ is in the span of H under this condition is centered around $1/p$ with error bound $O(\exp(-n^{c_4}))$.

Putting this all together, we have

$$\begin{aligned}
\mathbf{P}(A|B) &= \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \\
&= \frac{\mathbf{P}(A' \cap B')\mathbf{P}(B'|A')}{\mathbf{P}(B)} \\
&= \frac{\prod_{i=2}^{\infty}(1-p^{-i})(1/p)(\frac{1}{1-1/p})(1-\frac{1}{p})}{\prod_{i=2}^{\infty}(1-p^{-i})} \\
&= \frac{1}{p} + O(e^{-n^{c'}}),
\end{aligned}$$

as desired. □

Remark 7.1.3. *For ease of notation, the error bound analysis in the previous proof will be more precisely shown here. It is not immediately clear that*

$$\frac{\mathbf{P}(A' \cap B')\mathbf{P}(B'|A')}{\mathbf{P}(B)} := \frac{(x+\varepsilon_1)(y+\varepsilon_2)}{(z+\varepsilon_3)}$$

should yield

$$\frac{xy}{z} + \varepsilon_4$$

with $\varepsilon_4 = O(e^{-n^{c'}})$ as we claimed. Of course, for the numerator, we have

$$|(x+\varepsilon_1)(y+\varepsilon_2) - xy| = |y\varepsilon_1 + x\varepsilon_2 + \varepsilon_1\varepsilon_2| \leq O((e^{-n^{c}}))$$

using triangle inequality with the bounds $x, y \leq 1$. *So it now suffices to show that*

$$\frac{xy + \varepsilon'}{z + \varepsilon_3} = \frac{xy}{z} + \varepsilon_4.$$

We have

$$\begin{aligned} \frac{xy + \varepsilon'}{z + \varepsilon_3} &= \frac{xy + \varepsilon'}{z(1 + \varepsilon_3/z)} \\ &= \frac{(xy + \varepsilon')(1 - O(\varepsilon_3/z))}{z} \\ &= \frac{xy + \varepsilon''}{z} \\ &= \frac{xy}{z} + \varepsilon_4 \end{aligned}$$

and so we arrive at an error bound of $\varepsilon_4 = O(e^{-n^{c'}})$ *since* $z > 0$ *is fixed. We will omit this error analysis for future proofs, but the discourse would be analogous.*

Remark 7.1.4. *In the proof of* (7.1)*, we addressed the probability of random matrices achieving certain ranks. It should be noted that this logic can easily be generalized to the following: Let* $0 \leq u \leq d \leq n^c$ *for sufficiently small constant* c. *Then for random Bernoulli matrices we have*

$$\mathbf{P}(\text{rank}(M_{n \times (n-u)}) = n - d)) = \frac{1}{p^{d(d-u)}} \frac{\prod_{i=d+1}^{\infty}(1 - p^{-i})}{\prod_{i=1}^{d-u}(1 - p^{-i})} + O(e^{-n^{c'}}),$$

where c' *is another (sufficiently) positive constant depending on* c. *Results of this type are fairly common and can be directly applied - for instance [14, Theorem A.4] and*

[18, Theorem 5.3] can be utilized in order to improve our sub-exponential bound to an exponential one. This result can also be found using the methods of Maples in [11] and [12].

We now go further and assess the probability that for a fixed i, coordinate v_i meets value $a \in \mathbf{F}_p$, resulting in (7.2). As was the case before, we assume that we are under the condition that $M_{n\times(n-1)}$ has full rank.

Proposition 7.1.5. *For each $i \neq j \in \{1, \cdots, n\}$ and $a \in \mathbf{F}_p$, we can find a constant $c > 0$ such that we have $|\mathbf{P}(v_i = a \cap v_j = 1) - 1/p| \leq O(\exp(-n^c))$.*

Proof. Fix $i = 1$ and $j = 2$. The event that $v_1 = a$ and $v_2 = 1$ is equivalent to the event that $a\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 v_3 \cdots + \mathbf{r}_n v_n = 0$. If a is zero, we are done via the previous proposition, so assume a is nonzero.

Let H be the span of rows $\mathbf{r}_3, \cdots, \mathbf{r}_n$, which has full rank $n-2$ by our rank assumption on $M_{n\times(n-1)}$ and the fact that $a\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 v_3 \cdots + \mathbf{r}_n v_n = 0$.. Further, let π be the projection to the orthogonal complement $H^{\perp}$. For each evaluation of rows $\mathbf{r}_3, \cdots, \mathbf{r}_n$, π is deterministic and $\pi(\mathbf{r}) = \langle \mathbf{r}, \mathbf{n} \rangle$, where $\mathbf{n}$ is the deterministic normal vector. Applying this projection to the linear combination, we have

$$a\langle \mathbf{r}_1, \mathbf{n} \rangle + \langle \mathbf{r}_2, \mathbf{n} \rangle = 0.$$

Since $\mathbf{n}$ is deterministic, each inner product takes values $b, c \in \mathbf{F}_p$ with probability uniformly $1/p$ with error $O(\exp(-n^c))$ by Theorem 6.1.2. This means that a, as the ratio of the inner products, is also uniformly distributed with probability $1/p$ and similar error. □

7.2 Uniformity Among the Total Entries Meeting a Value

In (7.3), we now seek to show that for each $a \in \mathbf{F}_p$, the expected number of coordinates of $\mathbf{v}$ that take value a is roughly n/p. This is an intuitive conjecture based on the previous section, but unfortunately the entries of $\mathbf{v}$ are not independent. Once one coordinate is fixed, it may affect the probabilities that the remaining coordinates take a certain value.

Drawing inspiration from a recent paper by Huang [8], we adapt the following conventions. Let $(n_0, n_1, \cdots, n_{p-1})$ be the vector in $\mathbf{Z}^n_{\geq 0}$ that counts, within $\mathbf{v}$, the total number of evaluations of each element of $\mathbf{F}_p$; that is to say, $n_i := \#\{v_j : v_j = i\}$. We can decompose the $p-$tuples $(n_0, \cdots, n_{p-1})$ into two classes:

- (Equidistributed) The set of $p-$tuples $(n_0, \cdots, n_{p-1}) \in \mathbf{Z}^p_{\geq 0}$ such that

$$|n_j/n - 1/p| \leq \delta/p$$

  for each j.

- (Non-equidistributed) The set of $p-$tuples $(n_0, \cdots, n_{p-1})$ that are neither $(n, 0, 0, \cdots, 0)$ nor equidistributed.

This is relevant since if a $p-$tuple is equidistributed, then each n_j/n is roughly $1/p$, i.e. the number of instances of value j that appear in vector $\mathbf{v}$ is approximately n/p.

In order to prove (7.3), we must show that the contribution of ρ^{n-1} over the set $\mathcal{NE}$ is small, where $\mathcal{NE}$ is the set of non-equidistibuted vectors. More precisely, we can show using Chernoff bound analysis:

Proposition 7.2.1. $|\mathcal{NE}| \leq O(p^n e^{-\delta^2 n/3p})$

Proof. First, we note that each n_a has distribution $\mathrm{Bin}(n, \frac{1}{p})$ with variance

$$n(\frac{1}{p})(1-\frac{1}{p}) = \frac{n(p-1)}{p^2}.$$

Letting $0 < \delta < 1$ and μ denote the mean of this distribution, the upper-tail and lower-tail Chernoff inequalities combine to form the following bound:

$$\mathbf{P}(|n_a - \mu| \geq \delta\mu) \leq 2e^{-\mu\delta^2/3}.$$

For each a, let F_a denote the event that $|n_a - \frac{n}{p}| < \delta n/p$. Then

$$\mathbf{P}(F_0 \cap \cdots \cap F_{p-1}) \geq 1 - 2pe^{-\delta^2 n/3p}.$$

Since there are p^n different choices for $\mathbf{v}$ yet everything is preserved under $t-$dilations for each $t \in \{0, 1, \cdots, p-1\}$, this means that the number of non-equidistributed vectors $\mathbf{v}$ is at most $O(p^{n-1}pe^{-\delta^2 n/3p})$. □

We now have enough to prove (7.3).

Proof. We seek to upper bound $\mathbf{P}(\mathbf{v} \text{ is normal and } \mathbf{v} \in \mathcal{NE})$. Immediately we have $\mathbf{P}(\mathbf{v} \text{ is normal and } \mathbf{v} \in \mathcal{NE})$ is bounded above by

$$\sum_{\mathbf{v} \in \mathcal{NE}} \mathbf{P}(\mathbf{v} \perp X_1, \cdots, X_{n-1}) \leq \sum_{\mathbf{v} \in \mathcal{NE}} (\rho(\mathbf{v}))^{n-1}.$$

Similar to our previous sections, we may decompose the sum into classes where $\mathbf{v}$ is sparse and $\mathbf{v}$ is non-sparse. By Lemma 6.1.4, the contribution over our sparse vectors is negligible. For our non-sparse vectors, we appeal to Theorem 6.1.2. Let $(\mathcal{NE})'$ be the set of non-equidistributed vectors that are non-sparse. We can now bound the desired sum via:

$$\begin{aligned} \sum_{v \in (\mathcal{NE})'} (\rho(\mathbf{v}))^{n-1} &\leq \sum_{v \in (\mathcal{NE})'} (1/p + e^{-n^c})^{n-1} \\ &\leq \frac{p^n}{e^{\delta^2 n/3p}} (1/p + e^{-n^c})^{n-1} \\ &= \frac{1}{e^{\delta^2 n/3p}} (1 + pe^{-n^c})^{n-1} \\ &\leq \frac{1}{e^{\delta^2 n/6p}}, \end{aligned}$$

completing the proof as long as $p \ll n/\log n$. □

7.3 The Dot Product for Nonsparse Vectors

We can treat (7.4) with the motivation gained from Theorem 2.1.5. Drawing inspiration from the previous sections, suppose $\mathbf{f}$ is a fixed, deterministic vector that meets the following condition:

Condition 7.3.1. *Let* $\mathbf{f}$ *be a binary vector such that at least* $n - n^{\epsilon}$ *components are nonzero for some* $\epsilon > 0$*. That is,* $|\mathbf{supp}(\mathbf{f})| \geq n - n^{\epsilon}$*.*

For such an $\mathbf{f}$, a fixed binary vector with enough nonzero entries, we have a bound similar to the preceding sections:

Theorem 7.3.1. *With* $\mathbf{f}$ *as above, we can find an absolute constant* $c > 0$ *such that* $|\mathbf{P}(\langle \mathbf{f}, \mathbf{v} \rangle = a) - 1/p| \leq O(e^{-n^{c}})$

We will take an approach similar to Lemma 4.1.1, but in our mod p setting. As we will see, the same result holds and we can thus invoke extra randomness.

Lemma 7.3.2. *Suppose that we have a fixed evaluation of* M_n*, and let* $\mathcal{E}$ *be the event that this evaluation is reached. Let* $\mathbf{v}'$ *be the vector* $(\epsilon_1 v_1, \cdots, \epsilon_n v_n)$ *where each* ϵ_i *is independent and identically distributed Bernoulli* ± 1*. Then conditioning under the event* $\mathcal{E}$*, the law that* $\mathbf{v}$ *observes is the same as* $\mathbf{v}'$*. That is,*

$$\mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}) = \mathbf{P}(\langle \mathbf{v}', \mathbf{f} \rangle = a).$$

Proof. Fix a deterministic evaluation of the random matrix M_n. We can invoke the extra randomness via the following: suppose we multiply each row, say row i, independently by the Bernoulli random variable ϵ_i, and consider the new normal vector

$\mathbf{v}'_i$ that results.. It is easy to see that if $\mathbf{v}$ is the original normal vector orthogonal to each column of M_n and ϵ_i evaluates to 1, then $\mathbf{v}'_i = \mathbf{v}_i$. Similarly, if ϵ_i evaluates to -1, then $\mathbf{v}'_i = -\mathbf{v}_i$. Thus we may now consider the law under the new normal vector $\mathbf{v}'$. □

We can now invoke the extra randomness and use the previous sections to arrive at a nice result:

Lemma 7.3.3. *Let* $\mathbf{f}$ *and* $\mathbf{v}$ *be as above. Then there exists a* $c_1 > 0$ *such that*

$$\mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a) \leq \sum_{i:(\mathbf{vf})' \in \mathcal{NS}} \mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}_i)\mathbf{P}(\mathcal{E}_i) + \exp(-c_1 n),$$

where $\mathcal{NS}$ *denotes the event that a vector is non-sparse.*

Proof. We begin by breaking up the event $\langle \mathbf{v}, \mathbf{f} \rangle = a$ into the following sum, where each $\mathcal{E}_i$ is one of the possible evaluations of the matrix M_n:

$$\mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a) = \sum_i \mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}_i)\mathbf{P}(\mathcal{E}_i).$$

Let $\mathbf{v}', \mathbf{f}'$ be the truncated vectors where we restrict $\mathbf{v}$ and $\mathbf{f}$ to the $n - n^{\epsilon}$ components where $\mathbf{f}$ is nonzero. Furthermore, consider the vector $(\mathbf{vf})' := (\mathbf{v}'_i \mathbf{f}'_i)$ where the components consist of the entry-wise product of components of $\mathbf{v}'$ and $\mathbf{f}'$. We can further decompose the previous sum depending on whether $(\mathbf{vf})'$ is sparse (say $(\mathbf{vf})' \in \mathcal{S}$) or

$(\mathbf{vf})'$ is non-sparse (say $(\mathbf{vf})' \in \mathcal{NS}$):

$$\mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a) = \sum_{i:(\mathbf{vf})' \in \mathcal{NS}} \mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}_i)\mathbf{P}(\mathcal{E}_i) + \sum_{i:(\mathbf{vf})' \in \mathcal{S}} \mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}_i)\mathbf{P}(\mathcal{E}_i).$$

For the second sum, we bound each $\mathbf{P}(\langle \mathbf{v}, \mathbf{f} \rangle = a \,|\, \mathcal{E}_i)$ above by 1. The remaining sum gives the probability that $(\mathbf{vf})'$ is sparse, which can be bounded above by $\exp(-c_1 n)$ for some c_1. Indeed, since each f_i' is nonzero, it suffices to show that most of the v_i' are nonzero, which is Lemma 6.1.4. □

So it now suffices to consider the sum in our previous lemma. We can use Lemma 7.3.2 to upper bound the summand by

$$\mathbf{P}(\sum_j \epsilon_j f_j v_j = a)\mathbf{P}(\mathcal{E}_i),$$

for each non-sparse $(\mathbf{vf})'$. We wish to show that $\mathbf{P}(\sum_j \epsilon_j f_j v_j = a) = 1/p + O(e^{-n^c})$. We will assume for contradiction, similar to previous sections, that for some $\mathbf{w} := (\mathbf{vf})' \in \mathcal{NS}$, we have

$$|\mathbf{P}(\sum_j \epsilon_j f_j v_j = a)) - 1/p| \gg e^{-n^c}.$$

As it turns out, in this instance we can find a dilation similar to our previous chapter.

Lemma 7.3.4. *There exists a dilation $t\mathbf{w}$ of $\mathbf{w}$ such that $||\mathbf{w}'||_2 < \kappa$,*

$$\mathbf{ULCD}(\mathbf{w}') := L = O(\frac{e^{n^c}}{||\mathbf{w}'||_2}).$$

Furthermore, for such a $\mathbf{w}'$ *we also have*

$$\mathbf{ULCD}(\mathbf{w}') \geq \kappa^{5/4-c}.$$

Combining, we have

$$\kappa^{5/4-c} \leq \mathbf{ULCD}(\mathbf{w}') \leq \exp(n^c).$$

Proof. This lemma follows immediately by Corollary 5.1.5 and Corollary 5.1.13. □

As we did in Chapter 6, we now wish to show that this occurs with exponentially small probability. As $\mathbf{w}'$ is determined from $(\mathbf{vf})'$, the ability to do so will depend on the structure of $\mathbf{f}$. Note that if $\mathbf{f} = (1, 1, \cdots, 1)$, this is precisely the result of the previous section. Thus we can immediately say that

Theorem 7.3.5. *For* $\mathbf{f} = (1, 1, \cdots, 1)$*, there exists an absolute constant* $c > 0$ *such that* $|\mathbf{P}(\langle \mathbf{f}, \mathbf{v} \rangle = a) - 1/p| \leq O(e^{-n^c})$.

The next natural case to consider is when $\mathbf{f}$ consists of only 1's and 0's, resulting in (7.4):

Theorem 7.3.6. *For a binary vector* $\mathbf{f}$ *that satisfies our condition, we can find an absolute constant* $c > 0$ *such that* $|\mathbf{P}(|\langle \mathbf{f}, \mathbf{v} \rangle| = a) - 1/p| \leq O(e^{-n^c})$.

Proof. The proof follows by the methods in the first section of this chapter: dividing

$$\kappa^{1+(1/4-c)} < \mathbf{ULCD}_{\gamma,\kappa}(\mathbf{w}') \leq \exp(c'\kappa^2/2)/\kappa$$

into $O(\kappa^2)$ dyadic intervals and approximating via a sufficiently small net, we can thus show that the existence of such a $\mathbf{w} := (\mathbf{v}\mathbf{f})' \in \mathcal{N}\mathcal{S}$ with

$$|\mathbf{P}(\sum_j \epsilon_j f_j v_j = a)) - 1/p| \gg e^{-n^c}$$

occurs with extremely small probability, completing the proof. □

We now apply the result for $\mathbf{f} = (1, 1, \cdots, 1)$ and for binary vectors in order to prove our desired theorem for general $\mathbf{f}$ that meet our condition in (7.4).

Appendix A

Proof of Theorem 2.2.2

In this section, we prove that uncontrollability of a pair $(A, \mathbf{b})$ is equivalent to the existence of an eigenvector $\mathbf{v}$ of A such that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$.

The backward direction follows almost immediately. Indeed, if we can find an eigenvalue-eigenvector pair $(\lambda, \mathbf{v})$ of A such that $\mathbf{v}^T\mathbf{b} = 0$, then for each k, we have $\mathbf{v}^T A^k \mathbf{b} = \lambda^k \mathbf{v}^T \mathbf{b} = 0$. Letting A' denote the controllability matrix in Definition 2.2.1, we have that $\mathbf{v}^T A' = 0$ and thus A' is uncontrollable.

For the forward direction, suppose that each eigenvector $\mathbf{v}$ satisfies $\mathbf{v}^T\mathbf{b} \neq 0$. Then each eigenspace of A has dimension one (if we can find an eigenspace of dimension at least 2, then considering the intersection of that eigenspace with the orthogonal complement of the subspace spanned by $\mathbf{b}$ leads us to an eigenvector $\mathbf{v}$ such that $\mathbf{v}^T\mathbf{b} = 0$). Since A is symmetric, it thus follows that the eigenvalues are distinct so that A has simple spectrum. Now suppose that the spectrum of A is simple and assume that $(A, \mathbf{b})$ is uncontrollable, i.e. we can find a nonzero vector $\mathbf{a} =$

$(a_0, \cdots, a_{n-1})$ such that $A'\mathbf{a} = 0$, where

$$A' = (\mathbf{b} \quad A\mathbf{b} \quad \cdots \quad A^{n-1}\mathbf{b})$$

is our controllability matrix. Further suppose that our eigenvalue-eigenvector pairs are denoted $(\lambda_i, \mathbf{v}_i)$ with $\lambda_1 < \cdots < \lambda_n$. We begin to use the spectral theorem to decompose each $A^k\mathbf{b}$ as

$$A^k\mathbf{b} = \sum_{j=1}^{n} (\lambda_j^k \mathbf{v}_j^T \mathbf{b})\mathbf{v}_j.$$

Since $A'\mathbf{a} = 0$, we have that

$$0 = A'\mathbf{a} = \sum_{k=0}^{n-1} a_k A^k \mathbf{b} = \sum_{k=0}^{n-1} a_k \sum_{j=1}^{n} (\lambda_j^k \mathbf{v}_j^T \mathbf{b})\mathbf{v}_j = \sum_{j=1}^{n} \mathbf{v}_j (\sum_{k=0}^{n-1} \mathbf{v}_j^T \mathbf{b} \lambda_j^k a_k).$$

Letting

$$\beta_j = \sum_{k=0}^{n-1} \mathbf{v}_j^T \mathbf{b} \lambda_j^k a_k,$$

we have that each $\beta_j = 0$ by linear independence of our eigenbasis. Write

$$\beta_j = \mathbf{v}_j^T \mathbf{b} \begin{pmatrix} 1 \\ \lambda_j^1 \\ \vdots \\ \lambda_j^{n-1} \end{pmatrix}^T \mathbf{a}.$$

Since each $\mathbf{v}_j^T \mathbf{b} \neq 0$ by assumption, it must then be the case that

$$\begin{pmatrix} 1 \\ \lambda_j^1 \\ \vdots \\ \lambda_j^{n-1} \end{pmatrix}^T \mathbf{a} = 0.$$

But this implies that the Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}^T$$

is singular, and hence $\lambda_i = \lambda_j$ for some $i \neq j$, a contradiction.

Appendix B

Proof of Theorem 3.2.6

In order to prove this theorem, we will need to quickly introduce a more general notion of structure in additive groups, that of a coset progression:

Definition B.0.1. Let P be a set in a finite additive group G. We will say P is a coset progression of rank r if we can write

$$P = H + Q$$

where H is a finite subgroup of G and $Q = \{a_0 + x_1 a_1 + \cdots + x_r a_r | M_i \leq x_i \leq M_i', x_i \in \mathbf{Z}\}$ is a GAP of rank r. We will say that P with respect to this presentation is *proper* if each element of $H + Q$ is distinct. More generally, given a positive integer t, we will say that P with respect to this presentation is *t-proper* if $H + tQ$ is proper. Finally, we will say that P with this presentation is *symmetric* if the GAP Q is symmetric.

To prove Theorem 3.2.6 we will make use of two results from [27] by Tao and Vu. The first result allows one to pass from coset progressions to proper coset progressions without any substantial loss.

Theorem B.0.2. *[27, 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 B.0.3. *[27, Theorem 1.21] There exists a positive integer C_2 such hat 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* $0 \le d' \le d-1$ *and size* $|H+Q| \ge 2^{-2^{C_2 d^2 2^{6d}}} k^{d'} |X|$ *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 $Q = \{a_0 + x_1 a_1 + \cdots + x_r a_r | -N_i \le x_i \le N_i \text{ for all } 1 \le i \le r\}$ is contained in a symmetric GAP $Q' = \{x_0 a_0 + x_1 a_1 + \cdots + x_r a_r | -1 \le x_0 \le 1, -N_i \le x_i \le N_i \text{ for all } 1 \le i \le r\}$. Thus, by combining Theorem B.0.3 with Theorem B.0.2 we obtain the following

Corollary B.0.4. *Whenever* $d, k \ge 1$ *and* $X \subset G$ *is a non-empty finite set such that*

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

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

$$kX \subset H+P.$$

Proof. (of Theorem 3.2.6) First, for convenience we will pass to symmetric distributions. Let $\psi = \mu - \mu'$ be the symmetrization 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. We can check that $\max_x \mathbf{P}(\psi = x) \leq 1 - \alpha_n$. Indeed,

$$\mathbf{P}(\psi = x) = \sum_{c \in \mathbf{F}_p} \mathbf{P}(\psi = c)\mathbf{P}(\psi = c - x) \leq \sum_{c \in \mathbf{F}_p} \mathbf{P}(\psi = c)(1 - \alpha_n) = (1 - \alpha_n)$$

so 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 $\mathbf{P}(\psi' = t_j) = \mathbf{P}(\psi' = -t_j) = \beta_j/2$ for $1 \leq j \leq l$, and that $\mathbf{P}(\psi' = 0) = \beta_0$, where $t_{j_1} \pm t_{j_2} \neq 0 \mod p$ for all $j_1 \neq j_2$. Denote $S = \mu_1 w_1 + \cdots + \mu_n w_n$. Consider $a \in \mathbf{Z}/p\mathbf{Z}$ where $\mathbf{P}(S = a)$ is maximum (or minimum). Using the standard notation $e_p(x)$ for $\exp(2\pi\sqrt{-1}x/p)$, we have

$$\begin{aligned}
\mathbf{P}(S = a) &= \mathbf{E}\frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e_p(x(S - a)) \\
&= \mathbf{E}\frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e_p(xS)e_p(-xa) \\
&= \frac{1}{p} + \mathbf{E}\frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}, x \neq 0} e_p(xS)e_p(-xa).
\end{aligned}$$

So

$$\rho = |\mathbf{P}(S = a) - \frac{1}{p}| \leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}, x \neq 0} |\mathbf{E}e_p(xS)|. \tag{B.1}$$

By independence

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

It follows that

$$\begin{aligned}\rho &\leq \frac{1}{p} |\sum_{x \in \mathbf{Z}/p\mathbf{Z}, x \neq 0} \prod_{i=1}^{n} (\beta_0 + \sum_{j=1}^{l} \beta_j \cos \frac{2\pi x t_j w_i}{p})| \\ &\leq \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}, x \neq 0} \prod_{i=1}^{n} (\beta_0 + \sum_{j=1}^{l} \beta_j |\cos \frac{\pi x t_j w_i}{p}|),\end{aligned} \tag{B.2}$$

where we made the change of variable $x \to x/2$ (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,

$$|\cos \frac{\pi x}{p}| \leq 1 - \frac{1}{2} \sin^2 \frac{\pi x}{p} \leq 1 - 2\|\frac{x}{p}\|^2. \tag{B.3}$$

Hence for each w_i

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

Consequently, we obtain a key inequality

$$\begin{aligned}\rho &\le \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}, x \neq 0} \prod_{i=1}^{n} (\beta_0 + \sum_{j=1}^{l} \beta_j |\cos \frac{\pi x t_j w_i}{p}|) \\ &\le \frac{1}{p} \sum_{x \in F_p, x \neq 0} \exp(-2 \sum_{i=1}^{n} \sum_{j=1}^{l} \beta_j \|\frac{x t_j w_i}{p}\|^2).\end{aligned}$$

Large level sets. Now we consider the level sets

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

We have

$$n^{-C} \le \rho \le \frac{1}{p} \sum_{x \in F_p, x \neq 0} \exp(-2 \sum_{i=1}^{n} \sum_{j=1}^{l} \beta_j \|\frac{x t_j w_i}{p}\|^2) \le \frac{1}{p} \sum_{m \ge 1} \exp(-2(m-1)) |S_m|.$$

As $\sum_{m \ge 1} \exp(-m) < 1$, there must be is a large level set S_m such that

$$|S_m| \exp(-m+2) \ge \rho p. \tag{B.4}$$

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

We will now break up our proof into the two different cases depending on the size of p:

Case 1. Suppose that $p \ll n^C$. Since S_m is non-empty, choose a non-zero x_0 in S_m with

$$\sum_{i=1}^{n} \sum_{j=1}^{l} \beta_j \| \frac{x_0 t_j w_i}{p} \|^2 \leq m.$$

Then we have that most of the w_i satisfy

$$\sum_{j=1}^{l} \beta_j \| \frac{x_0 t_j w_i}{p} \|^2 \leq \frac{m}{n'} \tag{B.5}$$

Indeed, the set of w_i satisfying (B.5) has size at least $n - n'$. Let W' denote the members of this set. We claim that this is the W' in the conclusion of the theorem. Choose j_0 so that t_{j_0} is non-zero. Then as β_{j_0} is a positive constant, we have

$$\| \frac{x_0 t_{j_0} w_i}{p} \|^2 = O(\frac{m}{n'}) \tag{B.6}$$

Consider the rank one and size $O(p\sqrt{(\log n)/n'})$ arithmetic progression

$$P = \{x \in \mathbf{F}_p, \| \frac{x}{p} \| = O(\sqrt{\frac{m}{n'}})\}.$$

Indeed, after dilating by $x_0 t_{j_0}$, we see that W' belongs to P, as desired.

Case 2. Suppose that $p \gg n^C$. By double counting we have

$$\sum_{i=1}^{n} \sum_{x \in S_m, x \neq 0} \sum_{j=1}^{l} \beta_j \| \frac{x t_j w_i}{p} \|^2 = \sum_{x \in S_m} \sum_{i=1}^{n} \sum_{j=1}^{l} \beta_j \| \frac{x t_j w_i}{p} \|^2 \leq m|S_m|.$$

So, for most v_i

$$\sum_{x \in S_m, x \neq 0} \sum_{j=1}^{l} \beta_j \| \frac{x t_j w_i}{p} \|^2 \leq \frac{m}{n'} |S_m| \tag{B.7}$$

for some large constant C_0.

By averaging, the set of w_i satisfying (B.7) has size at least $n - n'$. We call this set W'. The set $\{w_1, \dots, w_n\} \backslash W'$ has size at most n' and this is the exceptional set that appears in Theorem 3.2.6. 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 $a \in kW'$

$$\sum_{x \in S_m, x \neq 0} \sum_{j=1}^{l} \beta_j \| \frac{x t_j a}{p} \|^2 \leq k^2 \frac{m}{n'} |S_m|. \tag{B.8}$$

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

$$\sum_{x \in S_m, x \neq 0} \sum_{j=1}^{l} \beta_j \| \frac{x t_j a}{p} \|^2 \leq k'^2 \frac{m}{n'} |S_m|. \tag{B.9}$$

Dual sets. Set

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

Then by definition of ξ, we have

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

Define

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

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

$$|S_m^*| \leq \frac{8p}{|S_m|}. \tag{B.10}$$

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 $a \in S_m^*$

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

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

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

Equation (B.10) 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 (B.9) 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

$$|kW''| = O(\frac{p}{|S_m|}) = O(\rho^{-1}\exp(-m+2)). \tag{B.11}$$

The long range inverse theorem. We are now in the position to apply Corollary B.0.4 with X as the set of distinct elements of W''. As $k = \Omega(\sqrt{\frac{\alpha_n' n'}{m}}) = \Omega(\sqrt{\frac{\alpha_n' n'}{\log n}})$,

$$\rho^{-1} \le n^C \le k^{4C/\varepsilon+1}. \tag{B.12}$$

It follows from Corollary B.0.4 that kX is a subset of a 2-proper symmetric coset progression $H+P$ of rank $r = O_{C,\epsilon_0}(1)$ and cardinality

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

Now we use the special property of $\mathbf{Z}/p\mathbf{Z}$ that it has only trivial proper subgroup. 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 $r = O_{C,\epsilon_0}(1)$ and cardinality

$$|P| \leq O_{C,\varepsilon}|kX|. \tag{B.13}$$

To this end, we apply the following dividing trick from [16, Lemma A.2].

Lemma B.0.5. *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*

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

Proof. (of Lemma B.0.5) 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 $0 \in X$, $2^{l-1}X \subset 2^l X \subset P$, 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 $0 \leq |2x_i| \leq N_i$. □

Combining (B.13) and Lemma B.0.5 we thus obtain a GAP Q that contains X and

$$|Q| = O_{C,\epsilon_0}(k^{-r}|kX|) = O_{C,\epsilon_0}(k^{-r}|kW''|) = O_{C,\epsilon_0}\left(\rho^{-1}\exp(-m)(\sqrt{\frac{\alpha'_n n'}{m}})^{-r}\right)$$
$$= O_{C,\epsilon_0}(\rho^{-1}(\alpha'_n n')^{-r}),$$

concluding the proof. □

Appendix C

Proof of Cauchy-Davenport

Before proving Cauchy-Davenport, we prove a quick inequality involving the sumset of two subsets of $\mathbf{R}$. When considering the sumset $A+B$, it is natural to ask how the size of this sumset relates to the sizes of the sets A and B. We typically expect the size to be large unless A and B are fairly compatible. The following inequality gives a lower bound on the size of our sumset:

Proposition C.0.1. *Let A, B be two subsets of the abelian group $(\mathbf{R}, +)$. Then $|A+B| \geq |A| + |B| - 1$.*

Proof. Without loss of generality, label the elements of A as $A = \{a_1, \cdots, a_i\}$ with $a_1 < \cdots < a_i$, and the elements of B as $B = \{b_1, \cdots, b_j\}$ with $b_1 < \cdots < b_j$. Then of course

$$a_1 + b_1 < a_1 + b_2 < \cdots < a_1 + b_j < a_2 + b_j < \cdots < a_i + b_j.$$

This gives at least $|A| + |B| - 1$ elements in $|A+B|$. □

It's easy to note that equality in the above lemma results if and only if A and B are both arithmetic progressions with the same difference.

Now suppose that that A and B are non-empty subsets of $\mathbf{F}_p$. It's natural to expect to see something similar over prime fields, and indeed, our inequality is almost unchanged:

Theorem C.0.2 (Cauchy-Davenport Theorem)**.** *Let A and B be non-empty subsets of $\mathbf{F}_p$. Then we have the inequality*

$$|A+B| \geq \min\{|A|+|B|-1, p\}.$$

Proof. We will approach the proof via induction on $|A|$. Of course, if $|A| = 0$ or 1, then the claim clearly holds.

Suppose $|A| > 1$. Note that if $|B| = 0$ or p, then the claim is also clearly true, so suppose not. Then since B is nontrivial, we know from before that $|A+B| > |B|$ and we can find some $b \in B$ such that $A + b \notin B$. Let $A' = \{a \in A : a + b \notin B\}$. We already know that $|A'| \geq 1$.

Consider $A_1 := A \setminus A'$ and $B_1 := B \cup (A' + b)$. Then since $(A' + b) \cap B = \emptyset$, we know that $|B_1| = |A' + b| + |B|$. Also, $|A_1| = |A| - |A'|$. Also by design, we have that $A_1 + B_1 \subset A + B$. Indeed, $(A \setminus A') + (A' + b) \subset A + B$, since if we have some $a \in A \setminus A'$ and a' in A', then $a + a' + b \in A + B$: $a \in A \setminus A'$ means we can choose some some $b' \in B$ such that $a + b = b'$. Then $a + a' + b = a' + b' \in A + B$.

The result follows via induction on $|A|$. □

Appendix D

Relevant Code For Chapter 7

Code D.0.1.

```
# Forms the normal vector v:

function Count = Unif(n,p)

A = randi([0 1], n,n-1);
A = transpose(A);
[L,D,U,rows,cols] = ModLU(A,p);
U;
N = NullPU(U,p);
Count = zeros(1,p);
Count1 = zeros(1,p);
Count2 = zeros(1,p);
if size(N,2) == 1
    if N(1) == 1;
        for i = 0:p-1
            Count(i+1) = sum(N(:) == i);
        end
    else
        N = N - N(1) + 1;
        N = mod(N,p);
        for i = 0:p-1
            Count(i+1) = sum(N(:) == i);
        end
    end
else
    if N(:,1)(1) == 1;
        for i = 0:p-1
            Count1(i+1) = sum(N(:,1) == i);
        end
    else
        N = N - N(1) + 1;
        N = mod(N,p);
        for i = 0:p-1
            Count1(i+1) = sum(N(:,1) == i);
        end
    end
    if N(:,2)(1) == 1;
        for i = 0:p-1
            Count2(i+1) = sum(N(:,2) == i);
        end
    else
        N = N - N(1) + 1;
        N = mod(N,p);
        for i = 0:p-1
            Count2(i+1) = sum(N(:,2) == i);
        end
    end
    Count = Count1 + Count2;
end

end
```

Code D.0.2.

```
# Forms the LU-factorization of A, modulo p:
# A(rows,cols) - mod(L * diag(D)*U,p)

function [L,D,U,rows,cols] = ModLU(A,p)

[m,n] = size(A);

# inverses in mod-p:
# mod(k*invp(k+1)) = 0 if k==0; 1 otherwise

invp = 2:p-2;
for i = 2:p-2; invp = mod(invp.*[2:p-2],p); end
invp = [0,1,invp,p-1];

# Initialize outputs:
L = eye(m); U = A;
rows = 1:m;
cols = 1:n;
```

```
# Sweep
for i = 1:m
    # Pivoting
    [row, col] = find(U(i:end,:));
    if isempty(row); break; end
    row = row(1)+i-1; col = col(1);

    r = 1:m; r(i) = row; r(row) = i;
    c = 1:n; c(i) = col; c(col) = i;
    ri = rows(i); rows(i) = rows(row); rows(row)=ri;
    ci = cols(i); cols(i) = cols(col); cols(col)=ci;

    rinv = 1:m; rinv(r) = 1:m;
    U = U(r,c); L=L(r,r);

    # Gaussian elimination
    L(i+1:end,i      ) = mod(invp(U(i,i)+1) * U(i+1:end,i),p);
    U(i+1:end,i:end) = mod(U(i+1:end,i:end) + (p-L(i+1:end,i)) *
U(i,i:end),p);
end

# Factorize diagonal
D = zeros(m,1); D(1:min(m,n)) = diag(U);
U = mod(diag(invp(D+1)) * U,p  );
```

Code D.0.3.

```
# for an upper triangular matrix U, calculates a base for
the null space  modulo p: U * N = 0.

function N = NullPU(U,p)

n = size(U,2);
rank = size(find(diag(U)),1);
A = U(1:rank,:);
for i=rank:-1:2
     A(1:i-1,:) = mod(A(1:i-1,:) + (p-1) * A(1:i-1,i) * A(i,:),p);
end
N = [mod(p-A(:,rank+1:end),p); eye(n-rank)];
```

Code D.0.4.

```
#Calculates the normal vector v 999 times
```

```
C = Unif(n,p);
for i = 1:999
     C = [C; Unif(n,p)];
end
```

Bibliography

[1] P. Bourgade and H.-T. Yau, The eigenvector moment flow and local quantum unique ergodicity, *Communications in Mathematical Physics*, **350** (2017), no. 1, 231-278.

[2] Y. Dekel, J. Lee, and N. Linial, Eigenvectors of random graphs: nodal domains, *Random Structures & Algorithms*, **39** (2011), no. 1, 39-58.

[3] P. Erdős, On a lemma of Littlewood and Offord, *Bulletin of the American Mathematical Society*, **51** (1945), 898–902.

[4] P. Erdős and L. Moser, Elementary problems and solutions, *Amer. Math. Monthly*, **54** (1947), no. 4, 229–230.

[5] L. Erdős, B. Schlein, and H. T. Yau, Local semicircle law and complete delocalization for Wigner random matrices, *Communications in Mathematical Physics*, **287** 2009, no. 2, 641-655.

[6] L. Erdős, B. Schlein, and H. T. Yau, Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices, *Ann. Probab.*, **37** 2009, no. 3, 815-852.

[7] G. Freiman, Foundations of a Structural Theory of Set Addition, *Translations of Mathematical Monographs*, **37** (1973), Amer. Math. Soc, Providence, RI, USA.

[8] J. Huang, Invertibility of adjacency matrices for random d-regular directed graphs, *arXiv:1806.0138* [math], June 2018

[9] J. Littlewood and A. Offord, On the number of real roots of a random algebraic equation (III), *Rec. Math. (Mat. Sbornik) N.S,* **54** (1943), 277-286.

[10] K. Luh, S. Meehan, and H. Nguyen, Random matrices in finite field: methods and results, *in preparation*

[11] K. Maples, Singularity of random matrices over finite fields, *arXiv:1012.2372 [math]*, December 2010.

[12] K. Maples, Cokernels of random matrices satisfy the Cohen-Lenstra heuristics, *arXiv:1301.1239 [math]*, January 2013.

[13] S. Meehan and H. Nguyen, Eigenvectors of Wigner matrices of symmetric entry distributions, *Proceedings of the American Mathematical Society*, **147** (2019), no. 2, 835-847.

[14] H. Nguyen and E. Paquette, Surjectivity of near square matrices, *arXiv:1802.0 0001 [math]*, January 2018.

[15] H. Nguyen, T. Tao, and V. Vu, Random matrices: tail bounds for gaps between eigenvalues, *Probability Theory and Related Fields*, **167** (2017), no. 3, 777-816.

[16] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, *Advances in Mathematics*, **226** (2011), no. 6, 5298-5319.

[17] H. Nguyen and V. Vu, Small ball probability, inverse theorems, and applications, *Erdős Centennial Proceeding*, Eds. L. Lovász et. al., Springer 2013.

[18] H. Nguyen and M. M. Wood, Random integral matrices: universality of surjectivity and the cokernel, *arXiv:1806.00596 [math]*, June 2018.

[19] S. O'Rourke and B. Touri, On a conjecture of Godsil concerning controllable random graphs, *SIAM Journal on Control and Optimization,* **54** (2016), no. 6, 3347-3378.

[20] S. O'Rourke and B. Touri, Controllability of random systems: universality and minimal controllability, *arXiv:1506.03125 [math]*, June 2015.

[21] S. O'Rourke, V. Vu and K. Wang, Eigenvectors of random matrices: a survey, *Journal of Combinatorial Theory Series A,* **144** (2016), issue C, 361-442.

[22] M. Rudelson and R. Vershynin, The Littlewood-Offord problem and invertibility of random matrices, *Advances in Mathematics*, **218** (2008), no. 2, 600-633.

[23] M. Rudelson and R. Vershynin, Smallest singular value of a random rectangular matrix, *Communications on Pure and Applied Mathematics*, **62** (2009), 1707-1739.

[24] A. Sárkőzy and E. Szemerédi, Uber ein Problem von Erdős und Moser, *Acta Arithmetica*, **11** (1965), 205-208.

[25] T. Tao, *Topics in Random Matrix Theory*, Amer. Math Soc., 2012.

[26] T. Tao and V. Vu, *Additative Combinatorics*, Cambridge Univ. Press, 2006.

[27] T. Tao and V. Vu, John-type theorems for generalized arithmetic progressions and iterated sumsets, *Adv. Math*, **219** (2008), no. 2, 428-449.

[28] T. Tao and V. Vu, Random matrices have simple spectrum, *Combinatorica*, **37** (2017), no. 3, 539-553.

[29] T. Tao and V. Vu, Random matrices: universal properties of eigenvectors, *Random Matrices: Theory and Applications*, **1** (2012), no. 1.

[30] T. Tao and V. Vu, A sharp inverse Littlewood-Offord theorem, *Random Structures and Algorithms*, **37** (2010), no. 4, 525-539.

[31] R. Vershynin, Invertibility of symmetric random matrices, *Random Structures and Algorithms*, **44** (2014), no. 2, 135-182.

[32] V. Vu and K. Wang, Random weighted projections, random quadratic forms and random eigenvectors, *Random Structures and Algorithms*, **47** (2015), no. 4, 792-821.