Centro de Investigación en Matemáticas, A.C.

Contributions on Non–Asymptotic Singularity of Random Matrices and on Backbend Percolation

TESIS
que para obtener el grado académico de

Doctor en Ciencias
con Orientación en Probabilidad y Estadística

Presenta:

Paulo César Manrique Mirón

Directores de tesis:
Víctor Pérez Abreu
Rahul Roy

Guanajuato, Gto. México

Enero de 2017

Contents

1 Introduction **7**

2 On the Singularity of Random Matrices **13**
- 2.1 Introduction 13
- 2.2 Random matrices over finite fields 13
- 2.3 Random matrices with continuous random entries 17
- 2.4 Minimum singular value of square matrices 19

3 Concentration Inequalities **27**
- 3.1 Introduction 27
- 3.2 Concentration functions 28
- 3.3 Concentration inequalities for sum of random variables 30

4 Ginibre and Wigner Matrices **41**
- 4.1 Introduction 41
- 4.2 Main results and applications 41
- 4.3 Proof in the Ginibre case 43
- 4.4 Proofs in the Wigner case 46

5 Circulant Random Matrices **53**
- 5.1 Introduction 53
- 5.2 Singularity of a circulant random matrix 54
- 5.3 Minimum singular value of a circulant random matrix 56
- 5.4 Roots of a random polynomial 56
- 5.5 Singularity of a circulant random matrix with prime dimension 62
- 5.6 Some additional results on circulant random matrices 63

6 Oriented Percolation with Backbend **69**
- 6.1 Introduction 69
- 6.2 The model 69
- 6.3 A characterization of the critical probability p_c^b 71
- 6.4 Stationary distribution for the edge process 76
- 6.5 Strict monotonicity of the limiting direction α^b over b 78
- 6.6 Exponential estimates for $p < p_c^b$ 79
- 6.7 The nature of the limiting direction α^b as the number of backbends b goes to infinity 83
- 6.8 A construction for studying $p > p_c^b$ 86

A Sub-Gaussian Random Variables **89**

Bibliography **90**

Acknowledgment

I would like to express my gratitude to CONACYT for gave me the opportunity to study a PhD under grant 224994. I would like to thank CIMAT for supported me in the last months of my PhD. I would like to say to thank Victor, Sergio, Gerardo, Rahul, Marceline, Sophie, Michelle, HocicoNegro, Marcos, María, Nicolás, Humbertina, Rafael, Stephanie, Lorena and all fictional creatures that I know, specially to Tenar, Philip K. Dick and H. P. Lovecraft. 

Chapter 1

Introduction

The main purpose of this thesis is the study of invertibility of unstructured and structured random matrices, which have been intensively investigated for at least five decades. One of the oldest references where the problem is mentioned goes back to 1964 in the work of Erdös and Rényi [23], and its pioneering study is found in the work by Komlós [39] in 1967. The problem of the singularity of random matrices arises in several areas of mathematics and its applications, such as the circular law [9], [29], compressed sensing [60], geometric functional analysis [59], [60], [75], smoothed analysis of algorithms [66], [67], and statistics [74],[60], among others.

Additionally, we include in this thesis a chapter with contributions on a different problem in backbend percolation, which was also worked as part of the PhD studies.

With regard to the singularity of random matrices, we consider the following random matrix models:

- **Ginibre matrix:** An $n \times n$ matrix $(\xi_{i,j})_{1 \le i,j \le n}$ is called a Ginibre matrix if $\xi_{i,j}, i, j = 1, \dots, n$ are independent random variables. It is an unstructured random matrix with n^2 independent entries.

- **Wigner matrix:** An $n \times n$ symmetric matrix $(\xi_{i,j})_{1 \le i,j \le n}$ is called a Wigner matrix if $\xi_{i,j}, 1 \le i \le j \le n$ are independent random variables. It is a structured random matrix with $n(n+1)/2$ independent entries.

- **Circulant random matrix:** An $n \times n$ matrix $(\xi_{i,j})_{1 \le i,j \le n}$ is called a circulant random matrix if $\xi_{i,j} = \xi_{1,j-i+1}$, where the subscripts are reduced modulo n and lie in the set $\{1, 2, \dots, n\}$, and the entries in the first row are independent random variables. It is a structured random matrix with n independent entries.

When these matrices have entries with continuous distributions, we have that they are invertible with probability one. But if the entries have discrete distributions, it is not immediate that they are invertible. This poses the question of what are the features of the discrete distributions that determine the invertibility of these models with high probability. Another question is how the singularity of random matrices depends on the number of independent random entries used in the construction of these matrices.

In the case of Ginibre matrices, Komlós [39] first considered a Ginibre matrix $GB(n, 1/2)$ whose entries are Bernoulli random variables, taking the values 0 or 1 with probability 1/2. Using a very clever "growing rank analysis" together with the Littlewood–Offord inequality (which is a concentration inequality[1]), Komlós proved that $\mathbb{P}\{\text{rank}(GB(n,1/2)) < n\} = \text{o}(1)$ as $n \to \infty$. Bollobás [8] presented an unpublished result due to Komlós about the rate of the probability of the singularity of $GB(n,1/2)$. Employing the concept of "strong rank" and the Littlewood–Offord inequality *viz*, Bollobás mentioned that $\mathbb{P}\{\text{rank}(GB(n,1/2)) < n\} = \text{O}(n^{-1/2})$ as $n \to \infty$. Komlós [40] was also the first to consider the singularity of a Ginibre matrix whose entries have a *common* arbitrary non-degenerate distribution, proving that the probability that such an $n \times n$ matrix is singular has order $\text{o}(1)$ as $n \to \infty$. This result was improved by Kahn, Komlós and Szemerédi [31] in the case of Ginibre matrices whose entries are i.i.d. taking the values -1 or 1 with probability $1/2$, showing that the probability of singularity is bounded above by θ^n for $\theta = .999$. The value of θ has been improved by Tao and Vu [69], [70] to $\theta = 3/4 + o(1)$ and by Bourgain, Vu and Wood [5] to $\theta = 1/\sqrt{2} + o(1)$. Slinko [65] considered Ginibre random matrices whose entries have the same uniform distribution taking values in a finite set, proving also that the probability of singularity is $\text{O}(n^{-1/2})$ as $n \to \infty$.

In the Wigner matrix case, the study of singularity was initiated by Costello, Tao and Vu [11], inspired by the work of Komlós [39]. They considered a Wigner matrix whose upper diagonal entries $\xi_{i,j}$ have Bernoulli distribution on $\{0,1\}$ with parameter $1/2$, and showed that $\mathbb{P}\{\text{rank}(W_n)) < n\} = \text{O}(n^{-1/8+\alpha})$, for any positive constant α, where the implicit constant in $\text{O}(\cdot)$ depends on α. They needed first to develop a quadratic Littlewood–Offord inequality, which is a concentration inequality for random quadratic forms. Nguyen [50] considered a Wigner matrix W_n with entries taking the values -1 or 1 with probability $1/2$, subject to the condition that each row has exactly $\lfloor n/2 \rfloor$ entries which are zero. He showed that the probability of W_n being singular is $\text{O}(n^{-C})$, for any positive constant C, and the implicit constant in $\text{O}(\cdot)$ depends on C. Vershynin [75] has considered the case of a Wigner matrix W_n whose entries satisfy the following property: the above-diagonal entries are independent and identically distributed with zero mean, unit variance, and are sub-Gaussian, while the diagonal entries satisfy $\xi_{ii} \leq K\sqrt{n}$ for some constant K. He showed that the probability of W_n being singular is bounded above by $2\exp(-n^c)$, where c depends only on the sub-Gaussian distribution and on K.

The previous results assume some restrictions on the distribution of the entries of the Ginibre or Wigner matrices. One of the first contributions of this thesis is to show that under the weaker conditions that the entries are non-degenerate independent random variables, not necessarily with equal distribution or moments, and such that the maximum jumps of their distributions are bounded by a number less than one (see Theorem 13 in Chapter 4), then the probability that the Ginibre or Wigner matrices are singular decreases to zero at least polynomially. Actually, in our investigation we establish universal rates of convergence and precise estimates for the probability of singularity of Ginibre and Wigner matrices, which depend only on the size of the maximum jumps of the distributions of the entries.

One of the main probability tools to prove that Ginibre and Wigner random matrices are invertible with high probability has been Levy's concentration function, which measures the maximum probability that a random variable lies in an interval. The problem of estimating the maximum probability that a linear combination of independent random variables belongs to a ball with given small radious is generally known as *Small Ball Probabilities* [51]. In 1943, Littlewood and Offord, in conection with their studies of random polynomials, estimated the small ball probability for a linear combination of Rademacher independent random variables [44]. Erdös studied the same case as Littlewood and Offord, but he analyzed the problem from a combinatorial point of view [22]. The small ball probabilty phenomenon was also studied in probability by Kolomogorov [37], [38], Rogozin [57], and others, and recently by Tao and Vu [72]. For the general assumptions we

[1]Some authors use the term "anti-concentration" instead "concentration."

consider in this work, we need first to establish an appropriate concentration inequality for a linear combination of independent random variables (linear Littlewood–Offord inequality) which is used in the Ginibre case. A suitable concentration inequality for a random quadratic form (quadratric Littlewood–Offord inequality) is also proved, which is used in the Wigner case. In both cases, we clearly exhibit the role of the maximum jumps of the distributions in these concentration inequalities.

Our Theorem 13 also handles the case when the entries of the Ginibre or Wigner matrices depend on the dimension of the matrix. This kind of random matrices appear in the study of random graphs [12], sparse matrices [13], [20], and some other models that have recently been extensively considered, like the so-called generalized, universal and banded Wigner ensembles [21], [64] among other works. See also the non i.i.d. Wigner case in, for example, [2, pp 26].

As another example where random matrices possess identifiable properties without moment assumptions, we have that a Ginibre matrix has large rank with exponentially small probability. Namely, let A be an $n \times n$ Ginibre matrix. Suppose that all the entries $\xi_{i,j}$ of A are random variables with different distributions, satisfying the following condition: for some $\varsigma \in (0,1)$

$$\sup_{x\in\mathbb{R}} \mathbb{P}\{\xi_{i,j} = x\} \leq \varsigma.$$

Then for all $\delta \in (0,1)$, we have

$$\mathbb{P}\left(\operatorname{rank}(A) < \delta n\right) < Cn\varsigma^{(1-\delta)^2 n^2},$$

for some suitable constant $C > 0$ which depends possibly on ς. This simple statement[2] shows that under weaker conditions, random matrices have "good qualities."

It is notable that if one considers the minimum singular value of a rectangular random matrix, we can see a similar phenomenon for the probability of the singularity of Ginibre and Wigner random matrices without moment assumptions. For a strictly rectangular matrix with i.i.d. random entries without moment assumptions, it was recently found by Tikhomirov [73] that the probability that its minimum singular value is large goes to one exponentially. For square matrices, estimating the minimum singular value (known also as the *hard edge* of spectrum) has been considerably more difficult [74].

Recall that if A is an $n \times n$ matrix with real or complex entries, the singular values[3] $s_k(A)$, $k = 1, \ldots, n$ of A are the eigenvalues of $|A| = \sqrt{A^*A}$ arranged in non-increasing order.

Note $s_n(A) > 0$ if and only if A is not singular, moreover $s_n(A)$ measures the distance of A to the set of singular matrices. The study of extreme singular values are interesting since, for example, they control the distortion of a Euclidean geometry under the action of the linear transformation A: the distance between any two points can increase by at most the factor $s_1(A)$ and decrease by at least the factor $s_n(A)$. The extreme singular values are clearly related to the operator norm of the linear operators A and A^{-1} acting between Euclidean spaces: $s_1(A) = \|A\|$ and if A is invertible, then $s_n(A) = 1/\|A^{-1}\|$. In numerical linear algebra, the condition number $\kappa_n(A) := s_1(A)/s_n(A)$ frequently serves as a measure of the stability of a matrix algorithm [60].

When A is a Ginibre matrix, one has, e.g., if its entries have sub-Gaussian moments with some additional weak assumptions, that for every $\varepsilon \geq 0$

$$\mathbb{P}(s_n(A)) \leq \varepsilon n^{-1/2}) \leq C\varepsilon + c^n, \tag{1.1}$$

where $C > 0$ and $c \in (0,1)$ depend on the sub-Gaussian moments [61]. Note that if $\varepsilon \neq 0$, the

[2] The idea of how to prove this statement is found in Lemma 13 in Chapter 4.

[3] This notation can be extended to rectangular matrices.

previous probabilities are not decreasing exponentially fast. If A is a Wigner random matrix, one has the same result as that in the expression (1.1) [75].

A second set of main contributions of this thesis is about the minimum singular value of a circulant random matrix. This $n \times n$ random matrix has a strong structure (great dependencies among the entries) since it can use at most n independent random variables. First, a result similar to (1.1) is obtained. Our approch to prove this new result for a circulant random matrix is different of the one used for Ginibre and Wigner matrices in [10], [61], [75], where they use the advantage that all or more than half of the entries are independent random variables.

Circulant matrices play a crucial role in the study of large-dimensional Toeplitz matrices. The study of random Toeplitz matrices is a relatively new field of research. The question of establishing the limiting spectral distribution of random Toeplitz matrices with independent entries was first posed in the review paper by Bai [1]. Bose, Subhra and Saha [6], [7], studied the probabilistic properties of the spectral norm (maximum singular value) of the scaled eigenvalues of circulant matrices. Sen and Virág [62] used circulant matrices to study the maximum singular value of a random symmetric Toeplitz matrix. Meckes [48], [49], who also studied the maximum singular value of symmetric Toeplitz matrices, was the first to give an estimate of the probability that a circulant random matrix is singular when its entries are Rademacher independent random variables, and he posed the problem of estimating the minimum singular value of a circulant random matrix.

For the study of the singularity of a circulant random matrix in this thesis, we use a remarkable relation among circulant random matrices and random polynomials. More specifically, the eigenvalues of a circulant matrix are the values that a certain polynomial takes on the roots of unity. A classic result in the theory of random polynomials says that the roots of a random polynomial are concentrated in the unit circle when the degree of the polynomial goes to infinity with probability one [4]. We show in Theorem 15 in Chapter 5 how the roots of a random polynomial move towards the unit circle at a certain speed, which implies that the minimum singular value is different from zero with high probability. Many results on the roots of random polynomials are about the behavior of the empirical distribution of the roots [4], [27], [63] without treating the speed with which the roots move towards the unit circle. Actually, the only references that we could find about the minimum value of a random polynomial on the unit circle were [34] and [42]. But in [34] there was no proof. Our Theorem 15 follows closely the ideas in [42]. In the proof of Theorem 15, we can extend the classic Salem–Zygmund's inequality for trigonometric random polynomial with i.i.d. coefficients, such that they have moment generating function. Actually, we show that a random variable ξ with $\mathbb{E}\left(\xi\right) = 0$, $\mathbb{E}\left(\xi^2\right) = \sigma^2 > 0$, and moment generating function $M_\xi(t)$ for $|t| < \Delta$ is locally sub-Gaussian random variable, i.e., there is $\Delta \geq \delta > 0$ such that $M_\xi(t) \leq e^{\gamma t^2/2}$ for $|t| < \delta$ and $\gamma > \sigma^2$.

Theorem 14 in Chapter 5 establishes that the mininum singular value $s_n(C_n)$ of an $n \times n$ circulant random matrix $\mathcal{C}_n$ whose entries have moment generating functions has the property that for all $\varepsilon > 0$ and for all large n, $\mathbb{P}(s_n(\mathcal{C}_n) \geq \varepsilon n^{-1/2}) \geq 1 - C\varepsilon$. This expression is similar to (1.1). This similitude was unexpected to the author since despite the large amount of dependency that exists among the entries, the minimum singular value of circulant matrices is roughly speaking similar as in the Ginibre and Wigner cases considered in [29], [75], also under some moment assumptions.

When the dimension of a circulant random matrix is a prime number, we can estimate the probability of its singularity when its entries have general distribution, without moment assumptions. Theorem 16 in Chapter 5 shows that a circulant random matrix with prime dimension is invertible with hign probability.

As a final contribution on circulant random matrices, we investigate the singularity phenomenon in g-circulant random matrices, which is a generalization of circulant matrices, with the same strong dependence among its entries. The g-circulant matrices have been an active research field of applied mathematics and computational mathematics [6], [7], [76]. There are many examples from statistics and information theory that illustrate applications of g-circulant matrices [76].

As a third and final set of contributions of this thesis, we include some results on backbend percolation that were also part of the PhD work of the author, which is different from the subject of random matrices. Backbend percolation is a generalization of oriented percolation by Durrett in [18], and was introduced by Roy, Sarkar and White [58]. Backbend model considers a path as defined in [18], with the diference that the path is allowed to go down until some depth b. We show that there exists a critical probability of this model as in oriented percolation, and we also study properties of the backbend model and similarities and diferences with unoriented percolation in two dimensions. Specifically, we establish the critical probability in terms of the "slope" of the right edge process. Our approach follows the ideas in [18], however the dependency inherent in the backbend model unlike in oriented percolation requieres a different analysis. Indeed, in oriented percolation the right edge process is built from independent random variables, while in our backbend model the right edge process is made from dependent random variables.

The remainder of this thesis is structured as follows.

- Chapter 2 contains a brief introduction to the problem of the singularity of random matrices. Section 2.2 mentions the singularity problem over finite fields and its differences from matrices over $\mathbb{R}$ or $\mathbb{C}$. Section 2.3 shows that some kind of models of random matrices are invertible with probability one when some of their entries have continuous distributions. Section 2.4 presents a result about the minimum singular value of a Ginibre matrix with i.i.d. entries, which was used in the proof of the circular law by Götze and Tikhomirov [29], with the goal of exemplifying the techniques used to analyze it.

- Chapter 3 presents the main probabilistic tools to prove that Ginibre and Wigner matrices are invertible with high probability. Section 3.2 introduces the notion of Levy's concentration function of a random variable. Section 3.3 presents our concentration inequalities for sums of random variables. This section includes: Theorem 11, which establishes a general concetration inequality for a linear combination of independent random variables (linear Littlewood–Offord inequality), and Theorem 12, which gives a general concentration inequality for random quadratic forms (quadratic Littlewood–Offord inequality).

- Chapter 4 presents some of the main results in this thesis. Sections 4.2 contains Theorem 13 where is established the universality rate of the probability of non-singularity of the Ginibre and Wigner matrices, for n large. Also, it is given an application of Theorem 13.b to Erdö-Rényi. Section 4.3 gives the proof of Theorem 13 for the case of the Ginibre matrices. The principal tools used for the Ginibre case are a suitable linear Littlewood–Offord inequality and the concept of "strong rank." Section 4.4 gives the proof of Theorem 13 for the case of Wigner matrices. The main tools used for the Wigner case are an appropiate quadratic Littlewood–Offord inequality and a slightly different concept of strong rank.

- Chapter 5 contains another set of the main contributions in this thesis. In Section 5.2 is presented the main theorems in this chapter, Theorem 14, 15 and 16. Theorem 14 determines the behavior of the minimum singular value of a circulant random matrix whose entries have moment generating functions. For our proof of Theorem 14, it is used a remarkable result about the roots of a random polynomial. This is found in Theorem 15. Theorem 16 shows that if the dimension of a circulant matrix is prime, it is possible to obtain an estimate for the probability of the singularity of a circulant matrix when its entries have general distributions. Sections 5.3, 5.4 and 5.5 are given the proofs of Theorem 14, 15 and 16, respectively. Section

5.6 includes some additional contributions about the extreme singular values of circulant random matrices and g-circulant matrices.

- Chapter 6 is about our contributions on backbend percolation. Section 6.2 describes the backbend percolation model. In Section 6.3 we give a characterization of the critical probability of backbend percolation. In Section 6.4 we show that there is an initial distribution on nonpositive integers such that the right edge process has stationary increments. Section 6.5 is about the strict monotocity of the "slope" of banckbend percolation. Section 6.6 contains our analysis of the sub–critical probability of backbend percolation model. Section 6.7 presents our first approach about the behavior of the "slope" of banckbend percolation when the models are near to unoriented percolation in two dimension. Section 6.8 considers the super–critical probability of the backbend percolation model.

- Appendix A provides some properties of sub-Gaussian random variables, material that is used in Chapter 5.

Every new result that the author provides is marked with a $\star$.

Chapter 2

On the Singularity of Random Matrices

2.1 Introduction

This chapter contains a brief introduction to the problem of the singularity of random matrices with the goal of exhibiting some different contexts where the problem arises and the distinct behavior where the probability of the singularity changes drastically. First, Section 2.2 refers to the case of random matrices over a finite field, since the probability of the singularity of some models is asymptotically non-zero. This is in contrast with the case where the random matrix is over $\mathbb{R}$ or $\mathbb{C}$—which is the one considered in this thesis—where this probability is asymptotically zero. For these situations, the study of non-singularity is commonly done using an analysis of the rank of the matrix ever since the pioneering work of Komlós [39]. Second, Section 2.3 presents first a proof of a result that is part of the folklore of the literature: unstructured random matrices whose entries have continuous distributions are invertible with probability one. Moreover, for some structured random matrices, we also prove that if the principal diagonal has entries which are independent continuous random variables and independent of the entries in the non-diagonal part (the distributions of the entries in the this part are arbitrary) are invertible with probability one. These proofs are carried out by an analysis of the determinant of the matrix. Hence, the non-trivial problem is when the entries have discrete distributions. Finally, in Section 2.4 we recall another feature of the singularity of matrices: the extreme singular values. We also include a result about the minimum singular value of a Ginibre matrix with i.i.d. entries, which is one of the keys behind the circular law as shown in Götze and Tikhomirov [29]. This is an example where the problem of singularity via minimum singular values comes out in very relevant situations in the theory of random matrices. It is not surprising that in this proof there appears naturally the use of concentration inequalities, which is the subject of Chapter 3.

2.2 Random matrices over finite fields

We would like to start by introducing the following "good" problem[1].

1. There are 16 2 by 2 matrices whose entries are 1's and 0's. How many are invertible?

[1] There is an affable anecdote about this problem [14], where we can note the interest in it. This problem appears as a "regular" exercise in [68].

2. (Much harder!) If you put 1's and 0's at random into the entries of a 10 by 10 matrix, is it more likely to be invertible or singular?

The first question is very easy. To check how many matrices are invertible we only need to list them. However, the second one is difficult, as we need to consider 2^{100} possibilities.

Binary matrices are studied in combinatorics, information theory, cryptology, and graph theory. In 1964, Erdös and Rényi [23] stated, at the end of their paper on the permanent of binary random matrices, the question of how many binary random matrices are non-singular (in $\mathbb{R}$). The invertibility of binary matrices is especially important in encoding, since it helps to encrypt messages and compress communication signals in an effective manner.

Note that the number of $n \times n$ binary matrices are 2^{n^2}. Let $F(n,k)$ be the number of $n \times n$ binary matrices of rank k, and $P(n) = F(n,n)/2^{n^2}$ the proportion of non-singular binary matrices. If we have a random matrix whose entries are independent Bernoulli random variables taking the value 1 with probability of 1/2 and the value 0 with probability 1/2, then $P(n)$ is the probability of non-singularity of this random matrix.

In 1967, Kómlos [39] showed that $P(n) \to 1$, but he studied the invertibility of this random matrix in $\mathbb{R}$. But in the case that the random matrix is over $\mathbb{F}_2$, finite field with two elements, we have that $P(n) \leq 1/2$ for all $n > 0$. This will be discussed in detail later. In Table 2.1 we can see the number of binary matrices which are invertible in $\mathbb{F}_2$ and $\mathbb{R}$.

n	Total	$F(n,n)$ in $\mathbb{F}_2$	$F(n,n)$ in $\mathbb{R}$	$P(n)$ in $\mathbb{F}_2$	$P(n)$ in $\mathbb{R}$
1	2^1	1	1	0.5	0.5
2	2^4	6	10	0.375	0.375
3	2^9	168	338	0.328125	0.33984...
4	2^{16}	20160	42976	0.307617...	0.34424...
5	2^{25}	9999360	21040112	0.298004...	0.37296...
6	2^{36}	2015870960	39882864736	0.293347...	0.41963...
7	2^{49}	163849992929280	292604283435872	0.291056...	0.48024...

Table 2.1: Number of non-singular matrices in $\mathbb{F}_2$ and $\mathbb{R}$.

Given a finite field $\mathbb{F}_q$ with q elements, the cardinality of the set $GL(n,q)$ of invertible matrices over $\mathbb{F}_q$ can be explicitly calculated and then we have the exact probability of non-singularity over $\mathbb{F}_q$ when the entries are independent discrete uniform random variables on all $\mathbb{F}_q$.

If $A \in GL(n,q)$, we can see A as a set of n linearly independent vectors in $\mathbb{F}_q$. We can construct A in the following way. The first vector in A should be different from zero, there are $q^n - 1$ choices. This vector spans a one-dimensional subspace, which contains q^1 elements. The second vector should not be in this subspace, so we have $q^n - q^1$ possibilities. In fact, if we have a set of $k-1$ independent vectors, there are $q^n - q^k$ possible ways to create an independent set with k vectors. Hence, the number of ways to choose vectors that will form an $n \times n$ invertible matrix is

$$\left(q^n - q^0\right)\left(q^n - q^1\right)\cdots\left(q^n - q^{n-1}\right) = \prod_{k=1}^{n}\left(q^n - q^{k-1}\right).$$

In other words, the probability of invertibility of an $n \times n$ matrix over the field $\mathbb{F}_q$ whose entries are independent random variables with discrete uniform distribution on all $\mathbb{F}_q$ is exactly

$$\frac{1}{q^{n^2}} \prod_{k=1}^{n} \left(q^n - q^{k-1}\right) = \prod_{k=1}^{n} \left(1 - q^{-k}\right) < 1 - \frac{1}{q}.$$

For $q = 2$, we have $\prod_{k=1}^{n} \left(1 - 2^{-k}\right) < 1/2$, i.e., less than half of the matrices over $\mathbb{F}_2$ are invertible. Moreover the number of invertible matrices decreases as n increases. Also, we can compute the number $\mathcal{G}(n, r, q)$ of $n \times n$ matrices over $\mathbb{F}_q$ with rank r. Note that if $r = 0$, then $\mathcal{G}(n, 0, q) = 1$.

Theorem 1.

$$\mathcal{G}(n, r, q) = (q^r)^{n-r} \prod_{k=1}^{r} \left(q^n - q^{k-1}\right).$$

Proof. The number of ways we can choose r linearly independent random vectors in $\mathbb{F}_q$ is $\prod_{k=1}^{r} \left(q^n - q^{k-1}\right)$. Now, observe that the other $n - r$ vectors should be a linear combination of these r first vectors. This gives us the result. □

From Theorem 1 we obtain the exact probability that a random matrix over $\mathbb{F}_q$ whose entries are independent random variables with discrete uniform distribution has rank r.

When $q \to \infty$, $\prod_{k=1}^{n} \left(1 - q^{-k}\right) \to 1$, i.e., if the number of elements of the finite field increases, the number of matrices which are invertible increases too. From Table 2.1, it can be observed that the probability of singularity decreases when we consider random matrices with independent Bernoulli random entries over $\mathbb{F}_2$ or $\mathbb{R}$. But, we will see that random matrices in $\mathbb{R}$ with independent continuous or discrete random entries are invertible with high probability. First, we want to show what happens with the probability of the singularity in the case of a symmetric random matrix over $\mathbb{F}_q$ with independent uniform random entries.

The next result[2] shows how to count the number of symmetric matrices over $\mathbb{F}_q$ with rank r. For this, we define $\mathcal{W}(n, r, q)$ as the number of $n \times n$ symmetric matrices over $\mathbb{F}_q$ with rank r. Write $d(n, j, q)$ for the number of j dimensional subspaces of $\mathbb{F}_q^n$. Define $\prod_n(x) = (1-x)(1-x^2)\cdots(1-x^n)$. It is well known [17] that

$$d(n, j, q) = \frac{\prod_n(q)}{\prod_{n-j}(q) \prod_j(q)}.$$

Theorem 2.

$$\mathcal{W}(n, n-j, q) = d(n, j, q)\mathcal{W}(n-j, n-j, q).$$

Proof. Let e_j be the n dimensional vector over $\mathbb{F}_q^n$ that has a 0 in each entry except for the jth entry. Let $E = \operatorname{span}\{e_1, \ldots, e_j\}$. Then there are $\mathcal{W}(n-j, n-j, q)$ $n \times n$ matrices with rank $n - j$ that have kernel E. To see this, we note that if M is an $n \times n$ matrix, then Me_j is the jth row of M. Hence, if A is a symmetric matrix with $Ae_j = 0$, then the jth row and the jth column of M is the zero vector.

Let A be a symmetric $n \times n$ matrix with rank $n - j$ that has kernel E. Since $Av = 0$ for all $v \in E$, then A has j rows and j columns equal to the zero vector. If we look at A without these j columns and the corresponding j rows, we have a symmetric $(n - j) \times (n - j)$ matrix that should have rank $n - j$. So, there are $\mathcal{W}(n - j, n - j, q)$ symmetric $n \times n$ matrices with rank $n - j$ that have kernel E.

[2]The proof of this theorem is the same as given by Prof. Robert C. Rhoades. The proof has been circulated on the Internet, `http://math.stanford.edu/~rhoades/FILES/rank_symmetric_matrices.pdf`

Let S be any j dimensional subspace of $\mathbb{F}_q^n$ with basis $\{v_1, \ldots, v_j\}$. We define $\mathcal{S}$ as the set of all $n \times n$ matrices of rank $n-j$ with kernel S and $\mathcal{E}$ is the set of all $n \times n$ matrices of rank $n-j$ with kernel E. Our goal is to show that $|\mathcal{S}| = |\mathcal{E}|$.

There are $k_1, \ldots, k_{n-j}$ such that $\{v_1, \ldots, v_j, e_{k_1}, \ldots, e_{k_{n-j}}\}$ is a basis for $\mathbb{F}_q^n$. Let B be the change of basis matrix defined by $e_s \mapsto v_s$ for $1 \le s \le j$ and $e_{j+t} \mapsto e_{k_t}$ for $1 \le t \le (n-j)$.

Define the map $\phi : \mathcal{S} \to \mathcal{E}$ by $\phi(A) = B^t AB$. Since B is an invertible matrix, B^t is too. Note, $B^t ABv = 0$ if and only if $ABv = 0$, but this implies $Bv \in S$. Since B is the change of basis matrix from $\{e_1, \ldots, e_j\}$ to $\{v_1, \ldots, v_j\}$, we have $v \in E$. Therefore, $B^t ABv = 0$ if and only if $v \in \mathcal{E}$, i.e., the map ϕ is well defined.

Since B is invertible, ϕ is a 1–1 map. Now, we consider $X \in \mathcal{E}$ and $Y = (B^t)^{-1}AB^{-1}$. We have $\phi(Y) = B^t YB = B^t((B^t)^{-1}XB^{-1})B = X$. Then it is enough to show that $Y \in \mathcal{S}$. But, $Yv = 0$ if and only if $XB^{-1}v = 0$, and since B^{-1} is also a change of the basis matrix from $\{v_1, \ldots, v_j\}$ to $\{e_1, \ldots, e_j\}$, we have $v \in S$. So $Y \in \mathcal{S}$, i.e., ϕ is surjective.

This completes the proof. Then $\mathcal{W}(n, n-j, q) = d(n, j, q)\mathcal{W}(n-j, n-j, q)$. $\square$

Theorem 2 gives the exact probability that a symmetric $n \times n$ matrix has rank r when its entries are independent uniform random entries over $\mathbb{F}_q$.

Random matrices over a finite field have been studied for many years [26]. Even though this thesis is about random matrices over $\mathbb{R}$ or $\mathbb{C}$, we would like to present some additional results about the singularity of random matrices over finite fields, since there are significance differences from the cases over $\mathbb{R}$ or $\mathbb{C}$.

Let $q = p^f$ be a prime power and let $\mathbb{F}_q$ be a finite field with q elements. Suppose ξ is a random variable that takes values in $\mathbb{F}_q$ with probability distribution μ. We say that μ is α-dense for $0 < \alpha < 1$ if for every additive subgroup $T \le \mathbb{F}_q$ and $s \in \mathbb{F}_q$,

$$\mathbb{P}(\xi \in s + T) \le 1 - \alpha.$$

Theorem 3. *[46] Let $\mathbb{F}_q$ with $q = p^j$ and suppose A is a $n \times n$ random matrix with i.i.d. entries which take values from an α-dense probability distribution. Then we have the estimate*

$$\mathbb{P}\left(A \textit{ is non-singular}\right) = \prod_{k=1}^{\infty}(1 - q^{-k}) + O\left(e^{-c\alpha n}\right),$$

where the implied constant and $c > 0$ are absolute.

$\square$

Theorem 4. *[47] Let Q_n be a $n \times n$ symmetric random matrix where the entries above and on the diagonal are independent copies of ξ, where ξ is a random variable with $\mathbb{P}(\xi \equiv t \mod q) < 1 - c$ for all $t \in \mathbb{F}_q$ with $c > 0$. Then we have the estimate for the total variation*

$$d_{TV}(n - \mathrm{rank}, \nu) = O\left(n^{-1/8}\right),$$

where the implied constant depends on c and q, and ν is the probability distribution on $\mathbb{Z}^+ := \{0, 1, 2, \ldots\}$ given by

$$\nu(k) := q^{-\frac{k^2+k}{2}} \prod_{l=k+1}^{\infty} (1 - q^{-l}).$$

$\square$

2.3 Random matrices with continuous random entries

If a random matrix A has entries [3] which are independent random variables with continuous distribution on $\mathbb{R}$ or $\mathbb{C}$, it is easy to check that $\mathbb{P}(\det(A) = 0) = 0$. A reason for this is that probability that the sum of independent random variables with continuous distribution takes a particular value is zero. The following results show that some structured random matrices with few continuous random entries are also invertible with probability one.

We first show that under the condition that the principal diagonal has entries which are independent with continuous distributions and independent of the upper and lower triangular parts, the random matrix is invertible with probability one, whatever the distributions of the non-diagonal entries. This theorem provides us a first guide when we want to find "good matrices" to test some algorithms, because it gaves an indication of what are the "bad matrices."

Theorem 5 ($\star$ Singularity: Continuous case)**.** *Let $G_n = (\xi_{i,j})_{1\leq i,j\leq n}$ be an $n \times n$ random matrix such that $\{\xi_{i,i} : i = 1, \dots, n\}$ is a set of independent random variables and independent of $\{\xi_{i,j} : 1 \leq i, j \leq n$ with $i \neq j\}$. If $\xi_{i,i}$ has continuous distribution for all i, then*

$$\mathbb{P}\left(G_n \textit{ is non-singular}\right) = 1.$$

Proof. The proof is by induction on n. For $n = 1$ the statement is trivial. We can suppose the statement is true for all $n \leq k$ for some k. Now, we consider $n = k + 1$. Since

$$\det(G_{k+1}) = \xi_{1,1}\det(G_k) + \sum_{j=2}^{k+1} \xi_{1,j} c_{1,j},$$

where $c_{1,j}$ is the $(1, j)$-cofactor of G_{k+1}, we have by independence and the inductive hypothesis

$$\begin{aligned}
\mathbb{P}\left(G_{k+1} \text{ is singular}\right) &= \mathbb{P}\left(\xi_{1,1}\det(G_k) + \sum_{j=2}^{k+1} \xi_{1,j} c_{1,j} = 0\right) \\
&= \mathbb{P}\left(\xi_{1,1} = -\frac{\sum_{j=2}^{k+1} \xi_{1,j} c_{1,j}}{\det(G_k)}\right) \\
&= \mathbb{E}\left(\mathbb{P}\left(\xi_{1,1} = -\frac{\sum_{j=2}^{k+1} \xi_{1,j} c_{1,j}}{\det(G_k)} \,\middle|\, \xi_{i,j} \text{ with } i,j = 1,\dots,n \text{ and } (i,j) \neq (1,1)\right)\right) \\
&= \mathbb{E}\left(0\right) \\
&= 0.
\end{aligned}$$

Therefore $\mathbb{P}\left(G_{k+1} \text{ is non-singular}\right) = 1$. $\square$

We note that if G is a Ginibre or Wigner matrix, then by Theorem 5 we have $\mathbb{P}(\det(G) = 0) = 0$.

Moreover, it is possible to consider a strong dependency among the entries of a random matrix. For example, if T is a random Toeplitz matrix, we can prove that the probability that $\det(T) = 0$ is zero only under the hypothesis that the entries of T are independent continuous random variables.

[3] In fact, this statement is also verified when the rows or columns are independent continuous random vectors.

We recall that a Toeplitz matrix is defined as $T_n = (\xi_{i-j})_{i,j=1}^n$, i.e., T_n looks like

$$T_n = \begin{bmatrix} \xi_0 & \xi_{-1} & \xi_{-2} & \cdots & \cdots & \xi_{-(n-1)} \\ \xi_1 & \xi_0 & \xi_{-1} & \ddots & & \vdots \\ \xi_2 & \xi_1 & \xi_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \xi_{-1} & \xi_{-2} \\ \vdots & & \ddots & \xi_1 & \xi_0 & \xi_{-1} \\ \xi_{n-1} & \cdots & \cdots & \xi_2 & \xi_1 & \xi_0 \end{bmatrix}.$$

Theorem 6 ($\star$ Singularity: Toeplitz case)**.** *Let $T_n = (\xi_{i,j})_{1\le i,j\le n}$ be an $n \times n$ random matrix such that $\{\xi_i : i = -(n-1), \ldots, n-1\}$ is an independent set of random variables. If ξ_i has continuous distribution for all i, then*

$$\mathbb{P}\left(T_n \text{ is non-singular}\right) = 1.$$

Proof. The proof is by induction on n. For $n = 1$ the statement is trivial. We can suppose the statement is true for all $n \le k$ for some k. Now, we consider $n = k + 1$. Since

$$\det(T_{k+1}) = \xi_{-k}\det(T_k) + \sum_{j=-(k-1)}^{0} \xi_j c_{1,|j|+1},$$

where $c_{1,|j|+1}$ is the $(1, |j| + 1)$-cofactor of T_{k+1}. We have by independence and the inductive hypothesis

$$\begin{aligned} \mathbb{P}\left(\det(T_{k+1}) = 0\right) &= \mathbb{P}\left(\xi_{-k}\det(T_k) + \sum_{j=-(k-1)}^{0} \xi_j c_{1,|j|+1} = 0\right) \\ &= \mathbb{P}\left(\xi_{-k} = -\frac{\sum_{j=-(k-1)}^{0} \xi_j c_{1,|j|+1}}{\det(T_k)}\right) \\ &= \mathbb{E}\left(\mathbb{P}\left(\xi_{-k} = -\frac{\sum_{j=-(k-1)}^{0} \xi_j c_{1,|j|+1}}{\det(T_k)} \,\middle|\, \xi_i \text{ with } i = k, \ldots, -(k-1)\right)\right) \\ &= \mathbb{E}\left(0\right) \\ &= 0. \end{aligned}$$

Therefore $\mathbb{P}\left(T_{k+1} \text{ is non-singular}\right) = 1$. □

Another example of a matrix with entries having a strong dependence is a circulant random matrix[4]. When the entries are independent continuous random variables, then a circulant random matrix is non-singular with probability one. However, the structure of a random matrix can drastically change the probability of singularity. This is shown in the following simple example.

Example 1. *Let $\xi_0, \xi_1, \ldots$ be independent random variables such that there is a $\varsigma \in (0, 1)$ with*

[4]This kind of matrix will be studied in more detail in Chapter 5.

$\sup_{x\in\mathbb{R}} \mathbb{P}\{\xi_k = x\} \leq \varsigma$ *for all* k. *Let* A *be the* $n \times n$ *random matrix defined as*

$$A_n = \begin{pmatrix} \xi_0 & \xi_1 & \xi_2 & \cdots & \xi_n \\ \xi_1 & \xi_0 & \xi_2 & \cdots & \xi_n \\ \xi_1 & \xi_2 & \xi_0 & \cdots & \xi_n \\ & & \cdots & & \\ \xi_1 & \xi_2 & \xi_3 & \cdots & \xi_0 \end{pmatrix}.$$

It is not difficult to show that $\det(A_n) = X_n Y_n$, *where*

$$X_n = \sum_{i=0}^{n} \xi_i \quad \text{and} \quad \det(Y_n) = \prod\nolimits_{i=1}^{n} (\xi_0 - \xi_i).$$

Suppose $\mathbb{P}(\xi_0 = \xi_1) = p \in [0,1]$ *and* $\mathbb{P}(\xi_0 = \xi_i) = 0$ *for all* $i \neq 0, 1$. *By the Kolmogorov–Rogozin inequality (Theorem 8 in Chapter 3),*

$$\begin{aligned} p = \mathbb{P}(Y_n = 0) &\leq \mathbb{P}(X_n = 0 \text{ or } Y_n = 0) \\ &\leq \mathbb{P}(X_n = 0) + \mathbb{P}(Y_n = 0) \\ &\leq Cn^{-1/2} + p. \end{aligned}$$

Then

$$\mathbb{P}(\det(A_n) = 0) \to p \quad \text{as} \quad n \to \infty.$$

□

In the proofs of the previous results we used the determinant. Then a natural question is: what is the behavior of the determinant of a random matrix? The study of the distribution of the determinant of a Ginibre matrix was considered in [52]. We do not pursue this problem in the present thesis.

The previous results show that under hypothesis that the random entries have continuous distribution, the random matrix is invertible with probability one. So, the complicated scenario appears when one considers discrete random entries. This situation will be analyzed in Chapters 4 and 5.

2.4 Minimum singular value of square matrices

We would like to finish this introductory chapter by mentioning some results on the extreme singular values of random matrices with independent entries. To understand the behavior of extreme singular values is actually one of the keys to the circular law theorem [9]. This is an example where the problem of singularity comes out in very relevant situations in the theory of random matrices [29]. We know that the minimum singular value of a matrix is positive if and only if the matrix is non-singular. The next result states that the matrix should be invertible if the maximum singular value is not large. This lemma and its proof can be found in the review of the circular law in [9]. We include it here to highlight the role of the minimum singular value in the study of the spectral asymptotic distribution of Ginibre matrices.

Recall that the extreme singular values are defined for a matrix A by the variational formulas

$$s_1(A) = \max_{x:\|x\|_2=1} \|Ax\|_2, \qquad s_n(A) = \min_{x:\|x\|_2=1} \|Ax\|_2.$$

Lemma 1 (Small singular values of random matrices with independent entries). *If $(X_{ij})_{1\le i,j\le n}$ is a random matrix with independent and non-constant entries in $\mathbb{C}$ and if $a > 0$ is a positive real number such that*

$$b := \min_{1\le i,j\le n} \mathbb{P}(|X_{ij}| \le a) \quad and \quad \sigma^2 := \mathrm{Var}(X_{ij}\mathbb{1}_{\{|X_{ij}|\le a\}}) > 0,$$

then there exists $c = c(a, b, \sigma) > 1$ such that for any $n \times n$ matrix M in $\mathbb{C}$, $n \le c$, $s \le 1$, $0 < t \le 1$,

$$\mathbb{P}\left(s_n(X + M) \le \frac{t}{\sqrt{n}}; s_1(X + M) \le s\right) \le c\sqrt{\log(cs)}\left(ts^2 + \frac{1}{\sqrt{n}}\right).$$

The proof of Lemma 1 is divided into two parts, which correspond to a subdivision of the unit sphere $\mathbb{S}^{n-1}$ of $\mathbb{C}^n$. For two real positive parameters $\delta, \rho > 0$ that will be fixed later, we define the set of sparse vectors

$$\mathrm{Sparse} := \{x \in \mathbb{C}^n : \mathrm{card}(\mathrm{supp}(x)) \le \delta n\},$$

where $\mathrm{supp}(x) := \{i : x_i \ne 0\}$, and we split $\mathbb{S}^{n-1}$ into the set of compressible vectors and the set of incompressible vectors as follows:

$$\mathrm{Comp} := \{x \in \mathbb{S}^{n-1} : \mathrm{dist}(x, \mathrm{Sparse}) \le \rho\} \quad \text{and} \quad \mathrm{Incomp} := \mathbb{S}^{n-1} \setminus \mathrm{Comp}.$$

We note that for A an $n \times n$ matrix over $\mathbb{C}$,

$$s_n(A) = \min_{x\in\mathbb{S}^{n-1}} ||Ax||_2 = \min\left(\min_{x\in\mathrm{Comp}} ||Ax||_2, \min_{x\in\mathrm{Incomp}} ||Ax||_2\right). \tag{2.1}$$

Compressible vectors

Lemma 2 (Distance of a random vector from a small subspace). *There exist $\varepsilon, c, \delta_0 > 0$ such that for all n sufficently large, all $1 \le i \le n$, any deterministic vector $v \in \mathbb{C}^n$ and any subspace H of $\mathbb{C}^n$ with $1 \le \dim(H) \le \delta_0 n$, we have, denoting $C := (X_{1i}, \dots, X_{ni}) + v$,*

$$\mathbb{P}(\mathrm{dist}(C, H) \le \varepsilon\sigma\sqrt{n}) \le c\exp(-c\sigma^2 n).$$

Proof. Let $\eta_k = \mathbb{1}_{\{|X_{ki}|\le a\}}$. Then $\eta_1 \dots, \eta_n$ are independent, $\eta_k \in \{0, 1\}$ and $\mathbb{E}(\eta_k) \ge b$ for all k. Then, from Hoeffding's deviation inequality [9],

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^n \eta_k \le \frac{nb}{2}\right) &= \mathbb{P}\left(\sum_{k=1}^n \eta_k - bn \le -\frac{nb}{2}\right) \\
&\le \mathbb{P}\left(\sum_{k=1}^n \eta_k - \sum_{k=1}^n \mathbb{E}(\eta_k) \le -\frac{nb}{2}\right) \\
&\le \exp\left(-\frac{nb^2}{2}\right).
\end{aligned}$$

Then it is enough to prove

$$\mathbb{P}(\,\mathrm{dist}(C, H) \le \varepsilon\sigma\sqrt{n}\,\big|\, E_m) \le c\exp(-c\sigma^2 n),$$

where $E_m := \{|X_{1i}| \le a, \dots, |X_{mi}| \le a\}$ with $m := \lceil nb/2\rceil$.

Let $\mathbb{E}_m[\cdot] := \mathbb{E}[\cdot|E_m; \mathcal{F}_m]$ denote the conditional expectation given E_m and the filtration $\mathcal{F}_m$ generated by $X_{m+1,i}, \dots, X_{n,i}$. Let $W = \operatorname{span}\{H, v, u, w\}$ where

$$u := (0, \dots, 0, X_{m+1,i}, \dots, X_{n,i}) \quad \text{and}$$

$$w := (\mathbb{E}[X_{1i}||X_{1i}| \leq a], \dots, \mathbb{E}[X_{mi}||X_{mi}| \leq a], 0, \dots, 0).$$

So, $\dim(W) \leq \dim(H) + 3$ and W is $\mathcal{F}_m$-measurable. Also, we have that

$$\operatorname{dist}(C, H) \leq \operatorname{dist}(C, W) = \operatorname{dist}(Y, W),$$

where

$$\begin{aligned} Y &:= (X_{1i} - \mathbb{E}[X_{1i}||X_{1i}| \leq a], \dots, X_{mi} - \mathbb{E}[X_{mi}||X_{mi}| \leq a], 0, \dots, 0) \\ &= C - u - v - w. \end{aligned}$$

By assumption, for $1 \leq k \leq m$,

$$\mathbb{E}_m[Y_k] = 0 \quad \text{and} \quad \mathbb{E}_m[|Y_k|^2] \geq \sigma^2.$$

Let $D = \{z : |z| \leq a\}$. We define the function $f : D^m \to \mathbb{R}_+$ by

$$f(x) = \operatorname{dist}((x_1, \dots, x_m, 0, \dots, 0), W).$$

This function is convex and 1-Lipschitz, and by Talagrand's inequality, for all $t \geq 0$

$$\mathbb{P}_m(|\operatorname{dist}(Y, M) - M_m| \geq t) \leq 4 \exp\left(-\frac{t^2}{16a^2}\right),$$

where M_m is the median of f under $\mathbb{P}_m$. From this inequality, we obtain for all $t \geq 0$

$$\mathbb{P}_m(-\operatorname{dist}(Y, M) + M_m \leq -t) \leq 4 \exp\left(-\frac{t^2}{16a^2}\right). \tag{2.2}$$

We want to prove that

$$M_m \geq \sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca.$$

We consider the event $\operatorname{dist}(Y, M) \geq \sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca$. Then

$$\mathbb{P}\left(\operatorname{dist}(Y, M) \geq \sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca\right) = \mathbb{P}\left(\operatorname{dist}^2(Y, M) \geq \left(\sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca\right)^2\right).$$

If we take $t = \left|\sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca\right|$ from (2.2), we have that there is a $c = c(a, b, \sigma) > 1$ such that

$$\mathbb{P}\left(\operatorname{dist}^2(Y, M) \geq \left(\sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca\right)^2\right) \leq \frac{1}{2}.$$

So, $M_m \geq \sqrt{\mathbb{E}_m[\operatorname{dist}^2(Y, M)]} - ca$. On the other hand, if P denotes the orthogonal projection on the orthogonal of W, we find

$$\begin{aligned}
\mathbb{E}_m[\mathrm{dist}^2(Y,M)] &= \sum_{k=1}^{m} \mathbb{E}_m[Y_k^2]P_{kk} \\
&\geq \sigma^2\left(\sum_{k=1}^{n} P_{kk} - \sum_{k=m+1}^{n} P_{kk}\right) \\
&\geq \sigma^2(n - \dim(H) - 3 - (n-m)) \\
&\geq \sigma^2\left(\frac{nb}{2} - \dim(H) - 3\right).
\end{aligned}$$

For n large enough, the last expression is bounded below by $c\sigma^2 n$ if $\delta_0 = b/4$. From (2.2) there follows the result. □

Let $0 < \varepsilon < 1$ and $s \geq 1$ be as in Lemma 2. We set

$$\rho = \frac{1}{4}\min\left\{1, \frac{\varepsilon\sigma}{s\sqrt{\delta}}\right\},$$

in particular $\rho \leq 1/4$. The parameter $\delta \in (0,1)$ is still to be specified, we only assume that $\delta < \delta_0$. We note that if A is an $n \times n$ matrix over $\mathbb{C}$ and $y \in \mathbb{C}^n$ is such that $\mathrm{supp}(y) \subset \pi \subset \{1,\ldots,n\}$, then

$$||Ay||_2 \geq ||y||_2 s_n(A_{|\pi}),$$

where $A_{|\pi}$ is an $n \times |\pi|$ matrix formed by the columns of A selected by π. So

$$\min_{x\in\mathrm{Comp}} ||A||_2 \geq \frac{3}{4}\min_{\pi\subset\{1,\ldots,n\}:|\pi|=\lfloor\delta n\rfloor} s_n(A_{|\pi}) - \rho s_1(A). \tag{2.3}$$

Write C_i for the ith column of A and

$$H_i := \mathrm{span}\{C_i : j \in \pi, j \neq i\}.$$

Then for any $x \in \mathbb{C}^{|\pi|}$,

$$||A_{|\pi}x||_2^2 = \left|\left|\sum_{i\in\pi} x_i C_i\right|\right|_2^2 \geq \max_{i\in\pi}|x_i|^2\mathrm{dist}^2(C_i,H_i) \geq \frac{1}{|\pi|}\sum_{i\in\pi}|x_i|^2\min_{i\in\pi}\mathrm{dist}^2(C_i,H_i).$$

In particular

$$s_n(A_{|\pi}) \geq \min_{i\in\pi}\mathrm{dist}(C_i,H_i)/\sqrt{|\pi|}.$$

Since H_i has dimension at most δn and is independent of C_i, by Lemma 2, we can see that the event $\min_\pi \mathrm{dist}(C_i,H_i) \geq \varepsilon\sigma\sqrt{n}$ has probability at least $1 - c\delta n\exp(-c\sigma^2 n)$ for n sufficiently large. Hence, for $A = X + M$,

$$\mathbb{P}\left(s_n((X+M)_{|\pi}) \leq \frac{\varepsilon\sigma}{\sqrt{\delta}}\right) \leq c\delta n\exp(-c\delta^2 n).$$

Therefore, using the union bound and our choice of ρ, we have from (2.3)

$$\begin{aligned}\mathbb{P}\left(\min_{x\in\text{Comp}}||(X+M)x||_2 \leq \frac{\varepsilon\sigma}{2\sqrt{\delta}}; s_1(X+M)\leq s\right) &\leq \binom{n}{\lfloor \delta n\rfloor} c\delta n e^{-c\sigma^2 n} \\ &= c\delta n \exp(H(\delta)(1+\mathrm{o}(1)) - c\sigma^2),\end{aligned}$$

where $H(\delta) := -\delta\log(\delta) - (1-\delta)\log(1-\delta)$. If we pick δ small enough so that $H(\delta) < c\sigma^2/2$, we have that there is a $c_1 := c_1(\sigma) > 0$ such that

$$\mathbb{P}\left(\min_{x\in\text{Comp}}||(X+M)x||_2 \leq \frac{\varepsilon\sigma}{2\sqrt{\delta}}; s_1(X+M)\leq s\right) \leq \exp(-c_1 n). \tag{2.4}$$

From now on, we fix

$$\delta = \frac{c_2\sigma^2}{|\log\sigma|},$$

small enough so $\delta < \delta_0$ and $H(\delta) > c\sigma^2/2$.

Incompressible vectors: Invertibility via distance

Lemma 3 (Incompressible vectors are spread)**.** *Let $x \in Incomp$. There exists a subset $\pi \subset \{1,\dots,n\}$ such that $|\pi| \geq \delta n/2$ and for all $i \in \pi$*

$$\frac{\rho}{\sqrt{n}} \leq |x_i| \leq \sqrt{\frac{2}{\delta n}}.$$

Proof. For $\pi \subset \{1,\dots,n\}$, we denote by P_π the orthogonal projection on span$\{e_i : i \in \pi\}$. Let $\pi_1 = \{k : |x_k| \leq \sqrt{2/(\delta n)}\}$ and $\pi_2 = \{k : |x_k| \geq \rho/\sqrt{n}\}$. Since $||x||_2^2 = 1$, we have

$$|\pi_1^c| \leq \frac{\delta n}{2}.$$

Also,

$$||x - P_{\pi_2}x||_2 = ||P_{\pi_2^c}x||_2 \leq \rho.$$

If $|\pi_2| < \delta n$, we would have $x \in$ Comb, and then $|\pi_2| \geq \delta n$. Write $\pi = \pi_1 \cap \pi_2$. From the previous,

$$|\pi| \geq n - |\pi_1^c| - |\pi_1^c| \geq n - \frac{\delta n}{2} - (n - \delta n) = \frac{\delta n}{2}.$$

□

Lemma 4 (Invertibility via mean distance)**.** *Let A be a random matrix over $\mathbb{C}$ with columns $C_1,\dots,C_n$ and for some arbitrary $1 \leq k \leq n$, let H_k be the span of all these columns except C_k. Then, for any $t \geq 0$*

$$\mathbb{P}\left(\min_{x\in\text{Incomp}}||Ax||_2 \leq \frac{t\rho}{\sqrt{n}}\right) \leq \frac{2}{\delta n}\sum_{k=1}^{n}\mathbb{P}(\text{dist}(C_k, H_k) \leq t).$$

Proof. Let $x \in \mathbb{S}^{n-1}$. From $Ax = \sum_k C_k x_k$, we get

$$||Ax||_2 \geq \max_{1\leq k\leq n} \operatorname{dist}(Ax, H_k) = \max_{1\leq k\leq n} |x_k| \operatorname{dist}(C_k, H_k).$$

Now, $x \in$ Incomp and π as in Lemma 3, we have

$$||Ax||_2 \geq \frac{\rho}{\sqrt{n}} \max_{k\in\pi} \operatorname{dist}(C_k, H, k).$$

Finally, note that for any real numbers $y_1, \ldots, y_n$ and $1 \leq m$,

$$\mathbb{1}_{\{(\max_{1\leq k\leq m} y_k)\leq t\}} \leq \frac{1}{m}\sum_{k=1}^{m} \mathbb{1}_{\{y_k\leq t\}} \leq \sum_{k=1}^{n} \mathbb{1}_{\{y_k\leq t\}}.$$

□

Let C be the kth column of $X + M$. We want to establish, for all $t \geq 0$, that

$$\mathbb{P}(\operatorname{dist}(C, H) \leq \rho t; s_1(X + M) \leq s) \leq \frac{c}{\sigma}\sqrt{\frac{|\log\rho|}{\delta}}\left(t + \frac{1}{\sqrt{t}}\right). \tag{2.5}$$

In order to obtain this, we consider a random vector η in $\mathbb{S}^{n-1} \cap H^{\perp}$ that is independent of C. Note that η is not unique, we just pick one and we call it the orthogonal vector to the subspace H. By the Cauchy–Schwarz inequality,

$$|\langle C, \eta\rangle| \leq \operatorname{dist}(C, H). \tag{2.6}$$

Lemma 5 (The random orthogonal vector is incompressible)**.** *For our choice of ρ and δ, and with c_1 as in (2.5), we have*

$$\mathbb{P}(\eta \in \text{Comp}; s_1(X + M) \leq s) \leq \exp(-c_1\sigma^2 n).$$

Proof. Let A be the $(n-1)\times n$ matrix obtained from $(X+M)^*$ by removing the kth row. Then by construction $A\eta = 0$, $||Ax||_2 \leq ||(X + M)^* x||$, and $||(X + M)^*|| = ||X + M||$. Hence if $\eta \in$ Comp, we have $\min_{x\in\text{Comp}} ||Ax||_2 = 0$. Note that (2.5) holds with $X + M$ replaced by A. □

Now, we will use the Berry–Essen theorem. In this step we assume that some coordinates are fixed, both the components of η and the random variables $X_{ik} + M_{ik}$ are well controlled. Namely, if $\eta \in$ Incomp, let $\pi \subset \{1, \ldots, n\}$ be as in Lemma 3 associated to vector η. Then conditioned on $\{\eta \in \text{Incomp}\}$, from Hoeffding's deviation inequality, the event that

$$\sum_{i\in\pi} \mathbb{1}_{\{|X_{ik}|\leq a\}} \geq \frac{|\pi|b}{2} \geq \frac{\delta bn}{4},$$

has conditional probability at least (since η and hence π are independent of C)

$$1 - \exp(-|\pi|b^2/2) \geq 1 - \exp(-c\delta n).$$

So, using our choice of δ and ρ, and using Lemma 5 and (2.6), it is sufficient, to prove (2.5), to show that for all $t \geq 0$

$$\mathbb{P}(|\langle \eta, C\rangle| \leq \rho t) \leq \frac{c}{\sigma}\sqrt{\frac{|\log \rho|}{\delta}}\left(t+\frac{1}{\sqrt{n}}\right),$$

where $\mathbb{P}_m(\cdot) = \mathbb{P}(\cdot|E_m, \mathcal{F}_m)$ is the conditional probability given $\mathcal{F}_m$ the σ-algebra generated by all variables except $(X_{1k}, \dots, X_{mk})$, $m := \lfloor \delta b n/4 \rfloor$, and

$$E_m := \left\{\frac{\rho}{\sqrt{n}} \leq |\eta_i| \leq \sqrt{\frac{2}{\delta n}} : 1 \leq i \leq m\right\} \bigcup \{|X_{ik}| \leq a : 1 \leq i \leq m\}.$$

Write

$$\langle \eta, C\rangle = \sum_{i=1}^{n} \bar{\eta}_i \langle C, e_i\rangle = \sum_{i=1}^{m} \bar{\eta}_i X_{ik} + u,$$

where $u \in \mathcal{F}_m$ is independent of $(X_{1k}, \dots, X_{mk})$. It follows that

$$\mathbb{P}_m(|\langle \eta, C\rangle| \leq \rho t) \leq \sup_{z\in\mathbb{C}} \mathbb{P}_m\left(\left|\sum_{i=1}^{m} \bar{\eta}_i (X_{ik} - \mathbb{E}_m X_{ik}) - z\right| \leq \rho t\right). \tag{2.7}$$

Now, we use the rate of convergence given by the Berry–Essen theorem to obtain an upper bound for this last expression.

Lemma 6 (The small ball probability via the Berry–Essen theorem)**.** *There exists a constant $c > 0$ such that if $Z_1, \dots, Z_n$ are independent centered complex random variables, then for all $t \geq 0$,*

$$\sup_{z\in\mathbb{C}} \mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i - z\right| \leq t\right) \leq \frac{ct}{\sqrt{\sum_{i=1}^{n} \mathbb{E}(|Z_i|^2)}} + \frac{c\sum_{i=1}^{n} \mathbb{E}(|Z_i|^3)}{(\sum_{i=1}^{n} \mathbb{E}(|Z_i|^2))^{3/2}}.$$

Proof. Let $\tau^2 = \sum_{i=1}^{n} \mathbb{E}|Z_i|^2$. Then either $\sum_{i=1}^{n} \mathbb{E}(\Re Z_i)^2$ or $\sum_{i=1}^{n} \mathbb{E}(\Im Z_i)^2$ is greater than or equal to $\tau^2/2$, where $\Re z$ and $\Im z$ are, respectively, the real and imaginary parts of z. Also

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i - z\right| \leq t\right) \leq \mathbb{P}\left(\left|\sum_{i=1}^{n} \Re(Z_i) - \Re(z)\right| \leq t\right)$$

and similarly with $\Im$. We can assume without loss of generality that the Z_i's are real random variables. Then, if G is a real centered Gaussian random variable with variance τ^2, the Berry–Essen theorem asserts that

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(\sum_{i=1}^{n} Z_i \leq t\right) - \mathbb{P}(G \leq t)\right| \leq c_0 \tau^{-3/2} \sum_{i=1}^{n} \mathbb{E}(|Z_i|^3).$$

For all $t \geq 0$ and $x \in \mathbb{R}$, we have

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i - x\right| \leq t\right) \leq \mathbb{P}(|G - x| \leq t) + 2c_0 \tau^{-3/2} \sum_{i=1}^{n} \mathbb{E}(|Z_i|^3).$$

To conclude, we note that $\mathbb{P}(|G - x| \leq t) \leq 2t/\sqrt{2\pi\tau^2}$. $\square$

Define $L = \frac{1}{2}\log_2 \frac{2}{\delta\rho^2}$. For our choice of ρ and δ, we can find a constant $c = c(a,b)$ such that

$$L \leq c|\log \rho|.$$

For $1 \leq j \leq L$, we define

$$\pi_j := \left\{1 \leq i \leq m : \frac{2^{j-1}}{\rho}\sqrt{n} \leq |\eta_i| \leq \frac{2^j \rho}{\sqrt{n}}\right\}.$$

From the pigeonhole principle, there exists j such that $|\pi_j| \geq m/L$. So, we have

$$\sigma_j := \sum_{i\in\pi_j} |\eta_i|^2 \mathbb{E}_m(|X_{ik} - \mathbb{E}_m(X_{ik})|^2) \geq \frac{2^{2j-2}\rho^2\sigma^2|\pi_j|}{n},$$

and

$$\sum_{i\in\pi_j} |\eta_i|^3 \mathbb{E}_m(|X_{ik} - \mathbb{E}_m(X_{ik})|^3) \leq \frac{2^{j+1}a\rho}{\sqrt{n}}\sigma_j.$$

Recall that X and Y are independent random variables, $\mathbb{P}(|X+Y-z| \leq r) \leq \mathbb{P}(|X-z| \leq r)$. Now, from (2.7) and Lemma 6 (by changing the value of c), we get, for all $t \geq 0$,

$$\begin{aligned}
\mathbb{P}_m(|\langle \eta, C\rangle| \leq \rho t) &\leq \frac{c\rho t}{\sigma_j} + \frac{c2^j a\rho}{\sigma_j\sqrt{n}} \\
&\leq \frac{ct\sqrt{n}}{\sigma\sqrt{|\pi_j|}} + \frac{c}{\sigma\sqrt{|\pi_j|}} \\
&\leq \frac{c}{\sigma}\sqrt{\frac{|\log\rho|}{\delta}}\left(t + \frac{1}{\sqrt{n}}\right).
\end{aligned}$$

The proof of (2.5) is complete.

Proof of Lemma 1. By Lemma 4 and (2.5), we find, for all $t \geq 0$,

$$\mathbb{P}\left(\min_{x\in\text{Incomp}} ||(X+M)x||_2 \leq \frac{\rho^2 t}{\sqrt{n}}; s_1(X+M) \leq s\right) \leq \frac{c}{\sigma}\sqrt{\frac{|\log\rho|}{\delta^3}}\left(t + \frac{1}{\sqrt{n}}\right).$$

Using our choice of ρ and δ, we obtain for some new constant $c = c(a,b,\sigma) > 0$,

$$\mathbb{P}\left(\min_{x\in\text{Incomp}} ||(X+M)x||_2 \leq \frac{t}{\sqrt{n}}; s_1(X+M) \leq s\right) \leq c\sqrt{\log cs}\left(ts^2 + \frac{1}{\sqrt{n}}\right).$$

The desired result follows from (2.1) and (2.5). □

Chapter 3

Concentration Inequalities

3.1 Introduction

In this chapter we present the main probability tools used to prove that Ginibre or Wigner matrices are invertible with high probability. Section 3.2 introduces the notation of Levy's concentration function, which is useful for understanding with what probability the sum of independent random variables takes a value in some interval. In Section 3.3, we mention some well known results about the concentration of sum of independent random variables and establish two proper concentration inequalities for the Ginibre and Wigner cases, respectively.

The main idea is to estimate the probability that a linear combination of independent random variables can take a particular value, which will be used in the proof of the Ginibre case. It is easy to see that in the continuous case, the probability that the sum takes a particular value is zero, hence the difficult stage is when we have discrete random variables. In Theorem 11 we establish, basically, that the probability that a linear combination of n independent random variables takes a particular value is at most $C_L n^{-1/2}$, where the constant C_L depends on the maximum jumps of the distribution of the random variables.

We also analyze the concentration of a random quadratic form, something which will be used in the proof for the Wigner case. Suppose that A is a Wigner matrix. Then we want to estimate the probability that the quadratic form $x^t Ax$ is zero. A random quadratic form is a sum of dependent random variables, but using a decoupling argument, it is possible to give a good estimate of the probability that $x^t Ax = c$ for any $c \in \mathbb{R}$.

In our Theorem 12, we obtain that the probability that a linear combination of n independent random variables takes a particular value is at most $C_Q n^{-1/4}$, where the constant C_Q also depends on the maximum jumps of the distribution of the random variables.

As mentioned in the Introduction of this thesis, the problem of estimating the maximum probability that a sum of random variables belongs to an interval is known as the topic of *small ball probability.* The study of the small ball probability goes back to the discovery made by Littlewood and Offord [44] and Erdös [22] almost 70 years ago. The set of these problems were studied by Doeblin, Lévy [15, 16], Erdös [22] (for the Bernoulli case, where it reduces to the Littlewood–Offord problem), Kolmogorov [38], Rogozin [57], Kesten [35] and Essen [24], and more recently by Tao and Vu [71], and Rudelson and Vershynin [59], [75].

Sections 3.2 and 3.3 are from the book [54]. The proof of Theorem 11 used the Kolmogorov–Rogozin inequality [24], [57], and the proof of Theorem 12 follows the ideas in [11], where only the Bernoulli case was considered.

The main results in this chapter were published by the author in joint work with Pérez–Abreu and Roy [45].

3.2 Concentration functions

The Levy's concentration function $Q(\xi;\lambda)$ of a random variable ξ is defined by

$$Q(\xi;\lambda) := \sup_{x\in\mathbb{R}} \mathbb{P}(\xi \in [x-\lambda/2, x+\lambda/2])$$

for every $\lambda \geq 0$. The function $Q(X;\lambda)$ is a non-decreasing function for λ. Also, it is clear $0 \leq Q(X;\lambda) \leq 1$ for every $\lambda \geq 0$.

We will show some properties about the concentration function.

Lemma 7. *If X and Y are independent random variables, then $Q(X+Y;\lambda) \leq \min\{Q(X;\lambda), Q(Y;\lambda)\}$ for every $\lambda \geq 0$.*

Proof. Writing $I_\lambda(x) = [x-\lambda/2, x+\lambda/2]$, we note for $y \in \mathbb{R}$

$$\mathbb{P}(X + y \in I_\lambda(x)) = \mathbb{P}(X \in I_\lambda(x+y)),$$

hence

$$\mathbb{P}(X+Y \in I_\lambda(x)) = \mathbb{E}\left(\mathbb{P}(X+Y \in I_\lambda(x)|Y)\right) \leq Q(X,\lambda).$$

Therefore $Q(X+Y;\lambda) \leq Q(X;\lambda)$. □

Lemma 8. *For every $\alpha \geq 0$ and $\lambda \geq 0$ we have $Q(\xi;\alpha\lambda) \leq (\lfloor\alpha\rfloor + 1)Q(\xi;\lambda)$*

Proof. Writing $I_\lambda(x) = [x-\lambda/2, x+\lambda/2]$. If $\alpha \in [0,1]$, then $\alpha\lambda \leq \lambda$ and

$$\mathbb{P}(\xi \in I_{\alpha\lambda}(x)) \leq \mathbb{P}(\xi \in I_\lambda(x)).$$

If $\alpha > 1$, we have

$$\mathbb{P}(\xi \in I_{\alpha\lambda}(x)) \leq \mathbb{P}(\xi \in I_\lambda(x)) + (\alpha - 1)Q(\xi;\lambda).$$

Thus, $Q(\xi;\alpha\lambda) \leq (\lfloor\alpha\rfloor + 1)Q(\xi;\lambda)$. □

Lemma 9. *Let ξ be a random variable with the characteristic function $f(t)$ and the concentration function $Q(\xi;\lambda)$. For $\delta \in (0,\pi)$*

$$Q(\xi;\lambda) \leq \left(\frac{a\sin(\delta/2)}{\delta/2}\right)^{-1} \int_{-a}^{a} |f(t)|dt \tag{3.1}$$

for every $\lambda \geq 0$ and $a > 0$ with $a\lambda \leq \delta$.

Proof. Let $h(t)$ be a function defined by

$$h(t) = \begin{cases} 1-|t| & \text{for } |t| \leq 1 \\ 0 & \text{for } |t| > 1 \end{cases}. \tag{3.2}$$

Note that $h(t)$ is a probability density function, which represent a symmetric random variable. Hence,

$$H(x) := \left(\frac{\sin(x/2)}{x/2}\right)^2 = \int_{-\infty}^{\infty} e^{itx}h(t)dt.$$

We denote by $F(x)$ the distribution function of the random variable ξ. For every real γ and $a > 0$, we have

$$\begin{aligned}\int_{-\infty}^{\infty} H(a(x-\gamma))dF(x) &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{ita(x-\gamma)}h(t)dt\right]dF(x)\\ &= \frac{1}{a}\int_{-a}^{a} e^{-i\gamma u}h\left(\frac{u}{a}\right)\left[\int_{-\infty}^{\infty} e^{iux}dF(x)\right]du\\ &= \frac{1}{a}\int_{-a}^{a} e^{-i\gamma u}h\left(\frac{u}{a}\right)f(u)du,\end{aligned}$$

whence

$$\int_{-\infty}^{\infty} H(a(x-\gamma))dF(x) \le \frac{1}{a}\int_{-a}^{a}|f(t)|dt.$$

Since $\lim_{x\to 0}(\sin(x)/x) = 1$, we have for $\delta \in (0,\pi)$ such that

$$\left(\frac{\sin(x/2)}{x/2}\right)^2 \ge \left(\frac{\sin(\delta/2)}{\delta/2}\right)^2 =: C_\delta \quad \text{for} \quad |x| \le \delta.$$

Let us denote

$$I_\lambda(\gamma) = [\gamma - \lambda/2, \gamma + \lambda/2].$$

If $a\lambda \le \delta$, then

$$\int_{-\infty}^{\infty} H(a(x-\gamma))dF(x) \ge \left(\frac{\sin(\delta/2)}{\delta/2}\right)^2 \mathbb{P}(\xi \in I_\lambda(\gamma))$$

and

$$\mathbb{P}(\xi \in I_\lambda(\gamma)) \le \frac{1}{aC_\delta}\int_{-a}^{a}|f(t)|dt.$$

In view that γ is arbitrary,

$$Q(\xi;\lambda) \le \frac{1}{aC_\delta}\int_{-a}^{a}|f(t)|dt.$$

□

We note one consequence of Lemma 9, corresponding to the value $\lambda = 0$. If ξ is an arbitrary random variable with the characteristic function $f(t)$ then for $a = \delta = \pi/2$

$$\sup_{x\in\mathbb{R}} \mathbb{P}(\xi = x) \le \sqrt{2}\int_{-\pi/2}^{\pi/2}|f(t)|dt. \tag{3.3}$$

Lemma 10. *Let ξ be a random variable with the characteristic function $f(t)$ and concentration function $Q(\xi;\lambda)$. Then*

$$Q(\xi;\lambda) \ge \frac{\lambda\sin(1)}{4\pi(1+2a\lambda)}\int_{-2a}^{2a}|f(t)|^2dt \tag{3.4}$$

for every non-negative λ and a.

Proof. Let $\xi^s = \xi - \eta$, where η is a random variable independent of ξ and having the same distribution as ξ, then $|f(t)|^2$ is the characteristic function of ξ^s. Note $h(\cdot)$, which was defined in (3.2), is the characteristic function of a distribution with the density $(1-\cos x)/\pi x^2$. Let U be a random variable independent of X and having the caracteristic function $h(t/4a)$, where $a > 0$. The random variable $V := \xi^s + U$ has a continuous distribution with chacracteristic function

$|f(t)|^2 h(t/4a)$. From lemma 7, we have $Q(V;\lambda) \leq Q(\xi,\lambda)$. By the Inversion Formula[1]

$$\mathbb{P}\left(|V| \leq \frac{1}{4a}\right) = \frac{1}{4a\pi}\int_{-4a}^{4a} |f(t)|^2 \left(1 - \frac{|t|}{4a}\right) \frac{\sin(t/4a)}{t/4a} dt.$$

Since $t \in [-4a, 4a]$, we have

$$Q\left(V; \frac{1}{2a}\right) \geq \mathbb{P}\left(|V| \leq \frac{1}{4a}\right) \geq \frac{\sin(1)}{4a\pi}\int_{-4a}^{4a} |f(t)|^2 \left(1 - \frac{|t|}{4a}\right).$$

By lemma 8, we have for $a, \lambda > 0$

$$\begin{aligned} Q(\xi;\lambda) &\geq Q(V;\lambda) \\ &\geq \left(\left\lfloor \frac{1}{2a\lambda} \right\rfloor + 1\right)^{-1} Q\left(V; \frac{1}{2a}\right) \\ &\geq \frac{2a\lambda\sin(1)}{4a\pi(1+2a\lambda)}\int_{-4a}^{4a} |f(t)|^2 \left(1 - \frac{|t|}{4a}\right) dt \\ &\geq \frac{\lambda\sin(1)}{4\pi(1+2a\lambda)}\int_{-2a}^{2a} |f(t)|^2 dt. \end{aligned}$$

In the case where $\lambda = a = 0$, the inequality (3.4) is satisfied. $\square$

3.3 Concentration inequalities for sum of random variables

Let ξ be a random variable with distribution function $F(x)$. For every $\lambda > 0$ write

$$D(\xi;\lambda) = \lambda^2 \int_{|x|<\lambda} x^2 dF(x) + \int_{|x|\geq\lambda} dF(x).$$

We define $D(\xi;0) := \mathbb{P}(\xi \neq 0)$. We have $D(\xi;\lambda) = 0$ for every $\lambda \geq 0$ if and only $\mathbb{P}(\xi = 0) = 1$.

If $0 < \lambda_1 < \lambda_2$, then

$$\begin{aligned} \lambda_2^{-1}\int_{|x|<\lambda_2} x^2 dF(x) &\leq \lambda_1^{-1}\int_{|x|<\lambda_1} x^2 dF(x) + \lambda_2^{-1}\int_{\lambda\leq|x|<\lambda_2} x^2 dF(x) \\ &\leq \lambda_1^{-1}\int_{|x|<\lambda_1} x^2 dF(x) + \int_{\lambda\leq|x|<\lambda_2} dF(x). \end{aligned}$$

Therefore $D(\xi;\lambda_2) \leq D(\xi;\lambda_1)$, i.e., $D(\xi;\lambda)$ is a non-increasing function. Additionally, if $u \geq \lambda$, then

$$\begin{aligned} D(\xi;\lambda) &\geq u^{-2}\int_{|x|<\lambda} x^2 dF(x) + \int_{\lambda\leq|x|<u} dF(x) \\ &\geq u^{-2}\int_{|x|\leq u} x^2 dF(x). \end{aligned}$$

[1] **Inversion Formula:** If f is the characteristic function of F. For $a < b$ points at which F is continuous, we have

$$F(b) - F(a) = \lim_{c\to\infty} \frac{1}{2\pi}\int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} f(t) dt.$$

If ξ is a random variable, we denote by ξ^s the corresponding symmetrized random variable, i.e., if ξ' is a random variable independent of ξ with the same distribution of ξ, $\xi^s = \xi - \xi'$.

Theorem 7. *Let $\xi_1, \ldots, \xi_n$ be independent random variables, $S_n = \sum_{k=1}^n \xi_k$. Let $\lambda_1, \ldots, \lambda_n$ be positive numbers, $\lambda_k \leq \lambda$, $k = 1, \ldots, n$. Then there exists an absolute positive constant A, such that*

$$Q(S_n; \lambda) \leq A\lambda \left(\sum_{k=1}^n \lambda_k^2 D(\xi_k^s; \lambda_k) \right)^{-1/2}. \tag{3.5}$$

Proof. Let $V_k(x)$ and $v_k(x)$ denote respectively the distribution function and the characteristic function of the random variable ξ_k, respectively. We apply lemma 9 to the sum S_n with $a = 1/\lambda$, we obtain

$$Q(\xi; \lambda) \leq A_1 \lambda \int_{|t| \leq 1/\lambda} \prod_{k=1}^n |v_k(t)| dt.$$

From the inequality $1 + x \leq e^x$ for every real x implies that

$$|v_k(t)|^2 \leq \exp(-(1 - |v_k(t)|^2)).$$

If $V_k^s(x)$ denote the distribution function of ξ_k^s, we have

$$1 - |v_k(t)|^2 = \int_{-\infty}^{\infty} (1 - \cos(tx)) dV^s(x).$$

Therefore,

$$Q(S_n; \lambda) \leq A_1 \lambda \int_{|t| \leq 1/\lambda} \exp\left(-\frac{1}{2} \sum_{k=1}^n \int_{-\infty}^{\infty} (1 - \cos(tx)) dV_k^s(x) \right) dt. \tag{3.6}$$

Let $L_k(x)$ be the function defined by $L_k(x) = V_k^s(x) - 1$ for $x > 0$ and $L_k(x) = V_k^s(x)$ for $x < 0$. $L_k(x)$ is a Lévy spectral function[2] (see [54] p. 35). In order to estimate the integral in (3.6), we will use the following lemma.

Lemma 11. *Let $L_k(x)$ be a Lévy spectral function for $k = 1, \ldots, n$. Let δ be a positive number, and let $0 < \lambda_k \leq \lambda_k$, $k = 1, \ldots, n$. Then*

$$\begin{aligned} &\int_{|t| \leq 1/\lambda} \exp\left\{ -\delta \sum_{k=1}^n \int_{|x|>0} (1 - \cos(tx)) dL_k(x) \right\} dt \\ &\qquad \leq A\delta^{-1/2} \left(\sum_{k=1}^n \left[\int_{0<|x|<\lambda_k} x^2 dL_k(x) + \lambda_k^2 \int_{|x| \geq \lambda_k} dL_k(x) \right] \right)^{-1/2}. \end{aligned} \tag{3.7}$$

[2]Let M be a function from $\mathbb{R}$ to $\mathbb{R}$. M is called Lévy spectral function if it has the following properties:

- M is defined on $\mathbb{R} \setminus \{0\}$
- M is nondecreasing on $(-\infty, 0)$ and on $(0, \infty)$ and is right continuous
- $M(-\infty) = 0 = M(\infty)$
- $\int_{0<|x|<\varepsilon} x^2 dM(x)$ is finite for all $\varepsilon > 0$.

Proof. If $|x| \leq 1$, then $1 - \cos x \geq \frac{11}{24}x^2$. For $|t| \leq 1/\lambda$, we have

$$\begin{aligned}\int_{|x|>0}(1-\cos(tx))dL_k(x) &= \int_{0<|x|<\lambda_k}(1-\cos(tx))dL_k(x) + \int_{|x|\geq\lambda_k}(1-\cos(tx))dL_k(x)\\ &\geq \frac{11}{24}t^2\int_{0<|x|<\lambda_k}x^2dL_k(x) + \int_{|x|\geq\lambda_k}(1-\cos(tx))dL_k(x).\end{aligned}$$

We denote the left-hand side of the inequality (3.7) by I and we write

$$\beta_0 = \delta\sum_{k=1}^{n}\int_{0<|x|<\lambda_k}x^2dL_k(x).$$

Then

$$I \leq \int_{|t|\leq 1/\lambda}\exp\left(-\frac{11}{24}\beta_0t^2\right)\prod_{k=1}^{n}\exp\left(-\delta\int_{|x|\geq\lambda_k}(1-\cos(tx))dL_k(x)\right)dt. \tag{3.8}$$

Let

$$\beta_k = \delta\lambda_k^2\int_{|x|\geq\lambda_k}dL_k(x) \quad 1 \leq k \leq n,$$

$$B = \sum_{k=0}^{n}\beta_k, \quad \alpha_k/B \quad 0 \leq k \leq n.$$

We have

$$B = \delta\sum_{k=1}^{n}\left(\int_{0<|x|<\lambda_k}x^2dL_k(x) + \lambda_k^2\int_{|x|\geq\lambda_k}dL_k(x)\right). \tag{3.9}$$

Without loss of generality we can assume that $\alpha_k > 0$ for all k. Appliying Hölder 's inequality to the right-hand side of (3.8), we have

$$I \leq \prod_{k=0}^{n}I_k^{\alpha_k}, \tag{3.10}$$

where

$$I_0 = \int_{|t|\leq 1/\lambda}\exp\left(-\frac{11}{24}Bt^2\right)dt,$$

$$I_k = \int_{|t|\leq 1/\lambda}\exp\left(-\frac{B}{\lambda_k^2}\int_{-\infty}^{\infty}(1-\cos(tx))dM_k(x)\right)dt, \qquad k = 1,\ldots,n,$$

where $M_k(x)$ is a distribution function such that

$$dM_k(x) = \begin{cases}\frac{1}{p_k}dL_k(x) & \text{if } |x| \geq \lambda_k\\ 0 & \text{if } |x| < \lambda_k\end{cases}$$

and

$$p_k = \int_{|x|\geq\lambda_k}dL_k(x).$$

We note

$$I_0 \leq \int_{-\infty}^{\infty}\exp\left(-\frac{11}{24}Bt^2\right)dt = A_2B^{-1/2}, \tag{3.11}$$

where $A_2 := \sqrt{24\pi/11}$. To estimate the integral I_k, $k = 1,\ldots,n$, we use the Jensens inequality with e^{-x}. Then

$$\exp\left(-\frac{B}{\lambda_k^2}\int_{-\infty}^{\infty}(1-\cos(tx))dM_k(x)\right) \le \int_{-\infty}^{\infty}\exp\left(-\frac{B}{\lambda_k^2}(1-\cos(tx))\right)dM_k(x)$$

and

$$I_k \le \int_{|x|\ge\lambda_k} J_k(t,x)dM_k(x),$$

where

$$J_k = J_k(t,x) := \int_{|t|\le 1/\lambda_k}\exp\left(-\frac{B}{\lambda_k^2}(1-\cos(tx))\right)dt.$$

To prove (3.7) it is sufficient to show that for some constant A_*,

$$J_k \le A_* B^{-1/2} \quad \text{if } |x| \ge \lambda_k \text{ and } 1 \le k \le n, \tag{3.12}$$

since that $I_k \le A_* B^{-1/2}$. In fact, from (3.10), (3.11), and $\sum_{k=0}^n \alpha_k = 1$ imply $I \le A_* B^{-1/2}$. Then (3.7) follows from the latter inequality and from (3.9).

If $\lambda_k \le |x| \le \pi\lambda$, then $|tx| \le \pi$ for $|t| \le 1/\lambda$. Using the inequality $\sin u/u \ge 2/\pi$ for $|u| \le \pi/2$, we obtain

$$1-\cos(tx) = 2\sin^2\frac{tx}{2} \ge \frac{2}{\pi^2}t^2x^2 \ge \frac{2}{\pi^2}t^2\lambda_k^2$$

and

$$J_k \le \int_{|t|\le 1/\lambda_k}\exp\left(-\frac{2}{\pi^2}Bt^2\right)dt \le A_3 B^{-1/2},$$

where $A_3 := \sqrt{\pi^3/2}$.

Now we consider $|x| > \pi\lambda$, but it is sufficient to suppose when $x > \pi\lambda$. Then

$$\begin{aligned}
\lambda J_k &= \frac{\lambda}{x}\int_{|u|\le x/\lambda}\exp\left(-\frac{B}{\lambda^2}(1-\cos u)\right)du \\
&\le \frac{2\lambda}{x}\left(\left\lfloor\frac{x}{2\pi\lambda}\right\rfloor+1\right)\int_{|u|\le\pi}\exp\left(-\frac{B}{\lambda^2}(1-\cos u)\right)du \\
&\le \frac{3}{\pi}\int_{|u|\le\pi}\exp\left(-\frac{B}{\lambda^2}(1-\cos u)\right)du,
\end{aligned}$$

because the function under the intengral sign has the period 2π. We have $1-\cos u \ge u^2/\pi^2$ for $|u| \le \pi$. Therefore,

$$\int_{|u|\le\pi}\exp\left(-\frac{B}{\lambda^2}(1-\cos u)\right)du \le \int_{|u|\le\pi}\exp\left(-\frac{B}{\pi^2\lambda^2}u^2\right)du \le A_4\lambda B^{-1/2},$$

where $A_4 := \sqrt{2\pi^3}$. These estimates imply (3.12). □

To complete the proof of theorem 7 we apply lemma 12 to the integral on the right-hans side of (3.6). So

$$Q(S_n;\lambda) \le A\lambda\left(\sum_{k=1}^n\left(\int_{|x|<\lambda_k}x^2dV_k^s(x)+\lambda_k^2\int_{|x|\ge\lambda_k}dV_k^s(x)\right)\right)^{-1/2}.$$

□

Note that for an arbitrary random variable ξ and every $\lambda > 0$, we have

$$\begin{aligned}
\lambda^2 D(\xi;\lambda) &= \int_{|x|<\lambda} x^2 dF(x) + \lambda^2 \int_{|x|\geq\lambda} dF(x) \\
&\geq \frac{\lambda^2}{4} \int_{\lambda/2\leq|x|<\lambda} + \lambda^2 \int_{|x|\geq\lambda} dF(x) \\
&\geq \frac{\lambda^2}{4} \mathbb{P}\left(|\xi| \geq \frac{\lambda}{2}\right). \qquad (3.13)
\end{aligned}$$

By lemma 7 we have

$$\mathbb{P}\left(|\xi| \geq \frac{\lambda}{2}\right) \geq 1 - Q(\xi^s;\lambda) \geq 1 - Q(\xi;\lambda), \qquad (3.14)$$

and hence

$$D(\xi^s;\lambda) \geq 1 - Q(\xi;\lambda). \qquad (3.15)$$

From inequalities (3.13) and (3.15), we have the next useful result.

Theorem 8 (Kolmogorov-Rogozin Inequality). *Let $\xi_1, \ldots, \xi_n$ be independent random variables, $S_n = \sum_{k=1}^n \xi_k$. Let $\lambda_1, \ldots, \lambda_n$ be positive numbers such that $\lambda_k \leq \lambda$, $k = 1, \ldots, n$. Then there exists an absolute positive constant A, such that*

$$Q(S_n;\lambda) \leq A\lambda \left(\sum_{k=1}^n \lambda_k^2 \mathbb{P}\left(|\xi_k^s| \geq \frac{\lambda_k}{2}\right)\right)^{-1/2} \qquad (3.16)$$

$$Q(S_n;\lambda) \leq A\lambda \left(\sum_{k=1}^n \lambda_k^2 \left(1 - Q(\xi_k;\lambda_k)\right)\right)^{-1/2} \qquad (3.17)$$

□

Kesten [36] obtained the following refinement of the above inequality.

Theorem 9. *For the constant A of the Kolmogorov–Rogozin inequality and any independent random variables $\xi_1, \ldots, \xi_n$, and real numbers $0 < \lambda_1, \ldots, \lambda_n \leq 2\lambda$, one has*

$$Q(S_n;\lambda) \leq 4 \cdot 2^{1/2}(1+A)L \frac{\sum_{i=1}^n \lambda_i^2 \left[1 - Q(\xi_i;\lambda_i)\right] Q(\xi_i,\lambda)}{\left\{\sum_{i=1}^n \lambda_i^2 \left[1 - Q(\xi_i;\lambda_i)\right]\right\}^{3/2}}.$$

□

Since $\xi \in [x - \lambda/2, x + \lambda/2]$ is equivalently to $|\xi - x| \leq \lambda/2$. If a is a real number with $|a| \geq 1$, we have

$$Q(a\xi;\lambda) = Q\left(\xi; \frac{\lambda}{|a|}\right) \leq Q(\xi;\lambda). \qquad (3.18)$$

Let $a_1, \ldots, a_n$ be real numbers such that $|a_k| \geq 1$, $k = 1, \ldots, n$ and let $\xi_1, \ldots, \xi_n$ be independent random variables. If there exists $\lambda > 0$ such that $Q(\xi_k;\lambda) \leq \rho \in (0,1)$ for every k, we obtain from

(3.21) and theorem 8

$$\begin{aligned}\mathbb{P}\left(\sum_{k=1}^{n} a_k\xi_k = x\right) &\leq \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k\xi_k - x\right| \leq \lambda\right)\\ &\leq Q\left(\sum_{k=1}^{n} a_k\xi_k; \lambda\right)\\ &\leq A(1-\rho)^{-1/2}n^{-1/2}.\end{aligned}$$

The above statement gives us the intuition that any linear combination of independent random variables takes one particular value has small probability to happend. The next result establish more formal this idea. This situation is called *Linear Littlewood-Offord problem.*

Theorem 10 (Linear Littlewood-Offord problem v.1)**.** *Let $a_1, \dots, a_n$ be real numbers diffenterent from zero. Let $\xi_1, \dots, \xi_n$ be independent non-degenerate*[3] *random variables. If for every $k = 1, 2, \dots, n$ there exists $\lambda_k > 0$ such that $Q(\xi_k; \lambda_k) \leq \rho \in [0, 1)$. Then there exists an absolute positive constant A such that*

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left(\sum_{k=1}^{n} a_k\xi_k = x\right) \leq A(1-\rho)^{-1/2}n^{-1/2}.$$

Proof. Let $m = \min_{\{1\leq k\leq n\}} |a_k|$ and $\lambda = \min_{\{1\leq k\leq n\}} \lambda_k$, by theorem 8 we have

$$\begin{aligned}\mathbb{P}\left(\sum_{k=1}^{n} a_k\xi_k = x\right) &\leq \mathbb{P}\left(\left|\sum_{k=1}^{n} \frac{a_k}{m}\xi_k - \frac{x}{m}\right| \leq \lambda\right)\\ &\leq Q\left(\sum_{k=1}^{n} \frac{a_k}{m}\xi_k; \lambda\right)\\ &\leq A(1-\rho)^{-1/2}n^{-1/2}.\end{aligned}$$

□

In the case that $\xi_1, \dots, \xi_n$ are i.i.d. random variables, we have in theorem 10 that $\lambda_k = \lambda$ for every k. Also, when r of the random variables $\xi_1, \dots, \xi_n$ are degenerate, we have

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left(\sum_{k=1}^{n} a_k\xi_k = x\right) \leq A(1-\rho)^{-1/2}(n-r)^{-1/2}.$$

If $\varsigma = \sup_{x\in\mathbb{R}} \mathbb{P}(\xi = x)$, then for every $\Delta > 0$, there exists $\delta = \delta(\Delta)$ such that

$$\varsigma \leq Q(\xi; \delta) \leq \varsigma + \delta.$$

We denote $\varsigma + \delta$ by ς_Δ. From this observation, we can obtain a little generalizarion of theorem 10 using theorem 9

Theorem 11 ($\star$ The Linear Concentration Inequality v.2)**.** *Let $\xi_1, \dots, \xi_n$ be independent random variables with non-degenerate distributions $F_1, \dots, F_n$, respectively, and let $\alpha_1, \dots, \alpha_n$ be real num-*

[3] A random variable ξ is degenerate if there is $x \in \mathbb{R}$ with $\mathbb{P}(\xi = x) = 1$.

bers with $\alpha_i \neq 0$, $i = 1, \ldots, n$. Then

$$\sup_{x \in \mathbb{R}} \mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i \xi_i = x\right\} = O\left(\frac{\sum_{i=1}^{n}(1-\varsigma(i))\varsigma_\Delta(i)}{\left\{\sum_{i=1}^{n}\left[1-\varsigma_\Delta(i)\right]\right\}^{3/2}}\right), \tag{3.19}$$

where the implicit constant in $O(\cdot)$ does not depend on F_i, $i = 1, \ldots, n$.

Proof. Let $a = \min_{1 \leq i \leq n}\{|\alpha_i|\}$ and $\delta = \min_{1 \leq i \leq n}\{\delta_i\}$, where $\delta_i > 0$ satisfies $\varsigma_\Delta(i) = Q(\xi_i, \delta_i)$, $i = 1, \ldots, n$. We have for $x \in \mathbb{R}$

$$\mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i \xi_i = x\right\} = \mathbb{P}\left\{\sum_{i=1}^{n} \frac{\alpha_i}{a} \xi_i = \frac{x}{a}\right\} = \mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i' \xi_i = x'\right\},$$

where $\alpha_i/a = \alpha_i'$ and $x/a = x'$. Now,

$$\begin{aligned}
\mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i' \xi_i = x'\right\} &\leq \sup_{y \in \mathbb{R}} \mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i' \xi_i \in [y, y+\delta]\right\} \\
&\leq 4 \cdot 2^{1/2}(1+9C)\frac{\sum_{i=1}^{n}(1-\varsigma(i))\varsigma_\Delta(i)}{\left\{\sum_{i=1}^{n}\left[1-\varsigma_\Delta(i)\right]\right\}^{3/2}},
\end{aligned}$$

the last expression following from theorem 9. □

If $\varsigma_\Delta(i) < \varsigma < 1$ for all i, from theorem 11

$$\sup_{x \in \mathbb{R}} \mathbb{P}\left\{\sum_{i=1}^{n} \alpha_i \xi_i = x\right\} = \mathrm{O}\left(\frac{\varsigma}{\sqrt{(1-\varsigma)^3 n}}\right). \tag{3.20}$$

When r of the random variables $\xi_1, \ldots, \xi_n$ are degenerate for some $1 \leq r < n$; n is replaced by $n - r$ in (3.19).

From theorem 11, it is natural to ask what happens if we consider a polynomial of degree k in $\xi_1, \ldots, \xi_n$. In the case $k = 2$, we have the *Quadratic Littlewood-Offord problem.*

In order to have a similar estimation as it was obtained in theorem 10, we will use a decouplig argument.

Lemma 12 (Decoupling)**.** *Let $X \in \mathbb{R}^{m_1}$ and $Y \in \mathbb{R}^{m_2}$ be independent random variables, with $m_1 + m_2 = n$, and let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a Borel function. Let X^* be a copy independent of X which is independent of Y. For any interval I of $\mathbb{R}$, we have*

$$\mathbb{P}^2\left(\varphi(X,Y) \in I\right) \leq \mathbb{P}\left(\varphi(X,Y) \in I, \varphi(X^*,Y) \in I\right).$$

Proof. We note

$$\begin{aligned}
\mathbb{P}\left(\varphi(X,Y) \in I, \varphi(X^*,Y) \in I \,|Y=y\right) &= \mathbb{P}\left(\varphi(X,y) \in I, \varphi(X^*,y) \in I\right) \\
&= \mathbb{P}^2\left(\varphi(X,y) \in I\right).
\end{aligned}$$

The above expression implies

$$\mathbb{P}\left(\varphi(X,Y) \in I, \varphi(X^*,Y) \in I \,|Y\right) = \mathbb{P}^2\left(\varphi(X,Y) \in I\right),$$

and finally

$$\mathbb{P}^2\left(\varphi(X,Y)\in I\right)\leq\mathbb{P}\left(\varphi(X,Y)\in I,\varphi(X^*,Y)\in I\right).$$

$\square$

Theorem 12 ($\star$ The Quadratic Littlewood-Offord Inequality)**.** *Let $\xi_1,\ldots,\xi_n$ be independent random variables with non-degenerate distributions $F_1,\ldots,F_n$, respectively, and let $(c_{ij})_{1\leq i,j\leq n}$ be a symmetric $n\times n$ array of constants. Suppose $S_1\sqcup S_2$ is a partition of $\{1,2,\ldots,n\}$ such that for each $j\in S_2$, the set $N_j:=\{i\in S_1:c_{ij}\neq 0\}$ is non-empty. Let*

$$\varphi=\varphi\left\{\xi_1,\ldots,\xi_n\right\}=\sum_{1\leq i,j\leq n}c_{ij}\xi_i\xi_j$$

be the quadratic form whose coefficients are c_{ij}. Then any $x\in\mathbb{R}$

$$\mathbb{P}\left\{\varphi=x\right\}=O\left(\left[\frac{1}{|S_2|}\sum_{j\in S_2}\left(\frac{\sum_{i\in N_j}(1-\overline{\varsigma}(i))\overline{\varsigma}_\Delta(i)}{\left\{\sum_{i\in N_j}\left[1-\overline{\varsigma}_\Delta(i)\right]\right\}^{3/2}}\right)+\sup_{D\subset S_2,|D|\geq|S_2|/2}\frac{\sum_{j\in D}(1-\varsigma(j))\varsigma_\Delta(j)}{\left\{\sum_{j\in D}\left[1-\varsigma_\Delta(j)\right]\right\}^{3/2}}\right]^{1/2}\right),$$

where for ξ_i' an independent copy of ξ_i, $\overline{\varsigma}(i)$ and $\overline{\varsigma}_\Delta(i)$ are the jumps associated with $\xi_i-\xi_i'$ and $\varsigma(j)$ and $\varsigma_\Delta(j)$ are the jumps associated with ξ_j. The implicit constant in $O(\cdot)$ does not depend on F_i, $i=1,\ldots,n$.

Proof. Let $\delta=\min_{1\leq i\leq n}\{\delta_i\}$ where $\delta_i>0$ satisfies $\varsigma_\Delta(i)=Q(\xi_i,\delta_i)$, $i=1,\ldots,n$. If $x\in\mathbb{R}$, we have

$$\mathbb{P}\left\{\varphi=x\right\}\leq\mathbb{P}\left\{\varphi\in[x,x+\delta/2]\right\}.$$

Write $I=[x,x+\delta/2]$, $X=(\xi_i:i\in S_1)$, $Y=(\xi_i:i\in S_2)$ and $X'=(\xi_i':i\in S_1)$, with X' independent of X and Y, but having the same distribution as X. By lemma 12,

$$\begin{aligned}\mathbb{P}^2\left\{\varphi(X,Y)\in I\right\}&\leq\mathbb{P}\left\{\varphi(X,Y)\in I,\varphi(X',Y)\in I\right\}\\&\leq\mathbb{P}\left\{\varphi(X,Y)-\varphi(X',Y)\in[-\delta/2,\delta/2]\right\}.\end{aligned}$$

We can rewrite $\varphi(X,Y)-\varphi(X',Y)$ as

$$\begin{aligned}\varphi(X,Y)-\varphi(X',Y)&=g(X,X')+2\sum_{j\in S_2}\xi_j\left(\sum_{i\in S_1}c_{ij}\left(\xi_i-\xi_i'\right)\right)\\&=g(X,X')+2\sum_{j\in S_2}\xi_j\eta_j,\end{aligned}$$

where $g(X,X')=\sum_{i,j\in S_1}c_{ij}(\xi_i\xi_j-\xi_i'\xi_j')$ and $\eta_j=\sum_{i\in S_1}c_{ij}\left(\xi_i-\xi_i'\right)$.

Let ζ be the number of η_j which are equal to zero. If $J=[-\delta/2,\delta/2]$, we have

$$\begin{aligned}\mathbb{P}\left\{\varphi(X,Y)-\varphi(X',Y)\in J\right\}\leq\mathbb{P}&\left\{\varphi(X,Y)-\varphi(X',Y)\in J,\zeta\leq\frac{|S_2|}{2}\right\}\\&+\mathbb{P}\left\{\zeta>\frac{|S_2|}{2}\right\}.\end{aligned}$$

Since $\zeta = \sum_{j \in S_2} \mathbf{1}_{\{\eta_j = 0\}}$, using theorem 11, we have

$$\begin{aligned}\mathbb{E}(\zeta) = \sum_{j \in S_2} \mathbb{P}\{\eta_j = 0\} &= \sum_{j \in S_2} \mathbb{P}\left\{\sum_{i \in N_j} c_{ij}\left(\xi_i - \xi_i'\right) = 0\right\} \\ &= \sum_{j \in S_2} \mathrm{O}\left(\frac{\sum_{i \in N_j}(1-\overline{\varsigma}(i))\overline{\varsigma}_\Delta(i)}{\left\{\sum_{i \in N_j}\left[1-\overline{\varsigma}_\Delta(i)\right]\right\}^{3/2}}\right),\end{aligned}$$

where $\overline{\varsigma}(i)$ and $\overline{\varsigma}_\Delta(i)$ are the jumps associated with $\xi_i - \xi_i'$. By Markov's inequality, we obtain

$$\mathbb{P}\left\{\zeta > \frac{|S_2|}{2}\right\} \leq \frac{2}{|S_2|}\mathbb{E}(\zeta) = \frac{1}{|S_2|}\sum_{j \in S_2} \mathrm{O}\left(\frac{\sum_{i \in N_j}(1-\overline{\varsigma}(i))\overline{\varsigma}_\Delta(i)}{\left\{\sum_{i \in N_j}\left[1-\overline{\varsigma}_\Delta(i)\right]\right\}^{3/2}}\right).$$

For $M := \{j \in S_2 : \eta_j \neq 0\}$, we note that (i) M is a random set which depends only on X, X' and (ii) $|M| \geq |S_2|/2$ whenever $\zeta \leq |S_2|/2$. Thus for a given realization x, x' of X, X' respectively, we have

$$\mathbb{P}\left\{\varphi(x,Y) - \varphi(x',Y) \in J \;\middle|\; \zeta \leq \frac{|S_2|}{2}\right\} = \mathbb{P}\left\{2\sum_{j \in S_2} \xi_j \eta_j \in J' \;\middle|\; \zeta \leq \frac{|S_2|}{2}\right\},$$

where $J' = [-g(x,x') - \delta/2, -g(x,x') + \delta/2]$. Then by theorem 11,

$$\mathbb{P}\left\{\varphi(x,Y) - \varphi(x',Y) \in J \;\middle|\; \zeta \leq \frac{|S_2|}{2}\right\} = \mathrm{O}\left(\frac{\sum_{j \in M(x,x')}(1-\varsigma(j))\varsigma_\Delta(j)}{\left\{\sum_{j \in M(x,x')}\left[1-\varsigma_\Delta(j)\right]\right\}^{3/2}}\right),$$

where $M(x,x')$ is the set M obtained for the realization x, x' of X, X'. So

$$\begin{aligned}&\mathbb{P}\left\{\varphi(X,Y) - \varphi(X',Y) \in J \;\middle|\; \zeta \leq \frac{|S_2|}{2}\right\} = \\ &= \mathbb{E}\left(\mathbb{P}\left\{\varphi(X,Y) - \varphi(X',Y) \in J \;\middle|\; \zeta \leq \frac{|S_2|}{2},\ X, X'\right\}\right) \\ &= \mathbb{E}\left(\mathrm{O}\left(\sup_{D \subset S_2, |D| \geq |S_2|/2} \frac{\sum_{j \in D}(1-\varsigma(j))\varsigma_\Delta(j)}{\left\{\sum_{j \in D}\left[1-\varsigma_\Delta(j)\right]\right\}^{3/2}}\right)\right) \\ &= \mathrm{O}\left(\sup_{D \subset S_2, |D| \geq |S_2|/2} \frac{\sum_{j \in D}(1-\varsigma(j))\varsigma_\Delta(j)}{\left\{\sum_{j \in D}\left[1-\varsigma_\Delta(j)\right]\right\}^{3/2}}\right).\end{aligned}$$

Hence

$$\mathbb{P}\{\varphi = x\} = \mathrm{O}\left(\left[\frac{1}{|S_2|}\sum_{j \in S_2}\left(\frac{\sum_{i \in N_j}(1-\overline{\varsigma}(i))\overline{\varsigma}_\Delta(i)}{\left\{\sum_{i \in N_j}\left[1-\overline{\varsigma}_\Delta(i)\right]\right\}^{3/2}}\right) + \sup_{D \subset S_2, |D| \geq |S_2|/2} \frac{\sum_{j \in D}(1-\varsigma(j))\varsigma_\Delta(j)}{\left\{\sum_{j \in D}\left[1-\varsigma_\Delta(j)\right]\right\}^{3/2}}\right]^{1/2}\right).$$

□

In theorem 12, if we suppose that $\varsigma_{\Delta}(i) < \varsigma < 1$ for every k, $|S_1| = |S_2| = n/2$, and $|N_j| \geq n^{1-\varepsilon}$ for all j and $\varepsilon > 0$, we have

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left\{\varphi = x\right\} = \mathrm{O}\left(\left[\frac{\varsigma}{\sqrt{(1-\varsigma)^3 n^{1-\varepsilon}}}\right]^{1/2}\right). \tag{3.21}$$

Even as theorem 11, we have if $1 \leq r \leq n$ of the random variables $\xi_1, \ldots, \xi_n$ are degenerate, in this situation, n is replaced by $n - r$ in (3.21).

Chapter 4

Ginibre and Wigner Matrices

4.1 Introduction

This chapter contains some of the main contributions of this thesis. We prove the universal asymptotically almost sure non-singularity of the Ginibre and Wigner matrices. The problem of estimating the probability that a Ginibre matrix is singular is a basic problem in the theory of random matrices and combinatorics [5]. Theorem 13.a establishes, under the assumptions that the entries of an $n \times n$ Ginibre matrix are independent random variables with possibly different distributions (possibly without moments and a few are allowed to be degenerate) which depend on the dimension of the matrix, that it is singular with probability at most $C_G n^{-1/2}$, where the constant C_G depends basically on the maximum of the jumps of the distributions of the entries. The proof of Theorem 13.a follows the ideas in [65], which considered the discrete uniform case and used a linear Littlewood–Offord inequality.

Theorem 13.b, under the same assumptions as for the Ginibre case, gives that an $n \times n$ Wigner matrix is singular with probabilty at most $C_W n^{-(1-\varepsilon)/4}$, for any $\varepsilon \in (0, 1)$, where the constant C_W depends basically on ε and the maximum of the jumps of the distributions of the entries. The proof of Theorem 13.b follows the ideas in [11], which considered the Bernoulli case and used a quadratic Littlewood–Offord inequality.

Our Wigner models include the adjacency matrices of Erdös–Rényi random graphs [8]. The distribution of the entries of the adjacency matrix of this random graphs depend on the dimension of the matrix. For this reason we present one example of the application of Theorem 13.b to the non-singularity of the adjacency matrix of a random graph.

Proposition 2 shows basically how to construct a sequence of Ginibre matrices from an arbitrary sequence of distribution functions, such that the probability of the singularity of the Ginibre matrices goes to one. The proof of this statement follows the pioneering ideas of Komlós in [40], which have been widely used in this area.

The main results in this chapter were published by the author in joint work with Pérez–Abreu and Roy [45].

4.2 Main results and applications

First, we introduce some definitions that it will be used in the rest of this chapter.

Definiton 1. *An $n \times n$ matrix $\mathcal{G}_n = (\xi_{i,j})_{1 \le i,j \le n}$ is called* **Ginibre matrix** *if $\xi_{i,j}, i, j = 1, \dots, n$ are independent random variables.*

Definiton 2. *An $n \times n$ symmetric matrix $\mathcal{W}_n = (\xi_{i,j})_{1\le i,j\le n}$ is called* **Wigner matrix** *if $\xi_{i,j}, 1 \le i \le j \le n$ are independent random variables.*

Given a collection of non-degenerate distribution functions $\{F_{ij}^{(n)} : \; i,j \ge 1, \; n \ge 1\}$ and a subsequence $\{m_n : n \ge 1\}$, we study the singularity of the $m_n \times m_n$ matrix with independent entries $\xi_{kl}^{(n)}$ governed by the distribution function $F_{kl}^{(n)}$ for every $1 \le k,l \le m_n$. Let us denote by ς_n the biggest jump of the distribution functions $F_{ij}^{(n)}$, $1 \le i,j \le m_n$, i.e., if $\varsigma_{i,j} = \sup_{x\in\mathbb{R}} \mathbb{P}\{\xi_{i,j}^n = x\}$, then

$$\varsigma_n = \max_{1\le i,j\le n} \{\varsigma_{i,j}\}. \tag{4.1}$$

We give a sufficient condition for $m_n = n$ in terms of the sequence of biggest jumps $(\varsigma_n)_{n\ge 1}$.

Theorem 13 ($\star$ Universality of the non-singularity of Ginibre and Wigner matrices)**.** *With the notation as above, let $G_r^{(n)}$ and $W_r^{(n)}$ be the $r \times r$ Ginibre and Wigner matrices respectively, each with entries $\xi_{i,j}^{(n)}$, $1 \le i \le j \le r$. Assume that $\varsigma_n < \varsigma \in [0,1)$ for all n*

a) As $n \to \infty$

$$\mathbb{P}\left\{ rank(G_n^{(n)}) < n \right\} = O\left(n^{-1/2}\right), \tag{4.2}$$

where the implicit constant in $O(\cdot)$ depends on ς.

b) For any $\varepsilon \in (0,1)$,

$$\mathbb{P}\left\{ rank(W_n^{(n)}) < n \right\} = O\left(n^{-(1-\varepsilon)/4}\right), \tag{4.3}$$

where the implicit constant in $O(\cdot)$ depends on ε and ς.

As an application of the Wigner case, we obtain an estimation of the probability that the adjacency matrix of a sparse random graph (not necessarily an Erdös-Rényi graph) is non-singular. Costello and Vu [12] have analyzed the adjacency matrices of sparse Erdös-Rényi graphs, where each entry is equal to 1 with the same probability $p(n)$ which tends to 0 as n goes to infinity (see also Costello and Vu [13] where a generalization of [11] is considered in which each entry takes the value $c \in \mathbb{C}$ with probability p and zero with probability $1-p$, and the diagonal entries are possibly non-zero). It is proved in [12] that when $c\ln(n)/n \le p(n) \le 1/2$, $c > 1/2$, then with probability $1-\mathrm{O}((\ln\ln(n))^{-1/4})$, the rank of the adjacency matrix equals the number of non-isolated vertices. Now we consider the following model extension of Erdös-Rényi graphs, where vertices i and j are linked with a probability that depends on i and j and the number of vertices. Furthermore, the rate of convergence is an improvement of the one given in [12] for $c\ln n/n^{\beta} \le p(n) \le 1/2$ with $c > 0$ and $\beta \in (0,1)$. From the proof of Theorem 13.b in section 4, if $\kappa_n = 1 - p(n)$, we have

$$\frac{\kappa_n^{\frac{3}{8}n - \frac{1}{2}n^{1-\varepsilon}}}{\kappa_n(1-\kappa_n)} \le \left(\frac{\kappa_n^2}{n^{1-\varepsilon}(1-\kappa_n)}\right)^{1/4} \le \left(\frac{(1 - c(\ln n/n^{\beta}))^2}{n^{1-\varepsilon-\beta}\ln n}\right)^{1/4} \to 0$$

as $n \to \infty$, if $\varepsilon + \beta < 1$.

Also, we notice that Theorem 13 provides us a first clue about what kind of matrices are "bad matrices" for testing a matrix algorithm. System of linear equations with Ginibre and Wigner matrices have solution with high probability for large variety of random entries. Hence, Theorem 13 tells us that we must understand the behavior of extreme singular values of a random matrix, as we will study in the next chapter for a circulant matrix, when we like to know the degree of singularity of a random matrix.

Proposition 1 ($\star$)**.** *Let $\{p_{ij} \in (0,1) : i,j = 1,2,\ldots\}$ be a double sequence of positive numbers with $p_n^* = \min_{1\le i\le j\le n}\{p_{ij}\} \in [c\ln n/n^\beta, 1/2]$, $c>0$, and $\varepsilon+\beta<1$, $\varepsilon,\beta\in(0,1)$, then there is a random graph with n vertices such that the vertex i is linked with the vertex j with probability p_{ij}, $1\le i<j\le n$, and if A_n is its the adjacency matrix, we have as $n\to\infty$*

$$\mathbb{P}\left\{\mathit{rank}(A_n) < n\right\} \le Cn^{-(1-\varepsilon-\beta)/4}, \tag{4.4}$$

for some constant $C>0$.

In the following sections we develop the proof of Theorem 13. For the rest of this chapter, all our random variables satisfy

$$\sup_{x\in\mathbb{R}} \mathbb{P}\{\xi = x\} \le \varsigma_\Delta(\xi) < \varsigma < 1.$$

4.3 Proof in the Ginibre case

We start with an extension of a result by Slinko [65], who treated the case of a discrete uniform distribution with parameter $1/q$ with $q\in\mathbb{Z}^+$.

Lemma 13. *Let $k\le m$ and let $A\in\mathbb{R}^{m\times k}$ be a (deterministic) matrix with $\mathrm{rank}(A)=k$. If $b\in\mathbb{R}^m$ is a random vector whose entries are independent random variables, then*

$$\mathbb{P}\left\{\mathit{rank}(A,b) = k\right\} \le \varsigma^{m-k}.$$

Proof. Since $\mathrm{rank}(A)=k$, we can decompose $[A\ b]$ in the following way

$$[A\ b] = \begin{pmatrix} A_k & b_k \\ A_{m-k} & b_{m-k} \end{pmatrix},$$

where $A_k\in\mathbb{R}^{k\times k}$, $A_{m-k}\in\mathbb{R}^{(m-k)\times k}$, $b_k\in\mathbb{R}^k$ and $b_{m-k}\in\mathbb{R}^{m-k}$. We note A_k is an invertible matrix. We have that there exists a random vector $\mathcal{D}\in\mathbb{R}^k$ such that $A_k\mathcal{D}=b_k$ and $A_{m-k}\mathcal{D}=b_{m-k}$, then $A_{m-k}A_k^{-1}b_k = b_{m-k}$. So

$$\begin{aligned}
\mathbb{P}\left\{r(A,b)=k\right\} &\le \mathbb{P}\left\{A_{m-k}A_k^{-1}b_k = b_{m-k}\right\} \\
&= \mathbb{E}\left\{\mathbb{P}\left\{A_{m-k}A_k^{-1}b_k = b_{m-k}\,\middle|A_{m-k}A_k^{-1}b_k\right\}\right\} \\
&\le \varsigma^{m-k},
\end{aligned}$$

the last inequality being due to the independence of every entry in b_{m-k}. $\square$

Lemma 14. *Let $k\le m$ and let $A\in\mathbb{R}^{m\times k}$ be a random matrix (whose entries are independent random variables). Then*

$$\mathbb{P}\left\{\mathit{rank}(A) < k\right\} < \frac{\varsigma}{1-\varsigma}\varsigma^{m-k}.$$

Proof. We note that if $A=[a_1|\cdots|a_k]$, $a_i\in\mathbb{R}^m$ $i=1,\ldots,k$, then

$$\begin{aligned}
\mathbb{P}\{\mathrm{rank}(A)=k\} &= \mathbb{P}\{a_1\notin\{0\}, a_2\notin\mathrm{span}\{a_1\},\ldots,a_k\notin\mathrm{span}\{a_1,a_2,\ldots,a_{k-1}\}\} \\
&= \mathbb{P}\{a_1\notin\{0\}\}\prod_{i=2}^{k}\mathbb{P}\{E_i\},
\end{aligned}$$

where we use the notation span$\{\cdot\}$ for the space generated by some vectors and

$$E_i = \{a_i \notin \text{span}\{a_1, a_2, \ldots, a_{i-1}\} \,|a_1 \notin \{0\}, a_2 \notin \text{span}\{a_1\}, \ldots, a_{i-1} \notin \text{span}\{a_1, a_2, \ldots, a_{i-2}\}\}.$$

Hence by lemma 13 and the Weierstrass product inequality[1],

$$\mathbb{P}\{\text{rank}(A) = k\} \geq \prod_{i=0}^{k-1}(1-\kappa^{m-i}) \geq 1 - \sum_{i=0}^{k-1}\kappa^{m-i} = 1 - \frac{\kappa}{1-\kappa}\kappa^{m-k}.$$

□

We consider the following concept used by Komlós [8]. Let $S = \{v_1, \ldots, v_n\}$ be a set of vectors. Let us define the strong rank of S, denoted $sr(S)$, to be n if S is a set of linearly independent vectors, and k if any k of the v_i's are linearly independent but some $k+1$ of the vectors are linearly dependent. For a matrix A, we denote the strong rank of the system of columns and the strong rank of the system of rows by $sr_c(A)$ and $sr_r(A)$, respectively.

Remark 1. *(a) Let A be an $m \times n$ random matrix with all entries independent random variables. It follows immediately from lemma 14 that*

$$\mathbb{P}\left\{sr_c(A) < k\right\} \leq \binom{n}{k}\frac{\kappa}{1-\kappa}\kappa^{m-k}.$$

(b) For every ς and $0 < \alpha \leq 1$ there exists $\beta > 0$ which satisfies

$$\frac{h(\beta)}{\log_2 \varsigma} + \beta < \alpha < 1, \tag{4.5}$$

where $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the entropy function. Indeed, let

$$g(x) = \frac{h(x)}{\log \varsigma} + x.$$

Now, since the function g is continuous and $g(0) = 0$, there exists a positive number $\beta > 0$, which depends on ς, such that $g(\beta) < \alpha < 1$.

c) We note from (a) and (b) that if $m = \lfloor \alpha n \rfloor$ and $k = \lceil \beta n \rceil$, then

$$\mathbb{P}\left\{rank(A) < \lceil \beta n \rceil\right\} < \binom{n}{\lceil \beta n \rceil}\frac{\varsigma}{1-\varsigma}\varsigma^{\lfloor \alpha n \rfloor - \lceil \beta n \rceil} < \frac{\varsigma}{1-\varsigma}2^{n(h(\beta)-(\alpha-\beta)\log_2(\varsigma))} < \frac{\varsigma}{1-\varsigma}2^{-n\gamma_\varsigma},$$

where we use $\binom{n}{\beta n} < 2^{nh(\beta)}$ and γ_ς is a positive constant which depends on ς.

Lemma 15. *Let $v_1, v_2, \ldots, v_k \in \mathbb{R}^m$ be (deterministic) linearly independent vectors. Let $B = [v_1|\ldots|v_k]$ and $sc_r(B) = s$. Then for a random vector $a \in \mathbb{R}^m$ whose entries are independent random variables,*

$$\mathbb{P}\left\{rank(v_1, v_2, \ldots, v_k, a) = k\right\} < C_1\varsigma^{m-k}s^{-1/2}.$$

[1]Weierstrass product inequality. For $0 \leq a_1, a_2, \ldots, a_n \leq 1$, then

$$\prod_{k=1}^{n}(1-a_k) + \sum_{k=1}^{n}a_k \geq 1.$$

Proof. Although simple, for the sake of completeness we include the proof. Let $b_1, b_2, \ldots, b_m$ be the rows of B. Without loss of generality we may assume that $b_1, b_2, \ldots, b_k$ are linearly independent and that all other rows are linear combination of them. We have

$$\sum_{i=1}^{k} \beta_i^{(r)} b_i = b^{(r)}$$

for $r = k+1, \ldots, m$. As $sc_r(B) = s$, at least s of the coefficients $\beta_1^{(r)} \ldots, \beta_k^{(r)}$ are nonzero.

Now, since we consider the event $[\text{rank}(v_1, v_2, \ldots, v_k, a) = k]$, we have

$$\sum_{j=1}^{k} \alpha_j v_j = a$$

for some $\alpha_1 \ldots, \alpha_k$ not all zero. In particular $\sum_{j=1}^{k} \alpha_j v_{k+1,j} = a_{k+1}$, where a_{k+1} is the $(k+1)$th entry of a. But

$$a_{k+1} = \sum_{j=1}^{k} \alpha_j v_{k+1,j} = \sum_{j=1}^{k} \alpha_j \left(\sum_{i=1}^{k} \beta_i^{(k+1)} v_{i,j} \right) = \sum_{i=1}^{k} \beta_i^{(k+1)} \left(\sum_{j=1}^{k} \alpha_j v_{i,j} \right) = \sum_{i=1}^{k} \beta_i^{(k+1)} a_i.$$

From the above and the independence of the entries of a,

$$\begin{aligned}
\mathbb{P}\left\{\text{rank}(v_1, v_2, \ldots, v_k, a) = k\right\} &\leq \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(r)} a_i = a_r, r = k+1, \ldots, m \right\} \\
&= \mathbb{E}\left\{ \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(r)} a_i = a_r, r = k+1, \ldots, m \mid a_1, \ldots, a_k \right\} \right\} \\
&= \mathbb{E}\left\{ \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(m)} a_i = a_m \mid a_1, \ldots, a_k \right\} \prod_{l=k+1}^{m-1} \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(l)} a_i = a_l \mid a_1, \ldots, a_k \right\} \right\} \\
&\leq \mathbb{E}\left\{ \varsigma^{m-k-1} \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(m)} a_i = a_m \mid a_1, \ldots, a_k \right\} \right\} \\
&= \varsigma^{m-k-1} \mathbb{P}\left\{ \sum_{i=1}^{k} \beta_i^{(m)} a_i = a_m \right\} \\
&\leq C_1 \varsigma^{m-k} s^{-1/2},
\end{aligned}$$

the last line being due to Theorem 11 and expression (3.20). □

Proof of Theorem 13.a. Let $\alpha \in (0,1)$ and $\beta > 0$ be as in expression (4.5) and let $n_0 = \lfloor \alpha n \rfloor$. Let B be the $n_0 \times n$ matrix whose columns are the first n_0 columns of G_n.

From lemma 14 we can assume that B has full rank. Since

$$\mathbb{P}\{\text{rank}(G_n) = n\} = \mathbb{P}\{\text{rank}(G_n) = n, sr_r(B) < \beta n\} + \mathbb{P}\{\text{rank}(G_n) = n, sr_r(B) \geq \beta n\},$$

by lemma 15 and remark 1, we have

$$\mathbb{P}\{\text{rank}(G_n) = n\} \geq \prod_{i=1}^{n-n_0} \left(1 - C_1(\beta n)^{-1/2}\varsigma^i\right) \geq 1 - \frac{C_1}{1-\varsigma}(\beta n)^{-1/2},$$

which proves Theorem 13.a. □

We now turn to Theorem 13. A natural question is to understand what happens when $\varsigma_n \to 1$. The following proposition says something about this situation.

Proposition 2 ($\star$). *For any sequence $\{\varsigma_n \in [0,1] : n \geq 1\}$ there is a sequence $\{G_{m_n} = (\xi_{i,j})_{1\leq i,j\leq m_n}\}$ such that:*

- *G_{m_n} is a $m_n \times m_n$ Ginibre matrix*
- *$\xi_{i,j}$, $1 \leq i,j \leq m_n$, have the same distribution F_{m_n}*
- *ς_n is the maximum jump of F_{m_n}*
- *$\mathbb{P}\{G_{m_n}$ has full rank $\} \to 1 \quad n \to \infty$*

Proof. Let F_1 be a distribution function whose biggest jump is ς_1. We take $m_n = 1$ and $\delta_1 = \varsigma_1/2$, then $\mathbb{P}\{G_{m_1}$ has full rank $\} > 1 - \delta_1$. Now, let F_n be a distribution function whose biggest jump is ς_n. By Lemma 2 in [40], there is $m_n \geq m_{n-1}$ and $\delta_n \leq 1/n \leq$ for $n > 1$ such that

$$\mathbb{P}\{G_{m_n} \text{ has full rank }\} > 1 - \delta_n,$$

where the entries of G_{m_n} have the same distribution and $\delta_n \to 0$ as $n \to \infty$. □

In the following examples we can see that if $\varsigma_n \to 1$ at some appropiate rate, the probability of a singularity can behave differently.

We write $GB(n,p)$ ($WB(n,p)$) for a $n \times n$ Ginibre (Wigner) matrix whose entries obey a Bernoulli distribution on $\{0,1\}$ with parameter p.

Let ZGB_n (ZWB_n) be the event that the first row of $GB(n,1/n)$, ($WB(n,1/n)$) contains only zeros. Then

$$\mathbb{P}\left\{ZGB_n\right\} = \left(1 - \frac{1}{n}\right)^n, \quad \mathbb{P}\left\{ZWB_n\right\} = \left(1 - \frac{1}{n}\right)^n,$$

and hence

$$e^{-1} \leq \lim_{n\to\infty} \mathbb{P}\left\{\text{rank}\left(GB\left(n, 1/n\right)\right) < n\right\},$$

$$e^{-1} \leq \lim_{n\to\infty} \mathbb{P}\left\{\text{rank}\left(WB\left(n, 1/n\right)\right) < n\right\}.$$

However, if $\alpha \in (0,1)$, then there is a constant $C_\alpha > 0$

$$\mathbb{P}\left\{\text{rank}\left(WB\left(n, n^\alpha/n\right)\right) < n\right\} \leq n^{-C_\alpha}. \tag{4.6}$$

In the Ginibre case it is not clear what happens when $\varsigma_n = n^\alpha/n$, but if $\gamma \in (0,1)$, then

$$\mathbb{P}\{\text{rank}\left(GB\left(n, n^\alpha/n\right)\right) > \gamma n\} \to 1 \text{ as } n \to \infty. \tag{4.7}$$

4.4 Proofs in the Wigner case

Following the terminology introduced in Costello, Tao and Vu [11], given n vectors $\{v_1, \ldots, v_n\}$, a linear combination of the v_i's is a vector $v = \sum_{i=1}^n c_i v_i$, where the c_i are real numbers. We say that a linear combination vanishes if v is the zero vector. A vanishing linear combination has degree k if exactly k among the c_i are nonzero.

A singular $n\times n$ matrix is called *normal* if its row vectors do not admit a non-trivial vanishing linear combination with degree less than $n^{1-\varepsilon}$ for a given $\varepsilon\in(0,1)$. Otherwise it is said that the matrix is *abnormal*. Furthermore, a row of an $n\times n$ non-singular matrix is called *good* if its exclusion leads to an $(n-1)\times n$ matrix whose column vectors admit a non-trivial vanishing linear combination with degree at least $n^{1-\varepsilon}$ (in fact, there is exactly one such combination, up to scaling, as the rank of this $(n-1)\times n$ matrix is $n-1$). A row is said to be *bad* otherwise. Finally, an $n\times n$ non-singular matrix A is *perfect* if every row in A is good. If a non-singular matrix is not perfect, it is called *imperfect*.

For the proof of Theorem 13.b, we first present three lemmas which generalize results in [11] for Wigner matrices $W_n=(\xi_{ij})$ with independent entries which need not be identically distributed and the appropriate estimates in these new cases are found in terms of the size of the biggest jump of the distribution functions governing the entries under the hypothesis $\varsigma_\Delta(i)<\varsigma<1$. We also obtain a better rate of convergence, which is universal. The proofs we give follow ideas in [11] but also take into account the size of the biggest jump.

Lemma 16. *Let $\varepsilon\in(0,1)$, then for all n large*

$$\mathbb{P}\left\{W_n \textit{ is singular and abnormal}\right\}\le\varsigma^{(n-n^{1-\varepsilon})/2} \tag{4.8}$$

and

$$\mathbb{P}\left\{W_n \textit{ is non-singular and imperfect}\right\}\le\varsigma^{(n-n^{1-\varepsilon})/2}. \tag{4.9}$$

Proof. If W_n is singular and abnormal the row vectors of W_n admit a non-trivial vanishing linear combination with degree at most $N:=n^{1-\varepsilon}$. For $i=1,\dots,N$, we have that if $i=1$, there is a row of W_n that contains only zeros, and if $i>1$, the ith row is a linear combination of the first $i-1$ rows of W_n that are linearly independent. We denote by $D(n,i)$ this last event and by T_{i-1} the upper triangular part of W_n until the row $i-1$ (included). The linear dependence of the ith row of W_n with the $i-1$ rows of W_n is determined only by its last $n-i+1$ entries. Then by the stochastic independence of T_{i-1} with the last $n-i+1$ entries of the row i

$$\begin{aligned}\mathbb{P}\left\{W_n \text{ is singular and abnormal}\right\}&\le\sum_{i=1}^{N}\binom{n}{i}\mathbb{P}\left\{D(n,i)\right\}\le\sum_{i=1}^{N}\binom{n}{i}\mathbb{E}\left\{\mathbb{P}\left\{D(n,i)\middle|\,T_{i-1}\right\}\right\}\\&\le\sum_{i=1}^{N}n^N\varsigma^{n-N+1}=Nn^N\varsigma^{n-N+1},\end{aligned}$$

and for all n large,

$$\mathbb{P}\left\{W_n \text{ is singular and abnormal}\right\}\le\varsigma^{\frac{3}{4}(n-n^{1-\varepsilon})}\le\varsigma^{\frac{1}{2}(n-n^{1-\varepsilon})}.$$

Now, we consider the case when W_n is non-singular and imperfect. We can suppose that the last row of W_n is the bad row. The $(n-1)\times n$-matrix obtained has rank $n-1$, hence there is a unique column that admits a non-trivial vanishing linear combination with degree at most $n^{1-\varepsilon}$. Then the last $n-k-1$ entries of this column are completely determined by its k first entries and k linearly independent columns, for $1\le k\le n^{1-\varepsilon}$. Since we can choose this bad row, we have as above for n large

$$\mathbb{P}\left\{W_n \text{ is non-singular and imperfect}\right\}\le n\varsigma^{\frac{3}{4}(n-1-(n-1)^{1-\varepsilon})}\le\varsigma^{\frac{1}{2}(n-n^{1-\varepsilon})}.$$

□

Lemma 17. *Let A be a deterministic $n \times n$ singular normal matrix. Then*

$$\mathbb{P}\left\{rank(W_{n+1}) - rank(W_n) < 2 \,|W_n = A\right\} = O_\varepsilon\left(\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right).$$

Proof. Since $r := \mathrm{rank}(A) < n$, without loss of generality it is possible to suppose that the first r rows of A are linearly independent. If $v_1, \dots, v_r$ are the first rows of A, then $v_n = \sum_{i=1}^r \alpha_i v_i$, and as A is normal, the number of coefficients in this linear combination is at least $n^{1-\varepsilon}$. If it does not hold that $\xi_n = \sum_{i=1}^r \alpha_i \xi_i$, where ξ_i are entries of the last column of W_{n+1}, by symmetry of W_{n+1} we have $\mathrm{rank}(W_{n+1}) = \mathrm{rank}(A) + 2$. Hence

$$\begin{aligned}\mathbb{P}\left\{\mathrm{rank}(W_{n+1}) - \mathrm{rank}(W_n) < 2 \,|W_n = A\right\} &\leq \mathbb{P}\left\{\xi_n = \sum_{i=1}^r \alpha_i \xi_i\right\} \\ &= \mathrm{O}_\varepsilon\left(\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right).\end{aligned}$$

The last expression follows from expression (3.20). □

Lemma 18. *Let A be a deterministic $n \times n$ non-singular perfect symmetric matrix. Then*

$$\mathbb{P}\left\{rank(W_{n+1}) = n \,|W_n = A\right\} = O_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right).$$

Proof. If $\mathrm{rank}(W_{n+1}) = n$, then $\det(W_{n+1}) = 0$, and we have

$$0 = \det(W_{n+1}) = (\det A)\xi_{n+1} + \sum_{i=1}^n \sum_{j=1}^n c_{ij} \xi_i \xi_j,$$

where ξ_i are entries of the last column of W_{n+1} and its transpose, and the c_{ij} are the cofactors of A. Since A is perfect, when we eliminate the ith row of A, the columns of the matrix thus obtained admit a vanishing linear combination of degree at least $n^{1-\varepsilon}$. When the column j is selected, where j is the index of a non-zero coefficient in this linear combination, we obtain an $(n-1) \times (n-1)$ non-singular matrix since there are at least $n^{1-\varepsilon}$ indices i such that there are at least $n^{1-\varepsilon}$ indices j with $c_{i,j} \neq 0$. Taking the partition of $\{1, 2, \dots, n\}$ as $S_1 = \{1, 2, \dots, \lfloor n/2 \rfloor\}$

and $S_2 = \{1, 2, \dots, n\} - S_1$, by expression (3.21)

$$\begin{aligned}
\mathbb{P}\{\operatorname{rank}(W_{n+1}) = n \,|W_n = A\} &\le \mathbb{P}\left\{(\det A)\xi_{n+1} + \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\xi_i\xi_j = 0\right\} \\
&= \mathbb{E}\left(\mathbb{P}\left\{(\det A)\xi_{n+1} + \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\xi_i\xi_j = 0 \,|\, \xi_{n+1}\right\}\right) \\
&= \mathbb{E}\left(\mathrm{O}_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right)\right) \\
&= \mathrm{O}_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right).
\end{aligned}$$

□

Now we consider the discrete stochastic process

$$X_n = \begin{cases} 0 & \text{if } \operatorname{rank}(W_n) = n \\ \left(\varsigma^{-1/8}\right)^{n-\operatorname{rank}(W_n)} & \text{if } \operatorname{rank}(W_n) < n, \end{cases}$$

for which we can prove the following result.

Proposition 3.

$$\mathbb{E}(X_n) = O_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right).$$

Proof. For $j = 0, \dots, n$, write $A_j = \{\operatorname{rank}(W_n) = n - j\}$ and let $1 + \gamma = \varsigma^{-1/8}$. We have

$$\begin{aligned}
\mathbb{E}(X_n) &= \sum_{j=1}^{n}(1+\gamma)^j \mathbb{P}\{A_j\} \\
&= \sum_{j=1}^{n}(1+\gamma)^j \mathbb{P}\{A_j,\, W_n \text{ normal}\} + S_1,
\end{aligned}$$

where

$$S_1 = \sum_{j=1}^{n}(1+\gamma)^j \mathbb{P}\{A_j,\, W_n \text{ abnormal}\}.$$

By lemma 16,

$$\begin{aligned}
S_1 &\le \sum_{j=1}^{n}(1+\gamma)^j \varsigma^{(n-n^{1-\varepsilon})/2} \\
&\le \varsigma^{(n-n^{1-\varepsilon})/2}\sum_{j=1}^{n}(1+\gamma)^j \\
&\le \frac{1-(\varsigma^{-1/8})^{n+1}}{1-\varsigma^{-1/8}}\varsigma^{(n-n^{1-\varepsilon})/2} \\
&= C\varsigma^{(3n-4n^{1-\varepsilon})/8}
\end{aligned}$$

for some constant $C > 0$.

So

$$\mathbb{E}\left(X_n\right) = \sum_{j=1}^{n}(1+\gamma)^j \mathbb{P}\left\{A_j,\, W_n \text{ normal}\right\} + \mathrm{O}_\varepsilon\left(\varsigma^{(3n-4n^{1-\varepsilon})/8}\right). \tag{4.10}$$

On the other hand,

$$\mathbb{E}\left(X_{n+1}\right) = S_2 + S_3 + S_4 + S_5,$$

where

$$\begin{aligned}
S_2 &= \mathbb{E}\left(X_{n+1}\,|A_0, W_n \text{ perfect}\right)\mathbb{P}\left\{A_0,\, W_n \text{ perfect}\right\} \\
S_3 &= \mathbb{E}\left(X_{n+1}\,|A_0, W_n \text{ imperfect}\right)\mathbb{P}\left\{A_0,\, W_n \text{ imperfect}\right\} \\
S_4 &= \sum_{j=1}^{n}\mathbb{E}\left(X_{n+1}\,|A_j, W_n \text{ normal}\right)\mathbb{P}\left\{A_j,\, W_n \text{ normal}\right\} \\
S_5 &= \sum_{j=1}^{n}\mathbb{E}\left(X_{n+1}\,|A_j, W_n \text{ abnormal}\right)\mathbb{P}\left\{A_j,\, W_n \text{ abnormal}\right\}.
\end{aligned}$$

By lemma 18 and the fact that $\operatorname{rank}(W_n) = n$,

$$\begin{aligned}
S_2 &\le (\varsigma^{-1/8})^{n+1-n}\mathbb{P}\{\operatorname{rank}(W_{n+1}) = n\,|W_n \text{ is perfect and non-singular}\} \\
&= \mathrm{O}_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right).
\end{aligned}$$

On the other hand, lemma 16 and the definition of X_{n+1} give

$$S_3 \le (\varsigma^{-1/8})^{n+1}\varsigma^{(n-n^{1-\varepsilon})/2} = \mathrm{O}_\varepsilon\left(\varsigma^{(3n-4n^{1-\varepsilon})/8}\right).$$

Using again lemma 16 and the definition of A_j,

$$S_5 \le \sum_{j=1}^{n}(\varsigma^{-1/8})^{j+1}\varsigma^{(n-n^{1-\varepsilon})/2} = \mathrm{O}_\varepsilon\left(\varsigma^{(3n-4n^{1-\varepsilon})/8}\right).$$

If $\operatorname{rank}(W_n) = n-j$, then $\operatorname{rank}(W_{n+1})$ is equal to $n-j+2$ or $n-j$ since W_{n+1} is a symmetric matrix. By lemma 17 and for n sufficiently large,

$$\begin{aligned}\mathbb{E}\left(X_{n+1}\,|A_j, W_n \text{ normal}\right) &= (1+\gamma)^{j+1}\mathbb{P}\{\operatorname{rank}(W_{n+1}) = \operatorname{rank}(W_n)\,|W_n \text{ normal and singular}\} \\ &\quad + (1+\gamma)^{j-1} \\ &= (1+\gamma)^j\left((1+\gamma)^{-1} + \mathrm{O}_\varepsilon\left(\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right)\right) \\ &\le \alpha(1+\gamma)^j\end{aligned}$$

for some $\alpha < 1$.

Then we have

$$\mathbb{E}(X_{n+1}) = \alpha\sum_{j=1}^{n}(1+\gamma)^j\mathbb{P}\{A_j, W_n \text{ normal}\} + \mathrm{O}_\varepsilon\left(f(\varsigma, n)\right),$$

where

$$f(\varsigma, n) := \frac{\varsigma^{\frac{3}{8}n - \frac{1}{2}n^{1-\varepsilon}}}{\varsigma(1-\varsigma)} + \left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}.$$

Using the expression (4.10)

$$\mathbb{E}(X_{n+1}) \le \alpha\mathbb{E}(X_n) + \mathrm{O}_\varepsilon\left(f(\varsigma, n)\right),$$

so

$$\mathbb{E}(X_{n+1}) \le \alpha^n\mathbb{E}(X_1) + \mathrm{O}_\varepsilon\left(f(\varsigma, n)\right).$$

This proves the proposition. □

Proof of Theorem 13.b. By Markov's inequality,

$$\begin{aligned}\mathbb{P}\left\{\operatorname{rank}(W_n) < n\right\} &= \mathbb{P}\left\{X_n \ge 1\right\} \\ &\le \mathbb{E}\left(X_n\right) \\ &= \mathrm{O}_\varepsilon\left(\left[\frac{\varsigma}{\sqrt{n^{1-\varepsilon}(1-\varsigma)^3}}\right]^{1/2}\right),\end{aligned} \tag{4.11}$$

where we have used proposition 3. □

Chapter 5

Circulant Random Matrices

5.1 Introduction

This chapter is about our research in the analysis of the minimum singular value $s_n(\mathcal{C}_n)$ of a circulant random matrix $\mathcal{C}_n$.

In Section 5.2 we give the main contributions to the question of the singularity of a circulant random matrix, and a new result of independent interest in the theory of random polynomials. Theorem 14 shows that when the first row of $\mathcal{C}_n$ has independent and identically distributed random variable entries with moment generating function and zero mean, we have $s_n(\mathcal{C}_n) \geq \varepsilon n^{-1/2}$ with high probability, for any $\varepsilon > 0$. A classic result in random polynomial theory says that the roots of a random polynomial become concentrated near the unit circle as the degree of the polynomial goes to infinity, with probability one [4]. Theorem 15 determines the speed of the movement of the roots of a random polynomial towards the unit circle. The proof of Theorem 14 is then a direct application of Theorem 15, whose proof follows ideas in [42]. Theorem 16 states that when a circulant random matrix has prime dimension and its entries are allowed to have general distribution (no moment assumptios), it is invertible with high probability. The proof of this result uses some properties of the concentration of a linear combination of Rademacher random variables mentioned in [51].

In Sections 5.3, 5.4, and 5.5, we give the proofs of the above three theorems in this chapter. Section 5.6 presents additional contributions. Theorem 19 gives an upper bound for the expectation of the maximum singular value of a circulant random matrix whose entries are sub-Gaussian random variables not necessarily indenpendent. Corollary 1 gives some additional results about the minimum singular value of g-circulant matrices, which are a generalization of circulant matrices. Finally, using a result on random polynomials whose coefficients are independent but not identically distributed random variables, we establish a condition for the singular value of a circulant random matrix to be large.

In a personal communication from S. V. Koyagin, I was told that he and his student, A. G. Karapetyan, did not continue to work on the minimum value of a random polynomial near the unit circle, and he encouraged me to study in detail the case of sub-Gaussian variables.

The main results in this chapter are collected in the coauthored with Gerardo Barrera in manuscript [3].

5.2 Singularity of a circulant random matrix

An $n \times n$ circulant matrix $\mathrm{circ}(c_0, \ldots, c_{n-1})$ has the form

$$\mathrm{circ}(c_0, \ldots, c_{n-1}) := \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-3} & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_1 & c_2 & \ddots & \vdots & c_0 \end{bmatrix},$$

where $c_0, \ldots, c_{n-1} \in \mathbb{C}$. It is well known that any circulant matrix can be diagonalized in $\mathbb{C}$ as follows: Let $\omega_n := \exp\left(i\frac{2\pi}{n}\right)$, $i^2 = -1$, and

$$F_n := \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_n & \omega_n^2 & \cdots & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^4 & \cdots & \omega_n^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \cdots & \omega_n^{(n-1)(n-1)} \end{bmatrix}.$$

The matrix F_n is called the *Fourier matrix* of order n. Note that F_n is a unitary matrix. By a straightforward computation, one can readily verify that

$$\mathrm{circ}(c_0, \ldots, c_{n-1}) = F_n^* \mathrm{diag}\left(G_n(1), G_n(\omega_n), \ldots, G_n(\omega_n^{n-1})\right) F_n,$$

where G_n is the polynomial given by

$$G_n(z) := c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}.$$

We have that the eigenvalues of $\mathrm{circ}(c_0, \ldots, c_{n-1})$ are $G_n(1), G_n(\omega_n), \ldots, G_n(\omega_n^{n-1})$, or equivalently

$$G_n(\omega_n^k) = \sum_{j=0}^{n-1} c_j \exp\left(i\frac{2\pi k j}{n}\right) \qquad k = 0, \ldots, n-1. \tag{5.1}$$

Now, we consider an $n \times n$ random circulant matrix $\mathcal{C}_n$ with independent entries, i.e., $\mathcal{C}_n := \mathrm{circ}(\xi_0, \ldots, \xi_{n-1})$, where $\xi_0, \ldots, \xi_{n-1}$ are independent random variables.

The maximum and minimum singular values of a circulant matrix $\mathcal{C}_n$ are given by

$$s_1(\mathcal{C}_n) = \max_{0 \leq k \leq n-1} \left| G_n(\omega_n^k) \right|$$

and

$$s_n(\mathcal{C}_n) = \min_{0 \leq k \leq n-1} \left| G_n(\omega_n^k) \right|.$$

If $\xi_0, \ldots, \xi_{n-1}$ have continuous distribution, then

$$\mathbb{P}\left(\mathcal{C}_n \text{ is singular}\right) = 0$$

since $\mathbb{P}\left(G_n(\omega_n^k) = 0\right) = 0$ for all k. If $\xi_0, \ldots, \xi_{n-1}$ have discrete distribution, estimate the probability that $\mathcal{C}_n$ is singular is not easy.

Note

$$\min_{z\in\mathbb{C}:|z|=1} |G_n(z)| \leq s_n(\mathcal{C}_n). \tag{5.2}$$

It is clear that ω_n^k satisfies $|\omega_n^k| = 1$ for every $k = 0, \ldots, n-1$. So, if G_n does not have any root in $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$, we have that $\mathcal{C}_n$ is a non-singular matrix.

In [27], it was shown that if $G_n(z) = \sum_{j=0}^{n-1} \xi_j z^j$ is a random polynomial with (real or complex) i.i.d. coefficients, its roots are asymptotically concentrated around $\mathbb{T}$ as $n \to \infty$ almost surely. Moreover, it was also proved that the condition $\mathbb{E}(\log(1 + |\xi_0|)) < \infty$ is necessary and sufficient for the roots of G_n to be asymptotically near the unit circle. If $z^*_{n,j}$, $j = 0, \ldots, n-1$ are the roots of G_n, we have for all $\varepsilon > 0$

$$\min_{0\leq j,k\leq n-1} \left| z^*_{n,j} - \omega_n^k \right| < \varepsilon \quad \text{as } n \to \infty \text{ a.s.}$$

The left–hand side of (5.2) was studied in [32], [34], [41] and [42]. In [42], it was shown that if G_n has i.i.d. Rademacher or standard normal random coefficients, then for all $\varepsilon > 0$ and large n, we have with high probability

$$\min_{z\in\mathbb{C}:|z|=1} |G_n(z)| > \varepsilon n^{-1/2}.$$

In [32] and [34] the sub-Gaussian case was studied, but there was no proof. Even so, we give a generalization of the main result in [42], which includes the sub-Gaussian case.

Theorem 15 is itself an interesting result about random polynomials because it provides a fine estimate of the distance between the roots of G_n and the unit circle. Many results on random polynomials are about the location of their roots via the convergence of the empirical distribution of the roots of G_n.

The main results in this chapter are the following:

Theorem 14 ($\star$ Minimum singular value of a circulant random matrix)**.** *Let ξ be a random variable with moment generating function such that $\mathbb{E}(\xi) = 0$ and $\mathbb{E}\left(\xi^2\right) = \sigma^2 > 0$. Let $\{\xi_k\}_{k\geq 0}$ be a sequence of independent random variables with $\xi_k \overset{D}{=} \xi$ for every $k \geq 0$. Let $\mathcal{C}_n := \operatorname{circ}(\xi_0, \ldots, \xi_{n-1})$ be an $n \times n$ circulant matrix. Then for all $\varepsilon > 0$ and for all large n,*

$$\mathbb{P}(s_n(\mathcal{C}_n) \geq \varepsilon n^{-1/2}) \geq 1 - C\varepsilon,$$

where C is a constant depending on ξ.

Theorem 15 ($\star$ Roots of a random polynomial)**.** *Let ξ be a random variable with moment generating function such that $\mathbb{E}(\xi) = 0$ and $\mathbb{E}\left(\xi^2\right) = \sigma^2 > 0$. Let $\{\xi_k\}_{k\geq 0}$ be a sequence of independent random variables with $\xi_k \overset{D}{=} \xi$ for every $k \geq 0$. Let $\phi : [0,1] \to \mathbb{R}$ be a non-zero Hölder continuous function of order $\varsigma \in (1/2, 1]$. Then for any $\varepsilon > 0$,*

$$\limsup_{n\to\infty} \mathbb{P}\left(\min_{z\in\mathbb{C}:||z|-1|<\varepsilon n^{-2}} \left| \sum_{j=0}^{n-1} \xi_j \phi(j/n) z^j \right| < \varepsilon n^{-1/2} \right) \leq C\varepsilon,$$

where C is a constant depending on ϕ and ξ.

The following result considers only the non-singularity of $\mathcal{C}_n$, when its size is a prime number. This result is interesting because it considers a more general kind of random variables.

Theorem 16 ($\star$)**.** *Let $\mathcal{X} = \{X_k\}_{k\geq 0}$ be a sequence of i.i.d. Rademacher random variables. Let $\mathcal{Y} = \{Y_k\}_{k\geq 0}$ be a set of independent random variables which are independent of those in $\mathcal{X}$. We define $\xi_k := X_k + Y_k$ for $k \geq 0$. If n is prime, then*

$$\mathbb{P}(\mathrm{circ}(\xi_0 \ldots, \xi_{n-1}) \text{ is singular}) = \mathrm{O}(n^{-1/2}),$$

where the constant implicit in O *is independent of the random variables in $\mathcal{Y}$.*

Remark. Throughout the proofs, absolute constants will be always denoted by C. So its particular value can be different in different instances.

5.3 Minimum singular value of a circulant random matrix

Proof of Theorem 14. If we take the Hölder function as $\phi \equiv 1$ in Theorem 15, we have

$$\liminf_{n\to\infty} \mathbb{P}\left(\min_{z\in\mathbb{C}:||z|-1|<\varepsilon n^{-2}} |G_n(z)| \geq \varepsilon n^{-1/2}\right) \geq 1 - C\varepsilon.$$

Then, there is $N := N(C\varepsilon) \in \mathbb{N}$ such that for all $n \geq N$

$$\begin{aligned} 1 - 2C\varepsilon &\leq \mathbb{P}\left(\min_{z\in\mathbb{C}:||z|-1|<\varepsilon n^{-2}} |G_n(z)| \geq \varepsilon n^{-1/2}\right) \\ &\leq \mathbb{P}\left(\min_{0\leq k\leq n-1} |G_n(\omega_n^k)| \geq \varepsilon n^{-1/2}\right). \end{aligned}$$

$\square$

5.4 Roots of a random polynomial

Proof of Theorem 15. The proof is based on [42]. We assume that $\varsigma \in (1/2, 1/2 + 1/20)$ and that

$$||\phi||_{C^\varsigma} := \max_{0\leq t\leq 1} |\phi(t)| + \sup_{0\leq t<s\leq 1} \frac{|\phi(t) - \phi(s)|}{|t-s|^\varsigma} = 1,$$

and we write

$$T_n(x) := \sum_{j=0}^{n-1} \xi_j \phi\left(j/n\right) e^{ijx}, \quad x \in [0, 2\pi].$$

Let

$$\{y_\beta\}_{\beta=1}^B = \left\{2\pi\frac{h}{k} : 1 \leq k \leq A, 0 \leq h \leq k-1, (k,h) = 1\right\},$$

where A and B are constants depending only on ϕ and where A is as specified in [42]. Fix $\varepsilon > 0$ and split $\mathbb{T} = [0, 2\pi]$ into non-overlapping intervals I_α of lengths between $\frac{1}{2}\varepsilon n^{-2}$ and εn^{-2}.

The intervals $J_\beta = \left[y_\beta - 2\pi n^{-1+\varsigma/20}, y_\beta + 2\pi n^{-1+\varsigma/20}\right]$, $\beta = 1, 2, \ldots, B$, will be called bad. We define I_α to be good provided $I_\alpha \not\subset \cup_{\beta=1}^B J_\beta$. For any such I_α, fix $x_\alpha \in I_\alpha \setminus \cup_{\beta=1}^B J_\beta$. Write

$$\begin{aligned} \mathcal{N} &:= \left\{z \in \mathbb{C} : ||z| - 1| < \varepsilon n^{-2}\right\}, \\ \mathcal{G} &:= \left\{||T_n'||_\infty \leq C_0 n^{3/2} \log^{1/2}(n), \ \sup_{z\in\mathcal{N}} |G_n''(z)| \leq n^{13/4}\right\}, \end{aligned}$$

where C_0 is a sufficiently large constant. For any interval $I \subset \mathbb{T}$, we write $\mathcal{E}(I) := \{e^{ix} : x \in I\}$. Let $D(z_0, \rho) := \{z \in \mathbb{C} : |z - z_0| < \rho\}$. Now, we have

$$\begin{aligned}\mathbb{P}\left(\min_{z\in\mathcal{N}} |G_n(z)| < \varepsilon n^{-1/2}\right) \quad &\le \quad \sum_{\alpha} \mathbb{P}\left(\min_{z\in D(e^{ix_\alpha}, 2\varepsilon n^{-2})} |G_n(z)| < \varepsilon n^{-1/2}, \mathcal{G}\right) \\ &\quad + \mathcal{P}_{0,n} + \mathbb{P}\left(\mathcal{G}^c\right),\end{aligned}$$

where

$$\mathcal{P}_{0,n} := \sum_{\beta=1}^{B} \mathbb{P}\left(\min_{z\in\mathcal{N}, z/|z|\in\mathcal{E}(J_\beta)} |G_n(z)| < \varepsilon n^{-1/2}, \mathcal{G}\right).$$

To avoid the loss of a logarithmic factor, we shall use the Taylor polynomials of T_n of order two around e^{ix_α} to estimate the sum over α. If the event $\mathcal{G}$ occurs, then

$$G_n(z) = T(x_\alpha) - (z - e^{ix_\alpha}) i e^{-ix_\alpha} T_n'(x_\alpha) + \mathrm{O}(\varepsilon^2 n^{-3/2}) \text{ for all } z \in D(e^{ix_\alpha, 2\varepsilon n^{-2}}).$$

Hence, if $|G_n(z)| < \varepsilon n^{-1/2}$ for some $z \in D(e^{ix_\alpha, 2\varepsilon n^{-2}})$, then

$$\left|T_n(x_\alpha) - (z - e^{ix_\alpha}) i e^{-ix_\alpha} T_n'(x_\alpha)\right| < 2\varepsilon n^{-1/2}$$

for large n. Consequently, if also $|T_n(x_\alpha)| \ge 4\varepsilon n^{-2} |T_n'(x_\alpha)|$, then $|T_n(x_\alpha)| < 4\varepsilon n^{-1/2}$. We conclude that for each I_α,

$$\mathbb{P}\left(\min_{z\in D(e^{ix_\alpha}, 2\varepsilon n^{-2})} |G_n(z)| < \varepsilon n^{-1/2}, \mathcal{G}\right) \le \mathcal{P}_{1,n} + \mathcal{P}_{2,n}, \tag{5.3}$$

where

$$\begin{aligned}\mathcal{P}_{1,n} \quad &:= \quad \mathbb{P}\left(|T_n(x_\alpha)| < 4\varepsilon n^{-1/2}\right), \\ \mathcal{P}_{2,n} \quad &:= \quad \mathbb{P}\left(|T(x_\alpha)| \le 4\varepsilon n^{-2} |T_n'(x_\alpha)|, ||T_n'||_\infty \le C_0 n^{3/2} \log^{1/2}(n)\right).\end{aligned}$$

We show that $\mathcal{P}_{1,n} + \mathcal{P}_{2,n} = \mathrm{O}(\varepsilon^2 n^{-2})$ as $n \to \infty$. Since the number of good intervals does not exceed $4\pi\varepsilon^{-1} n^2$, this will imply that the sum over α in (5.3) is $\mathrm{O}(\varepsilon)$ as $n \to \infty$. Also, we shall establish that the sum over the bad intervals in (5.3) is $\mathrm{o}(1)$ as $n \to \infty$. The proof will be complete provided $\lim_{n\to\infty} \mathbb{P}(\mathcal{G}^c) = 0$. Note that

$$\mathbb{P}(\mathcal{G}^c) \le \mathcal{P}_{3,n} + \mathcal{P}_{4,n},$$

where

$$\begin{aligned}\mathcal{P}_{3,n} \quad &:= \quad \mathbb{P}\left(||T_n'||_\infty > C_0 n^{3/2} \log^{1/2}(n)\right), \\ \mathcal{P}_{4,n} \quad &:= \quad \mathbb{P}\left(\sup_{z\in\mathcal{N}} |G_n''(z)| > n^{13/4}\right).\end{aligned}$$

$\mathcal{P}_{4,n} = \mathbf{o}(1)$ as $n \to \infty$

By the Markov inequality,

$$\begin{aligned}
\mathbb{P}\left(\sup_{z\in\mathcal{N}} |G_n''(z)| > n^{13/4}\right) &\leq \mathbb{P}\left(\sum_{j=0}^{n-1} j(j-1)|\xi_j||\phi(j/n)|\left(1+\frac{1}{n^2}\right)^{j-2} > n^{13/4}\right) \\
&\leq n^{-13/4}\sum_{j=0}^{n-1}\mathbb{E}\left(|\xi|\right) ej^2 \\
&\leq Cn^{-1/4}.
\end{aligned}$$

□

$\mathcal{P}_{0,n} = \mathbf{o}(1)$ as $n \to \infty$

The proof of this statement is the same as that provided by Lemma 3.3 in [42]. The auxiliary Lemma 3.2 in [42] should be taken with the mean covariance matrix of $T_n(x)$ multiplied by $\mathbb{E}\left(\xi^2\right)$. Then

$$\sup_{1\leq\beta\leq B} \mathbb{P}\left(\min_{z\in\mathcal{N}, z/|z|\in\mathcal{E}(J_\beta)} |G_n(z)| < \varepsilon n^{-1/2}, \mathcal{G}\right) = \mathrm{o}(1).$$

□

$\mathcal{P}_{3,n} = \mathbf{o}(1)$ as $n \to \infty$

In order to adapt the classical Salem–Zygmund Theorem to estimate $\mathcal{P}_{3,n}$, we show that if ξ is a random variable with moment generating function (mgf) and $\mathbb{E}\left(\xi\right) = 0$, then the mgf of ξ has a similar behavior around the origin to that of a sub-Gaussian random variable. Recall, a real-valued random variable ξ is said to be sub-Gaussian if there is some $b > 0$ such that for every $t \in \mathbb{R}$

$$\mathbb{E}\left(e^{t\xi}\right) \leq e^{b^2t^2/2}.$$

When this condition is satisfied for a particular value of $b > 0$, we say that ξ is b-sub-Gaussian, or sub-Gaussian with parameter b[1].

Lemma 19 ($\star$ Locally sub-Gaussian). *If ξ is a random variable with moment generating function M_ξ such that $\mathbb{E}\left(\xi\right) = 0$, $\mathbb{E}\left(\xi^2\right) = \sigma^2 > 0$, then there is a δ*

$$M_\xi(t) \leq e^{\gamma t^2/2} \quad \text{for} \quad |t| < \delta,$$

where $\gamma > \sigma^2$.

Proof. Define $g(t) := e^{\gamma t^2/2}$ for $t \in \mathbb{R}$. Then $g(0) = 1$, $g'(0) = 0$, $g''(0) = \gamma$. Let $h(t) := g(t) - M_\xi(t)$ for all $t \in I_\xi$, where I_ξ is the neighborhood of definition of M_ξ. Since $h''(0) = \gamma - \sigma^2 > 0$, there exists $\delta > 0$ such that $h''(t) > 0$ for every $|t| < \delta$. As $h'(0) = 0$, and therefore it is non-negative for $|t| < \delta$, it follows that $h(t) \geq 0$ for every $|t| < \delta$. □

Lemma 20 ($\star$ Salem–Zygmund). *Let $\{\xi_k\}_{k\geq 0}$ be a sequence of independent and identically distributed random variables with moment generating function M_{ξ_0} such that $\mathbb{E}\left(\xi_0\right) = 0$ and $\mathbb{E}\left(\xi_0^2\right) =$*

[1] For more details, see Appendix A.

$\sigma^2 > 0$. Let $W_n(x) := \sum_{j=0}^{n-1} \xi_j f_j(x)$ be a random trigonometric polynomial where $f_j(x) = \phi(j/n)e^{ijx}$, and ϕ is as in Theorem 15. Then for all large n,

$$\mathbb{P}\left(||W_n||_\infty \geq C_0 \left(r_n \log(n)\right)^{1/2}\right) \leq \frac{8\pi}{n^2}$$

for $r_n := \sum_{j=0}^{n-1} ||f_j||_\infty^2$ and some absolute constant $C_0 > 0$.

Proof. By Lemma 19, there exists a $\delta > 0$ such that

$$M_{\xi_0}(t) \leq e^{\gamma t^2/2} \quad \text{for} \quad |t| < \delta,$$

where $\gamma > \sigma^2$.

At first, we suppose that the f_j are real (we consider only the real part or the imaginary part) and we write $M_n := ||W_n||_\infty$. Since $\|f_j\|_\infty \leq 1$ for every $j = 0, \ldots, n-1$ then

$$e^{\gamma t^2 r_n/2} \geq \prod_{j=0}^{n-1} e^{\gamma t^2 |f_j(x)|^2/2} \geq \prod_{j=0}^{n-1} \mathbb{E}\left(e^{t\xi_j f_j(x)}\right) = \mathbb{E}\left(\prod_{j=0}^{n-1} e^{t\xi_j f_j(x)}\right) = \mathbb{E}\left(e^{tW_n(x)}\right)$$

for every $|t| < \delta$. There exists an interval I (in $\mathbb{T}$) of length $1/\rho_n$ with $\rho_n = 2\pi n^2$, where $|W_n(x)| > \frac{1}{2}||W_n||_\infty$ (see Proposition 5 of chapter 5 in [30]). So, $W_n(x) \geq M_n/2$ or $-W_n(x) \geq M_n/2$ on I. Then for every $|t| < \delta$,

$$\begin{aligned}
\mathbb{E}\left(e^{tM_n/2}\right) &\leq \rho_n \mathbb{E}\left(\int_I \left(e^{tW_n(x)} + e^{-tW_n(x)}\right)\mu(dx)\right) \\
&\leq \rho_n \mathbb{E}\left(\int_{\mathbb{T}} \left(e^{tW_n(x)} + e^{-tW_n(x)}\right)\mu(dx)\right) \\
&\leq 2\rho_n e^{\gamma t^2 r_n/2},
\end{aligned}$$

where μ is the normalized Lebesgue measure in $\mathbb{T}$.

From the above inequality, we obtain

$$\mathbb{E}\left(\exp\left\{\frac{t}{2}\left(M_n - \gamma t r_n - \frac{2}{t}\log\left(2\rho_n k\right)\right)\right\}\right) \leq \frac{1}{k}$$

for any $k > 0$ and $|t| < \delta$. Hence

$$\mathbb{P}\left(M_n \geq \gamma t r_n + \frac{2}{t}\log\left(2\rho_n k\right)\right) \leq \frac{1}{k}$$

for any $k > 0$ and $|t| < \delta$. For all large n, $\left|\frac{\log(2\rho_n k)}{\gamma r_n}\right| < \delta^2$. By choosing $t_n = \left(\frac{\log(2\rho_n k)}{\gamma r_n}\right)^{1/2}$ we obtain

$$\mathbb{P}\left(M_n \geq 3\left(\gamma r_n \log\left(2\rho_n k\right)\right)^{1/2}\right) \leq \frac{1}{k}$$

for any $k > 0$. Since $f_j = \mathrm{Re}(f_j) + i\mathrm{Im}(f_j)$, for all large n we have

$$\mathbb{P}\left(||\mathrm{Re}W_n||_\infty \geq 3\left(\gamma\sum_{j=0}^{n-1}||\mathrm{Re}f_j||_\infty^2\log(2\rho_n k)\right)^{1/2}\right) \leq \frac{1}{k},$$
$$\mathbb{P}\left(||\mathrm{Im}W_n||_\infty \geq 3\left(\gamma\sum_{j=0}^{n-1}||\mathrm{Im}f_j||_\infty^2\log(2\rho_n k)\right)^{1/2}\right) \leq \frac{1}{k}.$$

for any $k > 0$. Lastly, as $\rho_n = 2\pi n^2$ and taking $k = \frac{n^2}{4\pi}$, we have

$$\mathbb{P}\left(||W_n||_\infty \geq C_0\left(r_n\log(n)\right)^{1/2}\right) \leq \frac{8\pi}{n^2}$$

for large n, where $C_0 > 0$ is a suitable constant. □

As $f_j(x) = \phi(j/n)e^{ijx}$, $j = 0, \dots, n$, we have

$$r_n = \sum_{j=0}^{n-1}||f_j||_\infty^2 = \sum_{j=0}^{n-1}|\phi(j/n)|^2 \approx n\int_0^1\phi^2(x)dx$$

for all large n. Using Lemma 20 and the Bernstein inequality (page 153 in [55]), we have

$$\mathbb{P}\left(||T_n'||_\infty \geq C_0 n^{3/2}\log^{1/2}(n)\right) \leq \mathbb{P}\left(||T_n||_\infty \geq C_0 n^{1/2}\log^{1/2}(n)\right) \leq \frac{8\pi}{n^2},$$

where $C_0 > 0$ is an absolute constant.

$\mathcal{P}_{1,n} = \mathbf{O}(\varepsilon n^{-2})$ as $n \to \infty$

For a random variable ξ to have a moment generating function it is necessary and sufficient that it have exponential decay.

Lemma 21. *The following statements are equivalent.*

1. *There exist positive constants b and c such that*

$$\mathbb{P}\left(|\xi| \geq x\right) \leq be^{-cx} \quad \textit{for all } x > 0.$$

2. *There exists a constant $H > 0$ such that*

$$\mathbb{E}\left(e^{t\xi}\right) < \infty \quad \textit{for } |t| < H.$$

Proof. See Section 7 in [43]. □

The proof of $\mathcal{P}_{1,n} = \mathrm{O}(\varepsilon n^{-2})$ as $n \to \infty$ is the same as that provided by Lemma 4.1 in [42]. The auxiliary Lemma 4.2 should be considered that the characteristic function $f_\alpha(s)$ of $\frac{1}{\sqrt{n}}T(x_\alpha)$ is

$$f_\alpha(s) = \prod_{j=0}^{n-1}\mathbb{E}\left(\cos\left(\pi\xi_j\psi_j\right)\right),$$

where $s = (s_1, s_2) \in \mathbb{R}^2$ and

$$\psi_j = \frac{1}{\pi}\frac{1}{\sqrt{n}}\phi(j/n)\left(s_1\cos(jx_\alpha) + s_2\sin(jx_\alpha)\right),$$

this is possible because we want to estimate the value of $|f_\alpha|$, hence we can suppose that ξ_j is a symmetric random variable.

From Lemma 4.3 in [42], we have that there is $\mathcal{J} \subset \{0, 1, \ldots, n-1\}$ such that $\sup_{j\in\mathcal{J}}|\phi_j| \leq 3n^{-\tau}$ ($\tau := (\varsigma - 1/2)/10$ where ς is the constant in Theorem 15). Thus, for large n and by Lemma 21, we have

$$\begin{aligned}
|f_\alpha(z)| &\leq \prod_{j\in\mathcal{J}} \mathbb{E}\left(\left|\cos\left(\xi_j\pi\psi_j\right)\right|\right)\\
&\leq \prod_{j\in\mathcal{J}} \mathbb{E}\left(\left|\cos\left(\xi_j\pi\psi_j\right)\right| \mathbb{1}_{|\xi_j|\leq n^\tau/3} + \left|\cos\left(\xi_j\pi\psi_j\right)\right| \mathbb{1}_{|\xi_j|\geq n^\tau/3}\right)\\
&\leq \prod_{j\in\mathcal{J}} \mathbb{E}\left(\left(1-\xi_j^2\psi_j^2\right)\mathbb{1}_{|\xi_j|\leq n^\tau/3} + \mathbb{1}_{|\xi_j|\geq n^\tau/3}\right)\\
&= \prod_{j\in\mathcal{J}} \left(\left(1-\psi_j^2\mathbb{E}\left(\xi^2\mathbb{1}_{|\xi|\leq n^\tau/3}\right)\right) + \mathbb{P}\left(|\xi|\geq n^\tau/3\right)\right)\\
&\leq \prod_{j\in\mathcal{J}} \left(\left(1-\psi_j^2\frac{\mathbb{E}\left(\xi^2\right)}{2}\right) + be^{-cn^\tau}\right)\\
&\leq \prod_{j\in\mathcal{J}} \left(1+\frac{1}{n}\right)\left(1-\psi_j^2\frac{\mathbb{E}\left(\xi^2\right)}{2}\right)\\
&\leq e\prod_{j\in\mathcal{J}} \left(1-\psi_j^2\frac{\mathbb{E}\left(\xi^2\right)}{2}\right).
\end{aligned}$$

So, we obtain

$$\sup_{n^{1/6}<|s|<n^{1+\tau}} |f_\alpha(s)| < \exp\left\{-n^\tau\right\},$$

which is an important part of the proof of Lemma 4.1 in [42].

Lastly, we have

$$\sup_\alpha \mathbb{P}\left(|T(x_\alpha)| < \varepsilon n^{-1/2}\right) \leq C\varepsilon^2 n^{-2}.$$

□

$\mathcal{P}_{2,n} = \mathbf{O}(\varepsilon n^{-2})$ as $n \to \infty$

The proof of this statement is similar to that given in Lemma 5.2 in [42]. In the auxiliary Lemma 5.1, the covariance matrix of

$$\frac{1}{\sqrt{n}}(T(x_\alpha), T_n'(x_\alpha)/in)$$

should be multiplied by $\mathbb{E}\left(\xi^2\right)$. Using the ideas given in subsection 5.4 of the present thesis, we have

$$\sup_\alpha \mathbb{P}\left(|T(x_\alpha)| \leq 4\varepsilon n^{-2}|T_n'(x_\alpha)|, ||T_n'||_\infty \leq C_0 n^{3/2}\log^{1/2}(n)\right) \leq C\varepsilon^2 n^{-2}.$$

□

5.5 Singularity of a circulant random matrix with prime dimension

Proof of Theorem 16. First, we determine the cardinality of $\left\{\cos\left(2\pi k \frac{j}{n}\right) \;:\; j=0,\ldots,n-1\right\}$.

Lemma 22 ($\star$)**.** *If n is prime, then*

$$\left|\left\{\cos\left(2\pi k \frac{j}{n}\right) \;:\; j=0,\ldots,n-1\right\}\right| \geq \left\lfloor \frac{n}{2} \right\rfloor - 1,$$

for every $k = 1, \ldots, n-1$.

Proof. Fix $k \in \{1,\ldots,n-1\}$. We consider the function $\cos(2\pi k x)$ defined on $[0,1]$. Let $y \in [0,1]$. Then

$$A := \{y, 1-y\} \cup \left\{\frac{m}{k} \pm y\right\}_{m=1}^{k-1}$$

is the set all possible values in $[0,1]$ that are equal to $\cos(2\pi k y)$.

For $j \in \{0,1,\ldots,n-1\}$, we take $y = j/n$. We need to check that there is a $j' \in \{0,1,\ldots,n-1\}$ such that

$$\frac{j}{n} \mp \frac{j'}{n} = \frac{m}{k}$$

for $m \in \{1,\ldots,k-1\}$. If we suppose that this happens, we have

$$\frac{k}{m}(j \mp j') = n.$$

If m divides k, we have a contradiction. Hence, $j \mp j' = \alpha m$ for some integer α, so $k\alpha = n$. But, since $k \geq 1$ implies that $\alpha < n$, we have again a contradiction.

Note $\cos(2\pi k x) = 0$ only at $x = m/k$ for $1 \leq m \leq k-1$. So $j/n = m/k$, or

$$n = \frac{k}{m} j$$

and using an argument similar to those before, we have $\cos(2\pi k j/n) \neq 0$. Therefore, we obtain the result. □

The next lemma shows how a sum of Rademacher random variables can take particular values [51].

Lemma 23. *Let $\{a_j\}_{j=1}^n$ be a set of distinct real numbers different from zero. If $\{\xi_k\}_{k=1}^n$ are independent and identically distributed Rademacher random variables, then*

$$\sup_{x \in \mathbb{R}} P\left(\sum_{k=1}^{n} a_k \xi_k = x\right) = \mathrm{O}(n^{-3/2}).$$

□

Now, we can give an upper bound for the probability that a circulant matrix is singular under the hypothesis its dimension is a prime number.

From (5.1), we have that the event $\{\text{circ}(\xi_0 \dots, \xi_{n-1})$ is singular$\}$ is equivalent to

$$\left\{ \prod_{k=0}^{n-1} |G_n(\omega_n^k)| = 0 \right\}.$$

Note $|G_n(\omega_n^k)| = 0$ implies that $\sum_{k=0}^{n-1} \xi_k \cos(2\pi kj/n) = 0$. Hence

$$\mathbb{P}(\text{circ}(\xi_0 \dots, \xi_{n-1}) \text{ is singular}) \le \mathbb{P}\left(\sum_{j=0}^{n-1} \xi_j = 0\right) + \sum_{k=1}^{n-1} \mathbb{P}\left(\sum_{k=0}^{n-1} \xi_k \cos(2\pi kj/n) = 0\right).$$

From Lemma 22 and Lemma 23, we have for $k \neq 1$

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=0}^{n-1} \xi_k \cos\left(2\pi k\frac{j}{n}\right) = 0\right) &= \mathbb{P}\left(\sum_{k=0}^{n-1} X_k \cos\left(2\pi k\frac{j}{n}\right) = S\right) \\
&= \mathbb{E}\left[\mathbb{P}\left(\sum_{k=0}^{n-1} X_k \cos\left(2\pi k\frac{j}{n}\right) = S \,\middle|\, Y_0, \dots, Y_{n-1}\right)\right] \\
&\le \mathbb{E}\left[\mathrm{O}(n^{-3/2})\right] \\
&= \mathrm{O}(n^{-3/2}),
\end{aligned}$$

where $S := -\sum_{k=0}^{n-1} \tau_k \cos(2\pi kj/n)$. By the properties of the Lévy concentration function (pages 22 and 68 in [54]), we have

$$\mathbb{P}\left(\sum_{j=0}^{n-1} \xi_j = 0\right) = \mathrm{O}(n^{-1/2}).$$

Therefore,

$$\mathbb{P}(\text{circ}(\xi_0 \dots, \xi_{n-1}) \text{ is singular}) = \mathrm{O}(n^{-1/2}).$$

□

5.6 Some additional results on circulant random matrices

A g-circulant matrix $\mathcal{C}_n^g$ is an $n \times n$ matrix with the following form

$$\mathcal{C}_n^g := \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-g} & c_{n-g+1} & \cdots & c_{n-g-1} \\ c_{n-2g} & c_{n-2g+1} & \cdots & c_{n-2g-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_g & c_{g+1} & \cdots & c_{g-1} \end{bmatrix},$$

where g is a positive integer and each of the subscripts is understood to be reduced modulo n. The first row of $\mathcal{C}_n^g$ is $(c_0, c_1, \dots, c_{n-1})$ and its $(j+1)$th row is obtained by giving the jth row a right circular shift by g positions. Note that $g = 1$ or $g = n+1$ yields the classical circulant matrix.

A g-circulant matrix $\mathcal{C}_n^g$ with first row $(c_0, c_1, \dots, c_{n-1})$ can be factored as $\mathcal{C}_n^g = \mathcal{Q}_n^g \mathcal{C}_n$, where $\mathcal{Q}_n^g$ is a g-circulant matrix with the first row $e^* = (1, 0, \dots, 0)$ and $\mathcal{C}_n$ is a circulant matrix whose first row is $(c_0, c_1, \ldots, c_{n-1})$. The matrix $\mathcal{Q}_n^g$ is an unitary matrix if and only if n and g are co-prime integers [76]. From Theorem 14, we get the following corollary.

Corollary 1 ($\star$). *Let $\{\xi_k\}_{k\geq 0}$ be a sequence of random variables as in Theorem 14. Let $\mathcal{C}_n^g$ be a g–circulant random matrix, whose first row is $(\xi_0, \ldots, \xi_{n-1})$. Then for all $\varepsilon > 0$ and for all large n such that n and g are co-prime, we have*

$$\mathbb{P}(s_n(\mathcal{C}_n^g) \geq \varepsilon n^{-1/2}) \geq 1 - C\varepsilon,$$

where C is a constant depending on the distribution of ξ_0.

□

Now, we present some results about the maximum singular value of a circulant random matrix when its entries are complex random variables. We can establish exact distributions for $s_1(\mathcal{C}_n)$ and $s_n(\mathcal{C}_n)$. Write $X_n^t = (\xi_0, \ldots, \xi_{n-1})$. Then $\sqrt{n}F_nX_n$ is the vector of eigenvalues of $\mathcal{C}_n$, where F_n is the Fourier matrix of order n. Since F_n is a unitary matrix, if we suppose that X_n is a complex random vector such that $(\mathrm{Re}(X), \mathrm{Im}(X)) \in \mathbb{R}^{2n}$ has a spherical distribution (chapter 2 in [19]), and we establish the distribution of $s_1(\mathcal{C}_n)$ and $s_n(\mathcal{C}_n)$.

Theorem 17 ($\star$). *Let $X_n^t = (\xi_0, \ldots, \xi_{n-1}) \in \mathbb{C}^n$ be a complex random vector such that $(X_n')^t := (\mathrm{Re}(X_n), \mathrm{Im}(X_n)) \in \mathbb{R}^{2n}$ has a spherical distribution. Then*

$$F_nX_n \overset{D}{=} X_n,$$

where F_n is the Fourier matrix of order n. Moreover, if all entries of X_n' are independent random variables, we have for all $x \in \mathbb{R}$

$$\mathbb{P}\left(s_1(\mathcal{C}_n) \leq x\right) = \prod_{j=0}^{n-1} \mathbb{P}\left(|\xi_j| \leq x/\sqrt{n}\right),$$

$$\mathbb{P}\left(s_n(\mathcal{C}_n) \geq x\right) = \prod_{j=0}^{n-1} \mathbb{P}\left(|\xi_j| \geq x/\sqrt{n}\right).$$

Proof. Note F_n is unitary if and only if

$$F_n' := \begin{bmatrix} \mathrm{Re}(F_n) & -\mathrm{Im}(F_n) \\ \mathrm{Im}(F_n) & \mathrm{Re}(F_n) \end{bmatrix}$$

is orthogonal. Hence, $W_n := F_nX_n$ if and only if $W_n' = F_n'X_n'$, where

$$X_n' := \begin{bmatrix} \mathrm{Re}(X_n) \\ \mathrm{Im}(X_n) \end{bmatrix}.$$

Since the distribution of X_n' is spherical and F_n' is orthogonal, we have $W_n' \overset{D}{=} X_n'$, which implies $W_n \overset{D}{=} X_n$. Since $\sqrt{n}F_nX_n \overset{D}{=} \sqrt{n}X_n$, then for the second part of the statement, we only need to take the modulus of every entry of $\sqrt{n}X_n$. □

Theorem 17 is similar to Theorem 7.1 in [53] and Proposition 3 in [49]. However, Theorem 7.1 in [53] assumed F_n to be an orthogonal matrix in order to deduce that $F_nX_n \overset{D}{=} X_n$ when X_n

is a normally distributed random vector in $\mathbb{R}^n$, which is not possible since $F_n X_n$ is, in general, a complex vector and X_n is a real vector[2].

In the following theorems, we study the behavior of $s_1(\mathcal{C}_n)$ in the case that the entries of $\mathcal{C}_n$ are identically distributed sub-Gaussian random variables.

Theorem 18 ($\star$). *Let $\{\xi_j\}_{j\geq 0}$ be identically distributed sub-Gaussian random variables with parameter b. Then for every $\varepsilon > 0$, $\varrho > 0$, and all n large,*

$$\mathbb{P}\left(s_1(\mathcal{C}_n) > \varepsilon n^{1+\varrho}\right) \leq \exp\left(-\frac{\varepsilon^2 n^{\varrho}}{4b^2}\right).$$

Proof. Note

$$\mathbb{P}\left(\max_{0\leq k\leq n-1}\left|\sum_{j=0}^{n-1}\xi_j \exp\left(ij\frac{2\pi k}{n}\right)\right| > \varepsilon n^{1+\varrho}\right) \leq \sum_{k=0}^{n-1}\mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \exp\left(ij\frac{2\pi k}{n}\right)\right| > \varepsilon n^{1+\varrho}\right)$$

and

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \exp\left(ij\frac{2\pi k}{n}\right)\right| > \varepsilon n^{1+\varrho}\right) &\leq \mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \cos\left(j\frac{2\pi k}{n}\right)\right| > \frac{\varepsilon}{2}n^{1+\varrho}\right)\\ &\quad+\mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \sin\left(j\frac{2\pi k}{n}\right)\right| > \frac{\varepsilon}{2}n^{1+\varrho}\right).\end{aligned}$$

Also, we have that $\sum_{j=0}^{n-1}\xi_j\cos\left(j\frac{2\pi k}{n}\right)$ and $\sum_{j=0}^{n-1}\xi_j\sin\left(j\frac{2\pi k}{n}\right)$ are sub-Gaussian random variables with parameter

$$b_{k,c} := b\sum_{j=0}^{n-1}\left|\cos\left(j\frac{2\pi k}{n}\right)\right| \quad \text{and} \quad b_{k,s} := b\sum_{j=0}^{n-1}\left|\sin\left(j\frac{2\pi k}{n}\right)\right|,$$

respectively. By sub-Gaussian tail estimation, we have that

$$\mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \cos\left(j\frac{2\pi k}{n}\right)\right| > \frac{\varepsilon}{2}n^{1+\varrho}\right) \leq 2\exp\left(-\frac{\varepsilon^2 n^{2(1+\varrho)}}{8b_{k,c}^2}\right),$$

$$\mathbb{P}\left(\left|\sum_{j=0}^{n-1}\xi_j \sin\left(j\frac{2\pi k}{n}\right)\right| > \frac{\varepsilon}{2}n^{1+\varrho}\right) \leq 2\exp\left(-\frac{\varepsilon^2 n^{2(1+\varrho)}}{8b_{k,s}^2}\right).$$

If $k=0$, then $b_{0,c}^2 = b^2n^2$ and $b_{0,s}^2 = 0$. If $k\neq 0$, we have that

$$b_{k,c}^2 \leq \frac{b^2n^2}{2} \quad \text{and} \quad b_{k,s}^2 \leq \frac{b^2n^2}{2}.$$

Therefore, for all large n,

$$\mathbb{P}\left(s_1(\mathcal{C}_n) > \varepsilon n^{1+\varrho}\right) \leq 4n\exp\left(-\frac{\varepsilon^2 n^{2\varrho}}{4b^2}\right) \leq \exp\left(-\frac{\varepsilon^2 n^{\varrho}}{4b^2}\right).$$

[2]The author of this thesis asked to the authors of [53] about this mistake, but as of now, I have not received a response.

□

In the next lemma, we determine the behavior of the moment generating functions of $|\xi|$ and ξ^2 when ξ is a sub-Gaussian random variable. This lemma will be useful for estimating the expectation of $s_1(\mathcal{C}_n)$.

Lemma 24. *Let ξ be a sub-Gaussian random variable with parameter $b > 0$. Then there is a $B > 0$ such that for $t \in \mathbb{R}$,*

$$\mathbb{E}\left(\exp\left(t\xi^2\right)\right) \leq \exp\left\{\frac{B}{2}^2 t^2 + t\mathbb{E}\left(\xi^2\right)\right\},$$

$$\mathbb{E}\left(\exp\left(t|\xi|\right)\right) \leq \exp\left\{\frac{b}{2}^2 t^2 + t\mathbb{E}\left(|\xi|\right)\right\}.$$

For $t \geq 0$,

$$\mathbb{E}\left(\exp\left(t\xi^2\right)\right) \leq \exp\left\{\frac{B}{2}^2 t^2 + 2C^2 b^2 t\right\},$$

$$\mathbb{E}\left(\exp\left(t|\xi|\right)\right) \leq \exp\left\{\frac{b}{2}^2 t^2 + Cbt\right\},$$

where C is an absolute constant that does not depend on ξ.

Proof. This follows from the definition of sub-Gaussian random variable. □

Theorem 19 ($\star$)**.** *Let $\{\xi_j\}_{j\geq 0}$ be identically distributed sub-Gaussian random variables with parameter b. Then*

$$\mathbb{E}\left(s_1(\mathcal{C}_n)\right) \leq nb\left(\sqrt{n\log n} + 2\sqrt{\log n} + 2C\right),$$

where C is an absolute constant that does not depend on X.

Proof. Let $\Re_k := \mathrm{Re}(G_n(\omega_n^k))$ and $\Im_k := \mathrm{Im}(G_n(\omega_n^k))$. From the proof of Theorem 18, we have that $\Re_k$ and $\Im_k$ are sub-Gaussian random variables such that their parameters satisfy

$$b_{k,c} \leq bn \quad \text{and} \quad b_{k,s} \leq bn \quad \text{for all } k.$$

If $Z_k := |\Re_k| + |\Im_k|$, then by Lemma 24 and the Cauchy–Schwartz inequality, for every $\lambda > 0$

$$\begin{aligned}\mathbb{E}\left(\exp\{\lambda Z_k\}\right) &\leq \sqrt{\mathbb{E}\left(\exp\left(2\lambda|\Re_k|\right)\right)\mathbb{E}\left(\exp\left(2\lambda|\Im_k|\right)\right)} \\ &\leq \exp\left\{2n^2b^2\lambda^2 + 2Cnb\lambda\right\}.\end{aligned}$$

Then, by the Jensen inequality, for every $\lambda > 0$,

$$\begin{aligned}\exp\left(\lambda\mathbb{E}\left(\max_{k=1,\ldots,n} Z_k\right)\right) &\leq \mathbb{E}\left(\exp\left(\lambda\max_{k=1,\ldots,n} Z_k\right)\right) = \mathbb{E}\left(\max_{k=1,\ldots,n} e^{\lambda Z_k}\right) \\ &\leq \sum_{k=1}^{n}\mathbb{E}\left(e^{\lambda Z_k}\right) \leq n\exp\left\{2n^2b^2\lambda^2 + 2Cnb\lambda\right\}.\end{aligned}$$

Taking logarithms of both sides and dividing by λ in the previous expression, we have

$$\mathbb{E}\left(\max_{k=1,\ldots,n} Z_k\right) \leq \frac{\log n}{\lambda} + 2n^2b^2\lambda + 2Cnb.$$

The upper bound is minimized for $\lambda = \frac{\sqrt{\log n}}{\sqrt{2nb}}$, which yields

$$\mathbb{E}\left(\max_{k=1,\ldots,n} Z_k\right) \leq nb\left(\sqrt{n\log n} + 2\sqrt{\log n} + 2C\right).$$

Lastly,

$$\begin{aligned}
\mathbb{E}\left(s_1(\mathcal{C}_n)\right) &= \mathbb{E}\left(\max_{0\leq k\leq n-1} |G_n(\omega_n^k)|\right) \\
&\leq \mathbb{E}\left(\max_{0\leq k\leq n-1}\left(\left|\mathrm{Re}(G_n(\omega_n^k))\right| + \left|\mathrm{Im}(G_n(\omega_n^k))\right|\right)\right) \\
&\leq nb\left(\sqrt{n\log n} + 2\sqrt{\log n} + 2C\right).
\end{aligned}$$

□

One question arises from the previous discussion. When are the roots of a random polynomial G_n not near the unit circle? This problem was studied in [63], which provides a useful statement for understanding the minimum singular value of a random circular matrix.

We consider a sequence of random polynomials $\left\{H_n(z) := \sum_{k=0}^n \xi_k z^k\right\}_{n\geq 1}$ whose coefficients form a sequence of independent real- or complex-valued random variables. The complex-valued coefficients are of the form $\xi_k = X_k + iY_k$ where X_k and Y_k are real-valued random variables. The means μ_k and variances σ_k^2 of ξ are given by

$$\mu_k = \mu_{X_k} + i\mu_{Y_k}$$

and

$$\sigma_k^2 = \sigma_{X_k}^2 + \sigma_{Y_k}^2.$$

Let δ, η, θ be arbitrary numbers such that $0 \leq \eta < \theta \leq 2\pi$ and $0 \leq \delta \leq 1$. Let $\mathcal{D}$ and $\mathcal{R}$ be the following subsets of the complex plane:

$$\mathcal{D} := \{z \in \mathbb{C} : \eta < \arg(z) < \theta\},$$

$$\mathcal{R} := \{z \in \mathbb{C} : 1-\delta \leq |z| \leq 1+\delta\}.$$

Define $N_n(\mathcal{D})$ and $N_n(\mathcal{R})$ to be the number of zeros of H_n contained in $\mathcal{D}$ and $\mathcal{R}$, respectively.

In [63], we find the next result.

Lemma 25. *Let $\{\xi_n\}_{n\geq 0}$ be a sequence of independent random variables with finite means and standard deviations such that*

$$\limsup_{k\to\infty} \sqrt[k]{|\mu_k|} > \limsup_{k\to\infty} \sqrt[k]{\sigma_k}.$$

Then the following equalities hold almost surely:

$$\lim_{n_j\to\infty} \frac{N_{n_j}(\mathcal{R}^*)}{n_j} = 1,$$

$$\lim_{n_j\to\infty} \frac{N_{n_j}(\mathcal{D})}{n_j} = \frac{\theta-\eta}{2\pi},$$

where

$$\mathcal{R}^* = \left\{ z : \left(\limsup_{k\to\infty} \sqrt[k]{|\mu_k|} \right)^{-1} - \delta < |z| < \left(\limsup_{k\to\infty} \sqrt[k]{|\mu_k|} \right)^{-1} + \delta \right\}$$

and H_{n_j} is a subsequence of H_n which consists of polynomials whose coefficients satisfy

$$\lim_{n_j\to\infty} \sqrt[n_j]{|\mu_{n_j}|} = \limsup_{k\to\infty} \sqrt[k]{|\mu_k|}.$$

□

From Lemma 25 and using the Fundamental Theorem of Algebra, we can construct random circulant matrices such that their minimum singular value is large, as the following theorem establishes.

Theorem 20 ($\star$)**.** *Let $\{\xi_n\}_{n\geq 0}$ be a sequence of independent (real- or complex-) random variables with finite means and standard deviations such that*

$$R := \lim_{k\to\infty} \sqrt[k]{|\mu_k|} > \limsup_{k\to\infty} \sqrt[k]{\sigma_k}.$$

Then for all $0 < \delta < 1$, we have with high probability

1. *If $0 < R < 1$,*
$$s_n(\mathcal{C}_n) \geq (R^{-1} - \delta - 1)^n.$$
2. *If $R > 1$,*
$$s_n(\mathcal{C}_n) \geq (1 - \delta - R^{-1})^n.$$

□

Chapter 6

Oriented Percolation with Backbend

6.1 Introduction

Oriented percolation with backbend is a generalization of oriented percolation, defined by Durrent in [18]. In this chapter, we analyze the properties of the backbend model and its similarities and diferences with unoriented percolation in two dimensions.

Section 6.2 describes the model of backbend percolation; which roughly speaking is similar to oriented percolation with the diference that the backbend path is allowed to go down until a depth b. Section 6.3 gives a characterization of the critical probabability p_c^b of backbend percolation in terms of the right edge process. Section 6.4 gives the proof that there exists an initial distribution on the infinite subsets of $\{\dots,-4,-2,0\}$ which contain 0, such that the right edge process has stationary increments. Section 6.5 shows the strict monotonicity of the "slope" of right edge process respect to depth of backbend percolation. Section 6.6 exhibits that in the sub–critical probability of backbend model, the probability that a backbend path reaches the level n descreases exponentially fast to zero. Section 6.7 shows some similarities of backbend model with the unoriented percolation in $\mathbb{Z}^2$, when the depth of backbend is going to infinity. In our first approach of this situation, we show that it is possible to construct a backben path such that it can go far away to the right side of zero. Section 6.8 studies the super–critical probability of backbend percolation with a renormalization argument. Also, we obtain that the "slope" of right edge process is zero, when the model takes the critical probability.

The main results in this chapter were obtained in joint work with Roy.

6.2 The model

We consider an undirected graph where $\mathcal{L} = \{(m,n) \in \mathbb{Z}^2 : m+n \text{ is even}, n \geq 0\}$ is the set of vertices and

$$\mathcal{E} = \{\langle (m,n),(m+1,n+1)\rangle, \langle (m,n),(m-1,n+1)\rangle; (m,n) \in \mathcal{L}\}$$

is the set of undirected edges. A path π in $\mathcal{L}$ is a sequence of finite or infinite distinct vertices $x_0, x_1, ..., (x_k) \in \mathcal{L}$ such that $\langle x_i, x_{i+1}\rangle \in \mathcal{E}$ for all $i = 0,1,\dots,(k)$.

The edges are open or closed, independently, with probability p or $1-p$. So, we have the space $(\{0,1\}^{\mathcal{E}}, \mathcal{B}, \mathbb{P}_p)$ where $\mathcal{B}$ is the σ–algebra generated by cylinder sets and $\mathbb{P}_p$ is the product measure with marginals $\mathbb{P}_p(\omega(e)=1) = p = 1-\mathbb{P}_p(\omega(e)=0)$ for all $e \in \mathcal{E}$.

For $0 \leq b < \infty$, a b–backbend path π^b in $\mathcal{L}$ is a finite or infinite path $x_0, x_1, ...$ such that $(x_j)_2 \geq (x_i)_2 - b$ for all $0 \leq i \leq j$, where $(z)_2$ denotes the second coordinate of the point $z \in \mathcal{L}$.

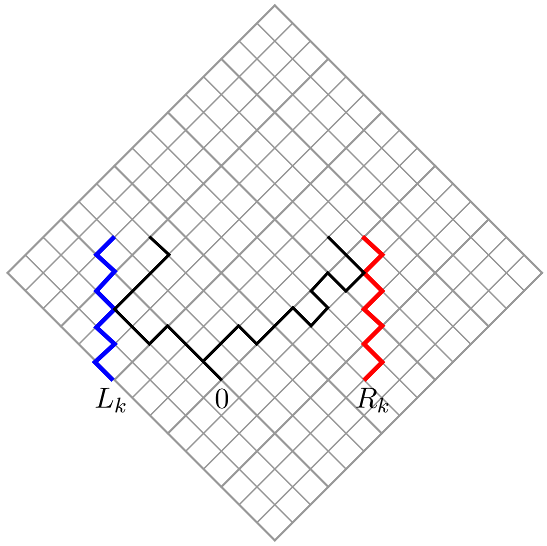

Figure 6.1: L_k and R_k

We call a path $\pi = x_0, x_1, ...$ open if all the edges $\langle x_i, x_{i+1}\rangle$ comprising the path are open. For $x, y \in \mathcal{L}$ we use the notation $x \to y$ to mean that the point x is connected to y through some open b–backbend path π^b.

We denote by C^b the random set of vertices $x \in \mathcal{L}$ that are connected to the origin through some open b–backbend path, i.e., $C^b = \{x \in \mathcal{L} : (0,0) \to x\}$. For $0 \le b < \infty$, the critical probability p_c^b is defined as $p_c^b = \sup\{p : \mathbb{P}_p(|C^b| = \infty) = 0\}$.

We define the following random variables:

- $u_b := \sup\{x : (y, 0) \to (x, b) \text{ for some } y \le 0\}$,
- $\bar{r}^b_{0,0} := 0$ and $\bar{r}^b_{0,n} := \sup\{x - u_b : (y, b) \to (x, n + b) \text{ for some } y \le u_b\}$ for $n > 0$,
- $\bar{r}^b_{m,n} := \sup\{x - \bar{r}^b_{0,m} : (y, m + b) \to (x, n + b) \text{ for some } y \le \bar{r}^b_{0,m}\}$ for $1 \le m \le n$.

The above definitions are meaningful only if $u_b < \infty$ a.s.; the following proposition shows this.

Proposition 4 (*). *For all $b \ge 0$, $u_b < \infty$ a.s.*

Proof. For a non-negative, even integer k let us define the path L_k by

$$L_k = (-k, 0), (-k - 1, 1), (-k, 2), (-k - 1, 3), \ldots, (-k, 2b)$$

and the path R_k by

$$R_k = (k, 0), (k + 1, 1), (k, 2), (k + 1, 3), \ldots, (k, 2b),$$

see Figure 6.1. We further define the random variables K_L, K_R by

$$K_L = \max\{k : L_k \text{ is an open path}\}$$

and

$$K_R = \min\{k : R_k \text{ is a closed path}\}.$$

The probability that the path L_k is an open path is $\gamma_l = p^{2b}$, and that the path R_k is a closed path is $\gamma_r = (1 - p)^{2b}$. By definition of K_L and K_R we have

$$|u_b| \le K_R - K_L.$$

Since K_R and K_L are geometric random variables with $\mathbb{E}(K_R) = 2(1-\gamma_r)/\gamma_r$ and $\mathbb{E}(K_L) = -2(1-\gamma_l)/\gamma_l$, we have $u_b < \infty$ a.s. $\square$

6.3 A characterization of the critical probability p_c^b

We prove the following properties of the process $\{\bar{r}^b_{m,n} : 0 \le m \le n\}_{n\ge 0}$.

Claim 1 (*). *For the process* $\{\bar{r}^b_{m,n} : 0 \le m \le n\}_{n\ge 0}$

1. $\bar{r}^b_{0,n} \le \bar{r}^b_{0,m} + \bar{r}^b_{m,n}$ *for all* $0 \le m \le n$.
2. $\{\bar{r}^b_{(n-1)k,nk} : n \ge 1\}$ *is stationary for all* k.
3. *The distribution of* $\{\bar{r}^b_{m,m+k} : k \ge 0\}$ *does not depend on* m.
4. $\mathbb{E}(|\bar{r}^b_{0,1}|) < \infty$.
5. *The process is ergodic.*

Proof.

1. By definition of u_b we have $\bar{r}^b_{0,0} = 0$. Since any b-path that starts from the line $y = b$ and reaches the line $y = n + b$ must cross the line $y = m + b$, we have $\bar{r}^b_{0,n} - \bar{r}^b_{0,m} \le \bar{r}^b_{m,n}$, or
$$\bar{r}^b_{0,n} \le \bar{r}^b_{0,m} + \bar{r}^b_{m,n}.$$
2. Let $E_{c,d}$ be the set of edges that lie between the levels c and d, i.e., $E_{c,d} := \{\langle x, y\rangle \in \mathcal{E} : c \le (x)_2, (y)_2 \le d\}$. Further, let $k \in \mathbb{N}$ be fixed and $x_1, x_2, \ldots, x_t \in \mathbb{R}$. Note that the probability of the event $\{\bar{r}^b_{0,k} < x_1, \bar{r}^b_{k,2k} < x_2 \ldots, \bar{r}^b_{(t-1)k,tk} < x_t\}$ depends only on what happens in $E_{0,tk+b}$. Since
$$E_{0,tk+b} \overset{\mathcal{D}}{=} E_{l,tk+b+l}$$
for all $l \in \mathbb{N}$, we have
$$\mathbb{P}_p\{\bar{r}^b_{0,k} < x_1, \ldots, \bar{r}^b_{(t-1)k,tk} < x_t\} = \mathbb{P}_p\{\bar{r}^b_{lk,(l+1)k} < x_1, \ldots, \bar{r}^b_{(l+t-1)k,(l+t)k} < x_t\}.$$
3. Let $0 \le l_1 < l_2 < \ldots < l_t$ be t integers and $x_1, \ldots, x_t \in \mathbb{R}$. As before, the probability of the event
$$\{\bar{r}^b_{m,m+l_1} < x_1, \ldots, \bar{r}^b_{m,m+l_t} < x_t\}$$
depends only on the configuration of edges in $E_{m-b,m+l_t+b}$. Since
$$E_{m-b,m+l_t+b} \overset{\mathcal{D}}{=} E_{m+1-b,m+1+l_t+b},$$
we have
$$\mathbb{P}_p\{\bar{r}^b_{m,m+l_1} < x_1, \ldots, \bar{r}^b_{m,m+l_t} < x_t\} = \mathbb{P}_p\{\bar{r}^b_{m+1,m+1+l_1} < x_1, \ldots, \bar{r}^b_{m+1,m+1+l_t} < x_t\}.$$
4. Let K_L and K_R be as in the proof of Proposition 4. For any non-negative, even integer k, let us define the path L_k by
$$L_k = (-k, 0), (-k-1, 1), (-k, 2), (-k-1, 3), \ldots, (-k-1, 2b+1)$$

and the path R_k as

$$R_k = (k,0),(k+1,1),(k,2),(k+1,3),\ldots,(k+1,2b+1).$$

We further define the random variables K_L, K_R by

$$K_L = \max\{k : L_k \text{ is an open path}\}$$

and

$$K_R = \min\{k : L_k \text{ is a closed path}\}.$$

The probability that the path L_k is an open path is $\gamma_l = p^{2b+1}$ and that the path R_k is a closed path is $\gamma_r = (1-p)^{2b+1}$. By definition of K_L and K_R we have

$$|\bar{r}^b_{0,1}| \le K_R - K_L.$$

Since K_R and K_L are geometric random variable with $\mathbb{E}(K_R) = 2(1-\gamma_r)/\gamma_r$ and $\mathbb{E}(K_L) = -2(1-\gamma_l)/\gamma_l$, we have

$$\mathbb{E}(|\bar{r}^b_{0,1}|) \le \mathbb{E}(K_R) - \mathbb{E}(K_L) = \frac{2(1-\gamma_r)}{\gamma_r} + \frac{2(1-\gamma_l)}{\gamma_l} < \infty$$

for $p \in (0,1)$.

5. Let us consider the process $\{\bar{r}^b_{kn,k(n+1)} : n \ge 1\}$ for fixed k. Let A be an invariant set, i.e., there is a $B \in \mathcal{B}$ such that for every $m \ge 1$

$$A = \{(\bar{r}^b_{km,k(m+1)}, \bar{r}^b_{k(m+1),k(m+2)}, \ldots) \in B\}.$$

Clearly $A \in \sigma(E_{km-b,\infty})$ for all $m \ge 1$, where $\sigma(E_{km-b,\infty})$ is the σ-algebra generated by the edges in the set $E_{km-b,\infty}$. Since $A \in \cap_{m=1}^{\infty} \sigma(E_{km-b,\infty})$, by Kolmogorov's 0–1 law, the probability $\mathbb{P}_p(A)$ is zero or one, implying that the process $\{\bar{r}^b_{kn,k(n+1)} : n \ge 1\}$ is ergodic.

□

Theorem 21 (*). *Consider the process* $\{\bar{r}^b_{m,n} : 0 \le m \le n\}_{n \ge 0}$. *Then*

$$\lim_{n\to\infty} \frac{\mathbb{E}(\bar{r}^b_{0,n})}{n} = \inf_{n \ge 1} \frac{\mathbb{E}(\bar{r}^b_{0,n})}{n} = \alpha^b(p)$$

for some constant $\alpha^b(p) \in [-\infty, \infty)$, *and*

$$\lim_{n\to\infty} \frac{\bar{r}^b_{0,n}}{n} = \alpha^b(p) \quad a.s.$$

Proof. From Claim 1, we have by the subadditive limit theorem that

$$\lim_{n\to\infty} \frac{\mathbb{E}(\bar{r}^b_{0,n})}{n} = \inf_{n \ge 1} \frac{\mathbb{E}(\bar{r}^b_{0,n})}{n} = \alpha^b(p)$$

for some constant $\alpha^b(p) \in [-\infty, \infty)$, and by the subadditive ergodic theorem

$$\lim_{n\to\infty} \frac{\bar{r}^b_{0,n}}{n} = \bar{r} \quad \text{a.s.}$$

for some random variable $-\infty \le \bar{r} < \infty$ such that $\mathbb{E}(\bar{r}) = \alpha^b(p)$. Moreover, if $\alpha^b(p) > -\infty$, then

$$\lim_{n\to\infty} \mathbb{E}\left(\left|\frac{\bar{r}^b_{0,n}}{n} - \bar{r}\right|\right) = 0.$$

From claim 5, we have $\bar{r} = \alpha^b(p)$ a.s. $\square$

We argue in the following that if $\mathbb{P}_p(|C^b| = \infty) > 0$, then $\alpha^b(p) > 0$. To show this, we define the following random variables:

- $w_b = \inf\{x : (y, 0) \to (x, b) \text{ for some } y \ge 0\}$,
- $\bar{l}^b_{0,0} = 0$ and $\bar{l}^b_{0,n} := \inf\{x - w_b : (y, b) \to (x, n + b) \text{ for some } y \ge w_b\}$ for $n \ge 1$,
- $\bar{l}^b_{m,n} = \inf\{x - \bar{l}^b_{0,m} : (y, m + b) \to (x, n + b) \text{ for some } y \ge \bar{l}^b_{0,m}\}$ for $1 \le m \le n$.

As in the proof of Proposition 4, we may show that $w_b > -\infty$ a.s. and by symmetry with $\bar{r}^b_{0,n}$ we have $\bar{l}^b_{0,n}/n \to -\alpha^b(p)$ as $n \to \infty$ a.s. Since $\bar{l}^b_{0,n} \le \bar{r}^b_{0,n}$, we have $-\alpha^b(p) \le \alpha^b(p)$, implying $\alpha^b(p) \ge 0$.

From our last statement, it follows that if $\alpha^b(p) < 0$, then $\mathbb{P}_p(|C^b| = \infty) = 0$. We show that if $\alpha^b(p) > 0$, then $\mathbb{P}_p(|C^b| = \infty) > 0$.

Claim 2 (*). *If $\alpha^b(p) > 0$, then $\mathbb{P}_p(|C^b| = \infty) > 0$.*

Proof. The idea of this proof is in principle the same as in the case of no backbend in Durrett [18]. However, we modify the proof to take care of the case with backbend.
Consider the random variable

$$M_1 = \max\{|M - w_b|, |M - u_b|\}.$$

As $\alpha^b(p) > 0$, we have $\bar{r}^b_{0,n} \to \infty$ a.s. This fact and the fact that the random variables u_b and w_b have geometric tails imply that there is an even integer $M < \infty$ so that

$$\mathbb{P}(\bar{r}^b_{0,n} > -M_1 \text{ for all } n) \ge 0.51.$$

For $A \subset (-\infty, \infty)$ let

$$\begin{aligned}
\xi_n^{b,A} &= \{x : (y, b) \to (x, n + b) \text{ for some } y \in A\},\\
r_n^{b,A} &= \sup \xi_n^{b,A}\,, \qquad l_n^{b,A} = \inf \xi_n^{b,A}\,, \qquad \tau^{b,A} = \inf\{n : \xi_n^{b,A} = \emptyset\}.
\end{aligned}$$

Clearly, for all $M \ge 0$

$$\begin{aligned}
\xi_n^{b,[-M_1,M_1]} &\subseteq \xi_n^{b,(-\infty,M_1]} \cap [l_n^{b,[-M_1,M_1]}, \infty),\\
\xi_n^{b,[-M_1,M_1]} &\subseteq \xi_n^{b,[-M_1,\infty)} \cap (-\infty, r_n^{b,[-M_1,M_1]}],\\
\xi_n^{b,[-M_1,M_1]} &\subseteq \xi_n^{b,(-\infty,\infty)} \cap [l_n^{b,[-M_1,M_1]}, r_n^{b,[-M_1,M_1]}].
\end{aligned}$$

Moreover, on $\{\xi_{n+b}^{b,[-M_1,M_1]} \ne \emptyset\}$ we have

$$\begin{aligned}
\xi_n^{b,[-M_1,M_1]} &= \xi_n^{b,(-\infty,M_1]} \cap [l_n^{b,[-M_1,M_1]}, \infty),\\
\xi_n^{b,[-M_1,M_1]} &= \xi_n^{b,[-M_1,\infty)} \cap (-\infty, r_n^{b,[-M_1,M_1]}],\\
\xi_n^{b,[-M_1,M_1]} &= \xi_n^{b,(-\infty,\infty)} \cap [l_n^{b,[-M_1,M_1]}, r_n^{b,[-M_1,M_1]}],
\end{aligned}$$

and

$$r_n^{b,[-M_1,M_1]} = r_n^{b,(-\infty,M_1]}, \qquad l_n^{b,[-M_1,M_1]} = l_n^{b,[-M_1,\infty)}.$$

Hence it follows that

$$\begin{aligned}\tau^{b,[-M_1,M_1]} &= \inf\{n : r_n^{b,[-M_1,M_1]} < l_n^{b,[-M_1,M_1]}\} \\ &= \inf\{n : r_n^{b,(-\infty,M_1]} < l_n^{b,[-M_1,\infty)}\}.\end{aligned}$$

This means

$$\{\tau^{b,[-M_1,M_1]} = \infty\} \supseteq \{l_n^{b,[-M_1,\infty)} \leq 0 \leq r_n^{b,(-\infty,M-u_b]} \text{ for all n}\}.$$

Hence, we have

$$\begin{aligned}\mathbb{P}_p(\xi_n^{b,[-M_1,M_1]} \neq \emptyset \ \text{ for all } n) &= \mathbb{P}_p(\tau^{b,[-M_1,M_1]} = \infty) \\ &\geq \mathbb{P}_p(r_n^{b,(-\infty,M_1]} \geq 0 \geq l_n^{b,[-M_1,\infty)} \text{ for all } n) \\ &\geq 2\mathbb{P}_p(r_n^{b,(-\infty,M_1]} > 0 \text{ for all } n) - 1 \\ &\geq 2\mathbb{P}_p(r_n^{b,(-\infty,0]} > -M_1 \text{ for all } n) - 1 \\ &\geq 2\mathbb{P}_p(r_n^{b,(-\infty,u_b]} > -M_1 + u_b \text{ for all } n) - 1 \\ &\geq 2\mathbb{P}_p(r_n^{b,(-\infty,u_b]} - u_b > -M_1 \text{ for all } n) - 1 \\ &\geq 2\mathbb{P}_p(\bar{r}_{0,n}^b > -M_1 \text{ for all } n) - 1 \\ &\geq 2(0.51) - 1 = 0.02.\end{aligned}$$

Here, the last inequality follows from the fact that $\bar{r}_{0,n}^b = r_n^{b,(-\infty,u_b]} - u_b$. Moreover, it is easily seen that $\mathbb{P}_p(\xi_b^{0,\{0\}} \supseteq (2\mathbb{Z} \cap [-M_1, M_1])) > 0$. Hence

$$\mathbb{P}_p(|C^b| = \infty) > P_1 \times P_2 > 0,$$

where $P_1 = \mathbb{P}_p(\xi_n^{b,[-M_1,M_1]} \neq \emptyset$ for all $n)$ and $P_2 = \mathbb{P}_p(\xi_{M_1+b}^{0,\{0\}} \supseteq (2\mathbb{Z} \cap [-M_1, M_1]))$. $\square$

By a simple coupling argument, we have that $\alpha_b(p)$ is a non-decreasing function of p. Then it follows from the above discussion that

$$\sup\{p : \alpha_b(p) < 0\} \leq p_c^b \leq \inf\{p : \alpha_b(p) > 0\}.$$

In the following theorem, we show that $p_c^b = \inf\{p : \alpha_b(p) > 0\}$.

Theorem 22 (*). *We have*

$$p_c^b = \inf\{p : \alpha_b(p) > 0\}.$$

Proof. To prove Theorem 22 it is sufficient to show that if $\alpha_b(p_2) > -\infty$ and $p_1 > p_2$, then

$$\alpha_b(p_1) - \alpha_b(p_2) \geq 2(p_1 - p_2). \tag{6.1}$$

We show this in three steps.

Step 1. In this step we show that if $A \supseteq B$ are infinite subsets of $\{-2, -4, \ldots\}$, then

$$\mathbb{E}\left(r_n^{b,B\cup\{0\}} - r_n^{b,B}\right) \geq \mathbb{E}\left(r_n^{b,A\cup\{0\}} - r_n^{A}\right) \geq 2. \tag{6.2}$$

From the definition of $\xi_n^{b,S}$ it is immediate that

$$\xi_n^{b,C\cup D} = \xi_n^{b,C} \cup \xi_n^{b,D},$$

which means

$$r_n^{b,C\cup D} = \max\{r_n^{b,C}, r_n^{b,D}\}$$

and

$$r_n^{b,C\cup D} - r_n^{b,C} = \max\{0, r_n^{n,D} - r_n^{n,C}\} = (r_n^{b,D} - r_n^{b,C})^+.$$

From the above we see

$$\begin{aligned} r_n^{b,B\cup\{0\}} - r_n^{b,B} = (r_n^{b,\{0\}} - r_n^{b,B})^+ &\geq (r_n^{b,\{0\}} - r_n^{n,A})^+ \qquad [r_n^{b,A} \geq r_n^{b,B}] \\ &= r_n^{b,A\cup\{0\}} - r_n^{b,A}. \end{aligned}$$

Now (6.2) follows from the observation that by translation invariance,

$$\mathbb{E}\left(r_n^{b,\{0,-2,\ldots\}} - r_n^{b,\{-2,-4,\ldots\}}\right) = 2. \tag{6.3}$$

Step 2. In this step we show that if $p_1 > p_2$ and $\alpha_n^b(p) = \mathbb{E}(\bar{r}_n^b)$, then

$$\alpha_n^b(p_1) - \alpha_n^b(p_2) \geq 2(1 - (1 - (p_1 - p_2))^n). \tag{6.4}$$

We construct the systems with parameters p_1 and p_2 on the same space in the same way as Durrett [18] has done for the no backbend case. For completeness, we present this here once again.

To each edge e, assign an independent random variable U_e that is uniformly distributed on $(0,1)$. Call an edge open if U_e is less than the parameter value, and closed otherwise. Let $\bar{r}_{1,n}^b$ and $\bar{r}_{2,n}^b$ be the location of $\xi_n^{b,(-\infty,-u_b]}$ in the systems with parameters p_1 and p_2, respectively. Let $\tau^b = \inf\{n : \bar{r}_{1,n}^b > \bar{r}_{2,n}^b\}$. We note that for the random time τ^b, the random variables $\bar{r}_{1,\tau^b}^b$ and $\bar{r}_{2,\tau^b}^b$ for the respective parameters p_1 and p_2 are independent of the edges in $E_{0,\tau_b - b}$.

So applying (6.2) and the strong Markov property we have

$$\begin{aligned} \mathbb{E}\left(\bar{r}_{1,n}^b - \bar{r}_{2,n}^b\right) &= \sum_{t=0}^{\infty} \mathbb{E}\left(\bar{r}_{1,n}^b - \bar{r}_{2,n}^b \,\middle|\, \tau^b = t\right) \mathbb{P}(\tau^b = t) \\ &= \sum_{t=0}^{n} \mathbb{E}\left(\bar{r}_{1,n}^b - \bar{r}_{2,n}^b \,\middle|\, \tau^b = t\right) \mathbb{P}(\tau^b = t) \\ &\geq 2\mathbb{P}(\tau^b \leq n). \end{aligned}$$

At each stage n, there is at least probability $p_1 - p_2$ that $\bar{r}_{1,n+1}^b - \bar{r}_{2,n+1}^b \geq \bar{r}_{1,n}^b - \bar{r}_{2,n}^b + 1$. Hence

$$\mathbb{P}(\tau^b \leq n) \geq (1 - (1 - (p_1 - p_2))^n).$$

This completes Step 2.

Step 3. We now complete the proof of Theorem 22 in this step. Let $\delta = (p_1 - p_2)/M$ where M is a large integer. Then using (6.4)

$$\begin{aligned}\alpha_n^b(p_1)-\alpha_n^b(p_2) &= \sum_{m=1}^{Mn}\left[\alpha_n\left(p_1+\frac{m\delta}{n}\right)-\alpha_n\left(p_1+\frac{(m-1)\delta}{n}\right)\right] \\ &\geq 2Mn\left(1-\left(1-\frac{\delta}{n}\right)^n\right).\end{aligned}$$

Dividing both sides of the inequality by n and letting $n\to\infty$ we see that

$$\alpha^b(p_1)-\alpha^b(p_2)\geq 2M\left(1-\exp\left(-\frac{p_1-p_2}{M}\right)\right).$$

Letting $M\to\infty$ yields

$$\alpha^b(p_1)-\alpha^b(p_2)\geq 2(p_1-p_2),$$

which proves Theorem 22. □

6.4 Stationary distribution for the edge process

In this section we discuss the stationary distribution for the edge process.

Theorem 23 (*). *If $p\geq p_c$, then there is an initial distribution μ concentrated on the infinite subsets of $\{\dots,-4,-2,0\}$ which contain 0, in such a way that r_n^μ has stationary increments.*

Proof. We start by introducing a family of "reset approximations" $\hat{\xi}_{n,m}^b$, which start from $\hat{\xi}_{n,0}^b=(-\infty,u_b]$ and evolve according to the following rules:

For $k=1,2,\dots$

1. If $(m+1)\notin n\mathbb{Z}$, then $\hat{\xi}_{n,m+1}^b=\{x:(y,m+b)\to(x,m+b+1)\text{ for some }y\in\hat{\xi}_{n,m}^b\}$.
2. If $(m+1)\in n\mathbb{Z}$, then $\hat{\xi}_{n,m+1}^b=(-\infty,\hat{r}_{n,m+1}^b]$, where $\hat{r}_{n,m+1}^b=\sup\{x:(y,m+b)\to(x,m+b+1)\text{ for some }y\in\hat{\xi}_{n,m}^b\}-u_b$.

Let $\hat{r}_{n,m}^b=\sup\hat{\xi}_{n,m}^b-u_b$ for all $n,m\in\mathbb{N}$. Then for fixed n, the increments $X_{n,k}^b:=\hat{r}_{n,k}^b-\hat{r}_{n,k-1}^b$ of these processes are not stationary, but they are periodic with periodicity n. To construct a stationary process out of $X_{n,k}^b$, we introduce an independent r.v. U_n with $\mathbb{P}(U_n=k)=1/n$ for $0\leq k<n$ and consider the process $Y_{n,k}^b:=X_{n,k+U_n}^b$. Let, for $a_1,\dots,a_s\in\mathbb{R}$,

$$\begin{aligned}\mathbb{P}(Y_{n,1}^b<a_1,\dots,Y_{n,s}^b<a_s) &= \mathbb{P}(X_{n,1+U_b}^b<a_1,\dots,X_{n,s+U_b}^b<a_s) \\ &= \sum_{i=0}^{n-1}\mathbb{P}(X_{n,1+U_b}^b<a_1,\dots,X_{n,s+U_b}^b<a_s\mid U_n=i)\mathbb{P}(U_n=i) \\ &= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{P}(X_{n,1+i}^b<a_1,\dots,X_{n,s+i}^b<a_s) \\ &= \mathbb{P}(X_{n,1}^b<a_1,\dots,X_{n,s}^b<a_s).\end{aligned}$$

Similarly, for $l\in\mathbb{N}$,

$$\begin{aligned}\mathbb{P}(Y_{n,1+l}^b<a_1,\dots,Y_{n,s+l}^b<a_s) &= \mathbb{P}(X_{n,1+l+U_b}^b<a_1,\dots,X_{n,s+l+U_b}^b<a_s) \\ &= \mathbb{P}(X_{n,1}^b<a_1,\dots,X_{n,s}^b<a_s).\end{aligned}$$

This shows that the process $\{Y^b_{n,k} : k \geq 1\}$ is stationary with

$$\begin{aligned}
\mathbb{E}(Y^b_{n,1}) &= \sum_{i=0}^{n} \mathbb{E}(X^b_{n,i+U_n} \mid U_n = i)\mathbb{P}(U_n = i) \\
&= \frac{1}{n}\sum_{i=0}^{n} \mathbb{E}(X^b_{n,i}) \\
&= \frac{1}{n}\mathbb{E}(\hat{r}^b_{n,n} - \hat{r}^b_{n,0}) \\
&\geq \alpha^b(p).
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\mathbb{E}|Y^b_{n,1}| &\leq \frac{1}{n}\sum_{k=1}^{n}(\mathbb{E}|\hat{r}^b_{n,k}| + \mathbb{E}|\hat{r}^b_{n,k-1}|) \\
&\leq \frac{1}{n}2n(2\frac{1-\gamma_1}{\gamma_1} + 2\frac{1-\gamma_2}{\gamma_2}) \\
&= 4(\frac{1-\gamma_1}{\gamma_1} + \frac{1-\gamma_2}{\gamma_2}) \\
&< \infty
\end{aligned}$$

for $p \in (0,1)$.
It follows from the above calculations that if we consider the processes $\{Y^b_{n,m} : m \geq 1\}$ as a sequence of random elements of $\mathbb{R} \times \mathbb{R} \times \cdots$, then the sequence is tight. So we can find a sequence $n_j \to \infty$ such that $\{Y^b_{n_j,m} : m \geq 1\}$ converges in distribution (in $\mathbb{R} \times \mathbb{R} \times \cdots$) to a limit $\{Y^b_m : m \geq 1\}$ with $\mathbb{E}(Y^b_1) \leq 4(\frac{1-\gamma_1}{\gamma_1} + \frac{1-\gamma_2}{\gamma_2})$.

To construct the measure μ, we have to take another sequence. Let us define the following random variables:

$$\hat{r}^b_{n,U_n} = \sup \hat{\xi}^b_{n,U_n} - u_b, \quad \tilde{\xi}^b_{n,m} = \hat{\xi}^b_{n,m+U_n} - \hat{r}^b_{n,U_n},$$
$$\tilde{r}^b_{n,m} = \sup \tilde{\xi}^b_{n,m}, \quad \tilde{Y}^b_{n,m} = \tilde{r}^b_{n,m} - \tilde{r}^b_{n,m-1}.$$

Since

$$\tilde{Y}^b_{n,m} = \tilde{r}^b_{n,m} - \tilde{r}^b_{n,m-1} = \hat{r}^b_{n,m} - \hat{r}^b_{n,m-1} = \hat{Y}^b_{n,m},$$

we have

$$\{\tilde{Y}^b_{n,m}\} \stackrel{\mathcal{D}}{=} \{\tilde{Y}^b_{n,m}\}.$$

Note that $\tilde{\xi}^b_{n,0}$ is a subset of $(-\infty, u_b]$ and $u_b \in \tilde{\xi}^b_{n,0}$. For $B \in \mathcal{B}$ and $k \in \mathbb{N}$ fixed, we have

$$\begin{aligned}
\mathbb{P}(\tilde{\xi}^b_{n,m} \in B) &= \mathbb{P}(\tilde{\xi}^b_{n,m} \in B, U_n \leq n-k) + \mathbb{P}(\tilde{\xi}^b_{n,m} \in B, U_n > n-k) \\
&= \mathbb{P}(\hat{\xi}^b_{n,m+U_n} - \hat{r}^b_{n,U_n} \in B, U_n \leq n-k) + \mathbb{P}(\tilde{\xi}^b_{n,m} \in B, U_n > n-k).
\end{aligned}$$

Note that for fixed k, $\mathbb{P}(U_n > n-k) \to 0$ as $n \to \infty$. This means that as n gets larger, the finite dimensional distribution of $\tilde{\xi}^b_{n,m}$ become arbitrarily close to those of $\hat{\xi}^{b,\mu_n}_m$ where μ_n is the distribution of $\tilde{\xi}^b_{n,0}$.

Our idea of constructing μ from μ_n is the same as in Durrett [18]. However, some care needs to be taken for the backbend model. We describe this here for the sake of completeness. Since

μ_n are probability measures on the compact space $\{0,1\}^{\{\dots,-4,-2,0\}}$, the sequence μ_{n_j} has a further subsequence $\mu_{n'_k}$ which converges weakly to a limit μ. Hence for each M the distribution of $\{\tilde{\xi}^b_{n'_k,m} : 0 \le m \le M\}$ converges weakly to $\{\xi^{b,\mu}_m : 0 \le m \le M\}$. Let

$$r^{b,\mu}_m = \sup \xi^{b,\mu}_m - u_b \qquad \text{and} \qquad X^{b,\mu}_m = r^{b,\mu}_m - r^{b,\mu}_{m-1}.$$

Then

$$\{X^{b,\mu}_m\} \stackrel{\mathcal{D}}{=} \{Y^b_m\}$$

and hence $X^{b,\mu}_m$, i.e., the increments of $r^{b,\mu}_m$ form a stationary sequence. We show

$$\frac{r^{b,\mu}_n}{n} \to \alpha^b(p) \text{ a.s.}$$

Since $\mathbb{E}(Y^b_{n,1}) \ge \alpha^b(p)$ and the process $Y^b_{n,1}$ converges in distribution to Y^b_1, we have

$$\mathbb{E}(Y^b_1) = \lim_{n\to\infty} \mathbb{E}(Y^b_{n,1}) \ge \alpha^b(p). \tag{6.5}$$

Noting the fact that $r^{b,\mu}_n \le \bar{r}^b_{0,n}$, we have

$$\limsup_{n\to\infty} \frac{r^{b,\mu}_n}{n} \le \alpha^b(p) \text{ a.s.}$$

By means of the Ergodic Theorem, we deduce that as $n \to \infty$

$$\frac{1}{n} r^{b,\mu}_n = \frac{1}{n} \sum_{k=1}^{n} Y^b_k \to \mathbb{E}[Y^b_1 | \mathcal{J}] \text{ a.s.}$$

where $\mathcal{J}$ is the shift invariant σ-algebra. Lastly, as

$$\mathbb{E}(Y^b_1) = \mathbb{E}(\mathbb{E}[Y^b_1|\mathcal{J}]) \le \alpha^b(p)$$

we have by (6.5)

$$\mathbb{E}(Y^b_1) = \alpha^b(p) \text{ and } \mathbb{E}[Y^b_1|\mathcal{J}]) = \alpha^b(p)$$

and hence

$$\frac{1}{n} r^{b,\mu}_n \to \alpha^b(p) \text{ a.s.}$$

□

6.5 Strict monotonicity of the limiting direction α^b over b

It is clear from the characterization of p^b_c in Section 6.3 that $\alpha^b(p) < 0 < \alpha^{b+1}(p)$ for $p^{b+1}_c < p < p^b_c$. In this section we show $\alpha^b(p) < \alpha^{b+1}(p)$ for $p > p^b_c$.

Theorem 24 (*). *For $p \ge p^b_c$*

$$\alpha^b(p) < \alpha^{b+1}(p).$$

Proof. We say that a b-path is a strict b-path if it is a b-path but not a $(b-1)$-path.

We take fixed $N = b+2$ and $p > p_c^b$. Let A^b_{mN}, $m = 1, \ldots$, the event that there exists a strict $(b+1)$-path below level mN from $\bar r^b_{mN} + 2(m-1)(b+1)$ to $\bar r^b_{mN} + 2m(b+1)$ (see Figure 6.2). Note

$$\pi_b := \mathbb{P}_p(\mathsf{A}^b_{mN}) = p^{2b}(1-p)^{2b-1} > 0 \text{ for all } m \geq 1.$$

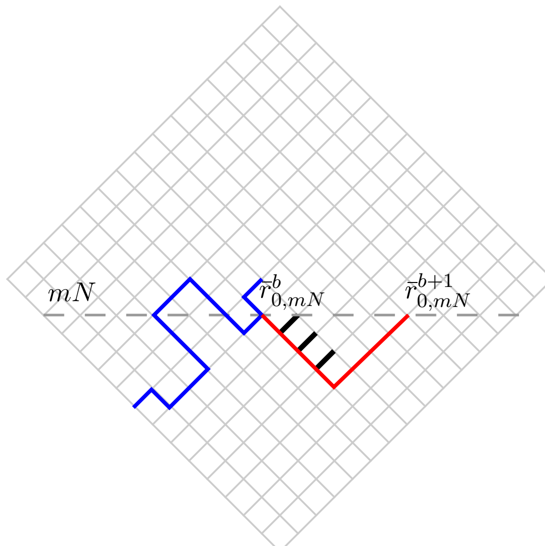

Figure 6.2: Event A^b_{mN} $(b=3)$

Note that $\mathbb{E}\left(\bar r^{b+1}_{0,N} - \bar r^b_{0,N}\right) \geq 2(b+1)\pi_b$ and by translation invariance

$$\mathbb{E}\left(\bar r^{b+1}_{0,N+k} - \bar r^b_{0,N+k}\right) \geq \mathbb{E}\left(\left(\bar r^{b+1}_{0,N+k} - \bar r^b_{0,N+k}\right)\mathbb{1}_{\mathsf{A}^b_N}\right) \geq 2(b+1)\pi_b$$

for all $k \geq 0$. Now, when $k = N$, we have $\mathbb{E}\left(\bar r^{b+1}_{0,2N} - \bar r^b_{0,2N}\right) \geq 4(b+1)\pi_b$ and, again, by translation invariance,

$$\mathbb{E}\left(\bar r^{b+1}_{0,N+l} - \bar r^b_{0,N+l}\right) \geq \mathbb{E}\left(\left(\bar r^{b+1}_{0,2N+l} - \bar r^b_{0,2N+l}\right)\mathbb{1}_{\mathsf{A}^b_{2N}}\right) \geq 4(b+1)\pi_b$$

for all $l \geq 0$.

So, by induction (see Figure 6.3), we have

$$\mathbb{E}\left(\bar r^{b+1}_{0,mN} - \bar r^b_{0,mN}\right) \geq 2(b+1)\pi_b m.$$

Dividing the above expression by mN and taking the limit as $m \to \infty$, we obtain

$$\alpha^{b+1}(p) - \alpha^b(p) \geq \frac{2(b+1)\pi_b}{N}.$$

□

6.6 Exponential estimates for $p < p_c^b$

In this section we provide some exponential estimates for p when $p < p_c^b$. Let $\xi_n^{b,0} = \{x : (0,b) \to (x, n+b)\}$.

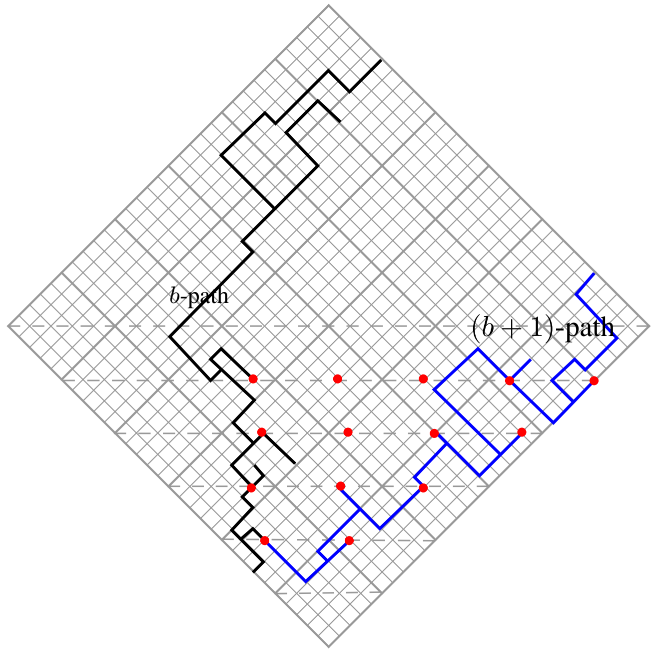

Figure 6.3: Average b-path and $(b+1)$-path

Claim 3 (*). *If $p < p_c^b$, then there is a positive constant $\gamma^b := \gamma^b(p)$ such that*

$$\mathbb{P}(\xi_n^{b,0} \neq \emptyset) \leq e^{-\gamma^b n} \tag{6.6}$$

and

$$\frac{1}{n} \log \mathbb{P}(\xi_n^{b,0} \neq \emptyset) \to -\gamma^b$$

as $n \to \infty$.

Proof. Consider the event $\{\xi_{m+n+2b}^{b,0} \neq \emptyset\}$. Then

$$\begin{aligned}
\mathbb{P}(\xi_{m+n+2b}^{b,0} \neq \emptyset) &\leq \mathbb{P}(\xi_m^{b,0} \neq \emptyset,\ \xi_{m+2b,m+n+2b}^{b,0} \neq \emptyset) \\
&= \mathbb{P}(\xi_m^{b,0} \neq \emptyset)\mathbb{P}(\xi_{m+2b,m+n+2b}^{b,0} \neq \emptyset) \\
&= \mathbb{P}(\xi_m^{b,0} \neq \emptyset)\mathbb{P}(\xi_n^{b,0} \neq \emptyset).
\end{aligned}$$

Taking the log of both sides and denoting $\log \mathbb{P}(\xi_n^{b,0} \neq \emptyset)$ by a_n, we have from the last inequality

$$a_{m+n+2b} \geq a_m + a_n.$$

For fixed $m > 2b$, we have for every $k \in \mathbb{N}$

$$\begin{aligned}
a_{km} &= a_{km-2b+2b} \\
&\geq a_{m-2b} + a_{(k-1)m} \\
&\geq a_{m-2b} + a_{m-2b} + a_{(k-2)m} \\
&\dots \\
&\geq k a_{m-2b}.
\end{aligned}$$

For fixed m consider $n = km + r$ where $r < m$ and $k \in \mathbb{N}$. Then using the above inequality it follows that

$$\begin{aligned}
\frac{a_{n+2b}}{n+2b} &= \frac{a_{km+r+2b}}{km+r+2b} \\
&\geq \frac{a_{km}+a_r}{km+r+2b} \\
&\geq \frac{ka_{m-2b}}{km+r+2b} + \frac{a_r}{km+r+2b} \\
&= \frac{a_{m-2b}}{m+\frac{r+2b}{k}} + \frac{a_r}{km+r+2b}.
\end{aligned}$$

Taking the limit infimum over n on both sides and noting that $k \to \infty$ as $n \to \infty$, we have

$$\liminf_n \frac{a_{n+2b}}{n+2b} \geq \frac{a_{m-2b}}{m}.$$

Taking the limit supremum over m on both sides we have

$$\liminf_n \frac{a_{n+2b}}{n+2b} \geq \limsup_m \frac{a_{m-2b}}{m}.$$

This means that the limit of the sequence $\{\frac{a_n}{n}\}$ exists and

$$\lim_{n\to\infty} \frac{a_n}{n} = \sup_{m\geq 2b} \frac{a_m}{m} =: -\gamma^b.$$

Note that up to this point everything holds for any $p \leq 1$. We only need to prove that $\gamma^b > 0$. Here, we use the fact that $p < p_c^b$. If $p < p_c^b$, there exists a (large) N such that $\mathbb{E}\bar{r}^b_{2b,N} < 0$, and from the subadditivity property and $\bar{r}^b_{0,n} \leq \bar{r}^b_{m,n}$ for all m, n, it follows that

$$\begin{aligned}
\bar{r}^b_{0,mN} &\leq \bar{r}^b_{0,N} + \bar{r}^b_{N,mN} \\
&\leq \bar{r}^b_{0,N} + \bar{r}^b_{N,2N} + \cdots + \bar{r}^b_{(m-1)N,mN} \\
&\leq \bar{r}^b_{2b,N} + \bar{r}^b_{N+2b,2N} + \cdots + \bar{r}^b_{(m-1)N+2b,mN} =: S_m,
\end{aligned}$$

where S_m is a random walk with $\mathbb{E}(S_1) < 0$. Note $\{\bar{r}^b_{(k-1)N+2b,N} : k = 0, \ldots, m\}$ is a set of independent and identically distributed random variables. We recall the random variable K_R defined in the proof of Proposition 4. We define K_R in a similar way but with the parameter $\gamma_R = (1-p)^{N+2b}$ $(p > 0)$. Hence $S_1 \leq K_R$ and for every $\theta \in [0, -1/2\log(1-\gamma_R))$, we have

$$\varphi(\theta) := \mathbb{E}(\exp(\theta S_1)) \leq \mathbb{E}(\exp(\theta K_R)) = \frac{\gamma_R}{1-(1-\gamma_R)e^\theta} < \infty.$$

For M large and positive,

$$\frac{\mathbb{E}\left(\exp(\theta S_1)\right) - 1}{\theta} \leq \frac{\mathbb{E}\left(\exp(\theta(S_1 \vee -M))\right) - 1}{\theta}.$$

Taking the limit supremum of both sides as $\theta \to 0$ we obtain

$$\limsup_{\theta\to 0} \frac{\mathbb{E}\left(\exp(\theta S_1)\right) - 1}{\theta} \le \mathbb{E}\left(S_1 \vee -M\right).$$

Lastly, letting $M \to \infty$,

$$\limsup_{\theta\to 0} \frac{\mathbb{E}\left(\exp(\theta S_1)\right) - 1}{\theta} \le \mathbb{E}\left(S_1\right) < 0.$$

Write $\varphi(\theta) = \mathbb{E}\left(\exp(\theta S_1)\right)$. Then there is a $\theta_0 > 0$ with $\varphi(\theta_0) < 1$. So

$$\mathbb{P}(S_m \ge 0) \le \mathbb{E}\exp(\theta_0 S_m) \le (\varphi(\theta_0))^m.$$

Hence

$$\mathbb{P}(\bar{r}^b_{0,mN} \ge 0) \le \mathbb{P}(\theta_0 S_m \ge 0) \le \mathbb{P}(\exp(\theta_0 S_m) \ge 1) \le \mathbb{E}\exp(\theta_0 S_m) \le (\varphi(\theta_0))^m,$$

i.e., $\mathbb{P}(\bar{r}^b_{0,mN} \ge 0) \to 0$ exponentially fast when $m \to \infty$.

To conclude that the same thing is true for $\mathbb{P}(\xi^{b,0}_n \ne \emptyset)$, observe that

$$\mathbb{P}(\xi^{b,0}_n = \emptyset) \ge \mathbb{P}(\bar{r}^b_{0,n} < 0 < \bar{l}^b_{0,mN}),$$

so

$$\mathbb{P}(\xi^{b,0}_n \ne \emptyset) \le \mathbb{P}(\bar{r}^b_{0,n} \ge 0) + \mathbb{P}(\bar{l}^b_{0,n} \le 0) = 2\mathbb{P}(\bar{r}^b_{0,n} \ge 0)$$

and since for every large n we have that $n \ge mN$ for large m

$$\mathbb{P}(\xi^{b,0}_n = \emptyset) \le \mathbb{P}(\xi^{b,0}_{mN} = \emptyset) \le 2\mathbb{P}(\bar{r}^b_{0,mN} \ge 0),$$

so the proof of (6.6) is complete. □

Claim 4 (*)**.** *If $s > \alpha^b$, then there are constants C_b and γ_b such that*

$$\mathbb{P}\left(\bar{r}^b_n > sn\right) \le C_b e^{-\gamma_b n}. \tag{6.7}$$

Proof. If $s > \alpha^b$, we have the following:

- $\mathbb{E}\left(\bar{r}^b_{2b,N_b} - sN_b\right) < 0$ for some $N_b > 0$,
- $\mathbb{P}\left(\bar{r}^b_n > sn\right) \le \mathbb{P}\left(\xi^{b,0}_n \ne \emptyset\right)$,
- $\bar{r}^b_{mN_b} - smN_b$ satisfies

$$\bar{r}^b_{0,mN_b} - smN_b \le (\bar{r}^b_{2b,N} - sN_b) + (\bar{r}^b_{N+2b,2N} - sN_b) + \cdots + (\bar{r}^b_{(m-1)N+2b,mN} - sN_b),$$

and now the conclusion follows from the proof of Claim 3. □

6.7 The nature of the limiting direction α^b as the number of backbends b goes to infinity

In this section we make some connections between the unoriented percolation model and the backbend percolation model. Intuitively it seems that when b gets larger and larger in a b-backbend percolation model then the model becomes closer and closer to the unoriented percolation model in $\mathbb{Z}^2$. In this spirit we show that for $p > 1/2$, the limiting direction $\alpha^b(p)$ diverges to infinity as b increases to infinity.

Theorem 25 (*)**.** *For $p > \frac{1}{2}$, $\alpha^b(p) \to \infty$ as $b \to \infty$.*

Proof (first approach). We start by defining a *top–bottom crossing* of a box $[a,b] \times [c,d] \subset \mathcal{L}$: it is a path in the box from $[a,b] \times \{c\}$ to $[a,b] \times \{d\}$; a *left–right crossing* is defined similarly.

Fix $q < \frac{1}{2}$. We define

$$\tau_1(n,q) = \mathbb{P}_q\{\exists \text{ an open left–right crossing of } [0,n] \times [0,3n]\},$$

$$\tau_2(n,q) = \mathbb{P}_q\{\exists \text{ an open top–bottom crossing of } [0,3n] \times [0,n]\}.$$

Since $q < \frac{1}{2}$, there is an N such that

$$\tau_i(N,q) \leq \kappa$$

for $i = 1,2$ and $\kappa := \frac{1}{4}(50e)^{-121}$, and so the hypotheses of Theorem 5.1 in [33] are satisfied. Therefore there is $0 < C_1, C_2 < \infty$ such that

$$\mathbb{P}\{\#W \geq n\} \leq C_1 e^{-C_2 n} \quad \text{for } n \geq 0,$$

where W is the set of all vertices which belong to the open cluster of 0 in $\mathcal{L}$ and $\#W$ denotes its cardinality.

We consider the next set of boxes. Let $L > 0$. Now define the boxes

$$A_k := [0,kL] \times [0,2kL], \qquad B_k := [0,L^k] \times [0,kL], \qquad k \geq 2.$$

For every A_k, we consider the parallelogram A'_k whose vertices are the points $(0,0), (kL,-kL)$, $(0,3kL), (kL,2kL)$, and for B_k we consider the parallelogram with vertices $(0,0), (L^k + kL, 0)$, $(L^k, kL), (-kL, kL)$ (see Figure 6.4). We define a top–bottom crossing in B'_k as a path which starts at the line that joins $(-kL,kL)$ with (L^k,kL) and finishes at the line that joins $(0,0)$ with $(L^k + kL, 0)$, and a left–right crossing in A'_k as a path which starts at the line that joins $(0,0)$ with $(0,3kL)$ and finishes at the line that joins $(kL,-kL)$ with $(kL,2kL)$.

Let E_k be the event that there exists a top–bottom crossing in B'_k and let F_k be the event that there exists a left–right crossing in A'_k. So, by the previous discussion, we have

$$\mathbb{P}_q(E_k) \leq (L^k + kL)L\mathbb{P}_q(\#W \geq \sqrt{2}kL) \leq c_1 L^k e^{-c_2 kL},$$

$$\mathbb{P}_q(F_k) \leq 3kL\mathbb{P}_q(\#W \geq \sqrt{2}kL) \leq c_1 k e^{-c_2 kL}$$

for $k \geq 0$ and suitable positive constants c_1 and c_2.

Now, change the model "open with probability q" to "open with probability $p = 1-q$." Note that if there being a top–bottom crossing in A'_k implies that there is a top–bottom crossing in A_k, and there being a left–right crossing in B'_k implies that there is a left–right crossing in B_k. From

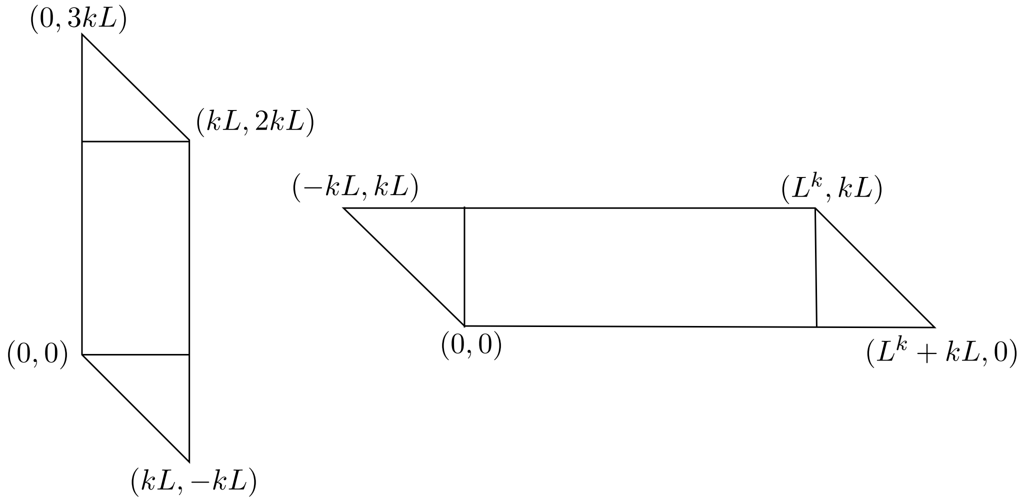

Figure 6.4: Parallelogram A_k', B_k'

the above inequalities we obtain

$$\mathbb{P}_p(\exists \text{ open top–bottom crossing in } A_k) \geq 1 - c_1 L^k e^{-c_2 kL},$$

$$\mathbb{P}_p(\exists \text{ open left–right crossing in } B_k) \geq 1 - c_1 L^k e^{-c_2 kL}.$$

We consider the configuration of the boxes A_k and B_k as is shown in Figure 6.5, such that

- up to k, the length of the configuration is

$$l := \sum_{j=2}^{k} L^j - \sum_{j=3}^{k} jL = \frac{L^{k+1}-1}{L-1} - \frac{L}{2}k(k+1) + 2L - 1.$$

- Up to k, the height of the configuration is

$$h := \sum_{j=2}^{k} 2jL - \sum_{j=3}^{k} (j-1)L = \frac{Lk}{2}(k+3) - L.$$

Suppose that the sequence (with respect to b) $\alpha^b(p)$ ($p > 1/2$) is bounded. Then $\alpha^\infty := \lim_{b\to\infty} \alpha^b(p)$ is finite. Hence, when L is large, $0 < l - \alpha^\infty h$ is also large.

If we concatenate a top–bottom crossing in A_j with a left–right crossing in B_j, for $j = 1, \ldots, k$, we obtain a $2kL$–path. So, by the FKG inequality, the probability that this $2kL$-path exists is at least

$$\pi_0 := \prod_{j=2}^{k} (1 - c_1 L^k e^{-c_2 kL}).$$

Since $c_1 L^k e^{-c_2 kL} = c_1 \left(L/e^{c_2 L}\right)^k$, we can choose L and k sufficiently large so that $\pi_0 > 1 - \varepsilon$ for some prescribed $\varepsilon > 0$. We choose a K such that $\pi_0 > 1 - \varepsilon$ for some fixed $\varepsilon > 0$ and L, and in the following we consider this $2KL$-path.

We can extend the $2KL$-path such that it still below of line with the slope α^∞ by adding a “copy of itself” in the following way (see Figure 6.6). We select one point in the original $2KL$-path.

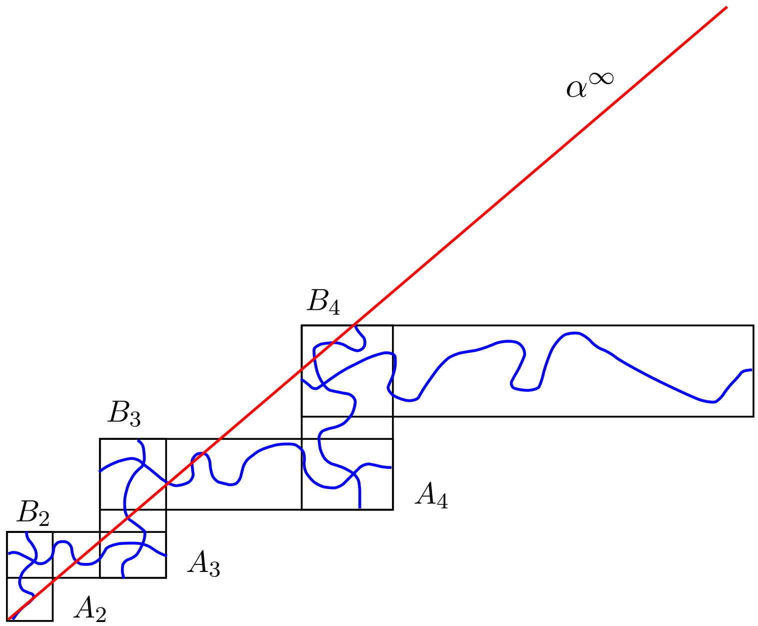

Figure 6.5: Configuration of the boxes A_k, B_k.

The probability that there is a $2KL$-path from the previous selected point is at least $1-(1-\pi_0)$. After that, we add j $2KL$-paths, we choose $j+1$ points on these paths, and now we have that the probability that there is a $2KL$-path starting from some of these points is at least $1-(1-\pi_0)^{j+1}$. By the FKG inequality, the probability that we can obtain an infinite $2KL$-path is at least

$$\pi_1 := \pi_0 \prod_{j=1}^{\infty} \left(1-(1-\pi_0)^j\right).$$

We can take L such that $\pi_1 > 0$, but this is a contradiction to the boundedness of $\alpha^b(p)$. Therefore $\alpha^b(p) \to \infty$ as $b \to \infty$. $\square$

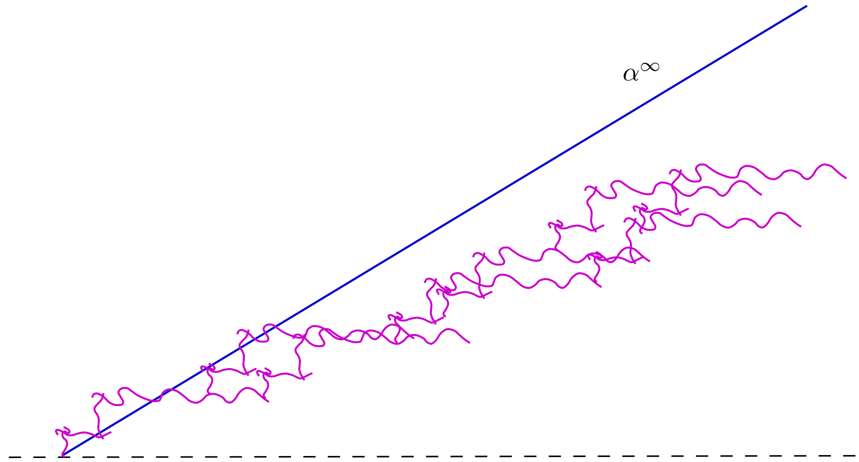

Figure 6.6: $2KL$-paths

6.8 A construction for studying $p > p_c^b$

Let $\mathcal{L} = \{(m,n) \in \mathbb{Z}^2 : m+n \text{ is even}, n \geq 0\}$. For δ small and L large, we define, for each $(m,n) \in \mathcal{L}$,

$$\begin{aligned} C_{m,n} &= ((1-\delta)\alpha^b Lm, Ln), \\ R_{m,n} &= C_{m,n} + [-(1+\delta)\alpha^b L, (1+\delta)\alpha^b L] \times [0, (1+\delta)L], \end{aligned}$$

where α^b is the constant defined in Theorem 21.

Let $A_{0,0}$ be the parallelogram with vertices

$$\begin{aligned} u_0 &= (-1.5\delta\alpha^b L, 0), & v_0 &= (-.5\delta\alpha^b L), \\ u_1 &= u_0 + (1+\delta)(\alpha^b L, L), & v_1 &= v_0 + (1+\delta)(\alpha^b L, L), \end{aligned}$$

and let $B_{0,0} = -A_{0,0}$. We say that the event $G_{0,0}$ occurs if there is a b-path from $[u_0, v_0]$ to $[u_1, v_1]$ which stays in $A_{0,0}$ and there is a b-path from $[-v_0, -u_0]$ to $[-v_1, -u_1]$ which stays in $B_{0,0}$ (see Figure 6.7).

The events $G_{m,n}$ are defined by translating the last definition by $C_{m,n}$. So, to every $z \in \mathcal{L}$ there is associated a random variable $\eta(z)$ as

$$\eta(m,n) = \mathbb{1}_{G_{m,n}}.$$

The next proposition mentions some properties of this η-system.

Proposition 5 (*). *The η-system satisfies the following:*

1. *If $\delta \leq .1$, the random variable $\eta(z)$ will be $1-$dependent, that is, if we let $||(m,n)|| = (|m| + |n|)/2$ and if $z_1, \ldots, z_m$ are points with $||z_i - z_j|| > 1$ for $i \neq j$, then $\eta(z_1), \ldots, \eta(z_m)$ are independent.*

2. *If percolation occurs in the η-system, then there is an infinite path in the original system which starts in $[-1.5\delta\alpha^b L, 1.5\delta\alpha^b L]$.*

3. *If $\delta, \varepsilon > 0$ and p with $\alpha(p) > 0$, then we can pick L large enough so that $\mathbb{P}(\eta(z) = 1) > 1 - \varepsilon$.*

Proof. (i). Note that $\eta(m,n)$ depends only on the configuration in $R_{m,n}$. We take $z_1 = (m_1, n_1), z_2 = (m_2, n_2) \in \mathcal{L}$ with $||z_1 - z_2|| > 1$. Suppose that $m_1 \leq m_2$ and $n_1 = n_2$. If

$$(1-\delta)\alpha^b Lm_1 + (1+\delta)\alpha^b L \geq (1-\delta)\alpha^b Lm_2 - (1+\delta)\alpha^b L,$$

then $R_{m_1,n_1} \cap R_{m_2,n_2} \neq \emptyset$, but this is not possible, since $m_2 - m_1 \geq 3$ and

$$\frac{5}{2} \geq 2\frac{1+\delta}{1-\delta} \geq m_2 - m_1.$$

Then, if $||z_1 - z_2|| > 1$ (other possible cases are similar to that above), we have that $\eta(z_1)$ and $\eta(z_2)$ are independent.

(ii). If $z_0 = (0,0)$ and $z_1 = (1,1)$ are open, then there is a b-path from $[-1.5\delta\alpha^b L, 1.5\delta\alpha^b L] \times \{0\}$ through $C_{1,1} + ([-0.5\delta\alpha^b L, 0.5\delta\alpha^b L] \times \{0\})$ and on up to $C_{1,1} + ([-1.5\delta\alpha^b L, 1.5\delta\alpha^b L] \times \{0\})$ and to $C_{2,2} + ([-0.5\delta\alpha^b L, 0.5\delta\alpha^b L] \times \{0\})$ (see Figure 6.7). From this observation and induction we have that if there is an infinite 0-path in the η-system and then there is a corresponding infinite b-path in the original system (but not conversely).

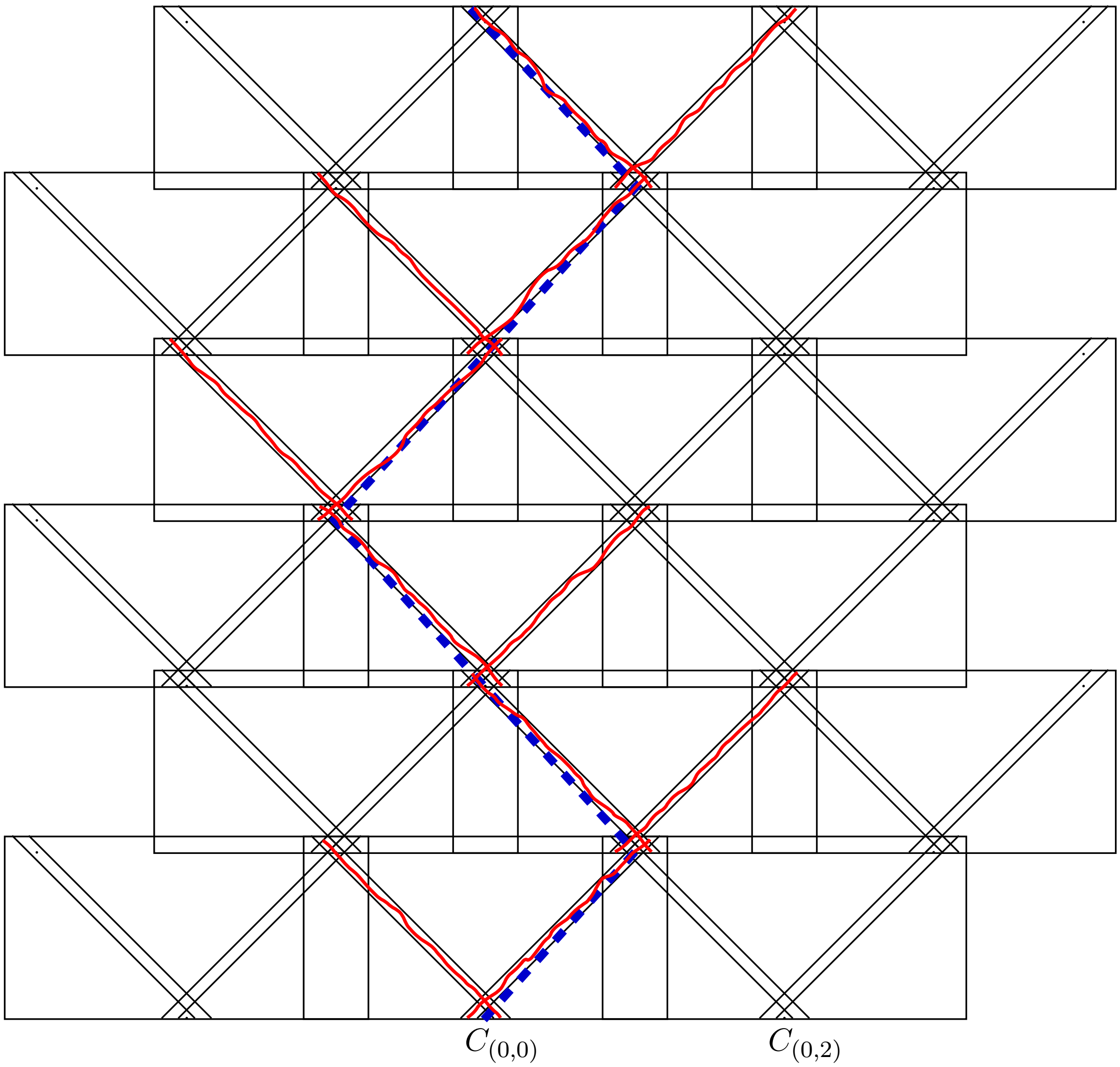

Figure 6.7: η-system

(iii). Let $\hat{r}_n^b = \sup \xi_n^{b,(-\infty,-.8\delta\alpha^b L]}$ and let $\bar{r}_n^b = \sup \xi^{b,(-\infty,0]}$. $\{\hat{r}_n^b + .8\delta\alpha^b L : n \geq 0\} \stackrel{d}{=} \{\bar{r}_n^b : n \geq 0\}$ and as $n \to \infty$, $\bar{r}_n^b/n \to \alpha^b$ a.s., so it follows that if we pick L large enough, then with probability $\geq 1 - \varepsilon/4$ we have

$$\hat{r}_{(1+\delta)L}^b > -.8\delta\alpha^b L + (1 + .9\delta)\alpha^b L$$

and for $n \leq (1+\delta)L$

$$\hat{r}_n^b \leq -.7\delta\alpha^b L + n\left(\frac{1+1.1\delta}{1+\delta}\right)\alpha^b.$$

The last two events guarantee that there is a b-path from $(-\infty, -.8\delta\alpha^b L] \times \{0\}$ up to $[(1 + .1\delta)\alpha^b L, (1+.4\delta)\alpha^b L] \times \{(1+\delta)L\}$ which does not cross the line between v_0 and v_1.

To prove that this b-path does not fall too far to the left, we observe that to travel from the line between u_0 and u_1 to $[\alpha^b L, \infty) \times \{(1+\delta)L\}$ a b-path must have an avarage slope $s := \alpha^b(1+1.5\delta)/(1+\delta) > \alpha^b$ and it follows from (6.7) that

$$\mathbb{P}\left(\bar{r}_n^b > sn\right) \leq Ce^{\gamma n},$$

so picking M large enough so that

$$\sum_{n=M}^{\infty} Ce^{\gamma n} \geq \frac{\varepsilon}{8}$$

and then considering separately the points on line between u_0 and u_1 with $y \leq (1+\delta)L - M$ and $y > (1+\delta)L - M$ (see Figure 6.8). We see that if L is large, the probability that there is a b-path connecting the line between u_0 and u_1 and $[\alpha^b L, \infty) \times \{(1+\delta)L\}$ is at most $\varepsilon/4$.

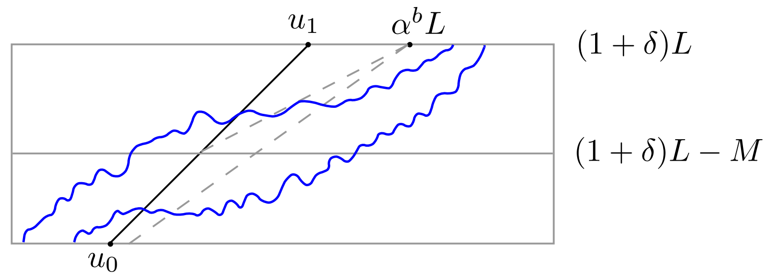

Figure 6.8: A b-path connecting the line between u_0 and u_1 and $[\alpha^b L, \infty) \times \{(1+\delta)L\}$

By the previous discussion, if L is sufficiently large, then the first half of the event $G_{0,0}$ occurs with probability $\geq 1 - \varepsilon/2$. Now the second half of the event $G_{0,0}$ has the same probability as the first. So, it follows that with probability $\geq 1 - \varepsilon$ the good event occurs. □

Theorem 26 (*). $\alpha^b(p_c^b) = 0$

Proof. In section 10 of Durrett [18], it was shown that $\mathbb{P}(\eta(z) = 1) > 1 - 3^{-36}$ and then there is a positive probability of percolation in the η-system. If $\alpha^b(p_c^b) > 0$, let $\delta = .1$ and pick L so large that $\mathbb{P}(\eta(z) = 1) > 1 - 3^{-37}$. Since there are only a finite number of bonds in $R_{0,0}$, we can choose p such that $p < p_c^b$ and $\mathbb{P}(\eta(z) = 1) > 1 - 3^{-36}$, but this is a contradiction because $\alpha^b(p) = 0$. □

Appendix A

Sub-Gaussian Random Variables

This appendix presents several facts about sub-Gaussian random variables and some of the properties which are used in Chapter 5. For details in this subject we recommend [56], [60].

A real-valued random variable ξ is said to be *sub-Gaussian* if there is some $b > 0$ such that for every $t \in \mathbb{R}$

$$\mathbb{E}\left(e^{t\xi}\right) \leq e^{b^2 t^2/2}.$$

When this condition is satisfied with a particular value of $b > 0$, we say that ξ is b-sub-Gaussian, or sub-Gaussian with parameter b.

From this definition, we have that sub-Gaussian random variables are centered and their variance has a natural upper bound in terms of the sub-Gaussian parameter.

Proposition 6. *If ξ is b-sub-Gaussian, then $\mathbb{E}(\xi) = 0$ and $\mathrm{Var}(\xi) \leq b^2$.*

Example 2. *If ξ has distribution $\mathcal{N}(0, \sigma^2)$, then an easy computation shows that any $t \in \mathbb{R}$*

$$\mathbb{E}\left(e^{t\xi}\right) = e^{\sigma^2 t^2/2},$$

i.e., ξ is sub-Gaussian with parameter σ.

Example 3. *Let ξ be a Rademacher random variable, i.e., the law of ξ is $\mathbb{P}_\xi = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ (here δ_x is the point mass at x). Then for any $t \in \mathbb{R}$*

$$\mathbb{E}\left(e^{t\xi}\right) = \frac{1}{2}e^{-t} + \frac{1}{2}e^{t} = \cosh(t) \leq e^{t^2/2},$$

so ξ is 1-sub-Gaussian.

Example 4. *Let ξ be a random variable with uniform distribution over the interval $[-a, a]$ for some fixed $a > 0$. The for any real $t \neq 0$*

$$\mathbb{E}\left(e^{t\xi}\right) = \frac{1}{2a}\int_{-a}^{a} e^{tx}dx = \frac{1}{2at}\left(e^{at} - e^{-at}\right) = \sum_{n=0}^{\infty} \frac{(at)^{2n}}{(2n+1)!},$$

since $(2n+1)! \geq n!2^n$, we see that ξ is a-sub-Gaussian.

More generally, any centered and bounded random variable is sub-Gaussian.

Theorem 27. *If ξ is a random variable with $\mathbb{E}(\xi) = 0$ and $|\xi| \leq 1$ a.s., then*

$$\mathbb{E}\left(e^{t\xi}\right) \leq \cosh(t) \quad \forall t \in \mathbb{R} \tag{A.1}$$

and so ξ is 1-sub-Gaussian. Moreover, if equality holds in (A.1) for some $t \neq 0$, then ξ is a Radamecher variable and hence equality holds for all $t \in \mathbb{R}$.

Corollary 2. *If ξ is a random variable with $\mathbb{E}(\xi) = 0$ and $|\xi| \leq b$ a.s. for some $b > 0$, then ξ is b-sub-Gaussian.*

The set of all sub-Gaussian random variables has a linear structure.

Theorem 28. *If ξ is b-sub-Gaussian, then for any $\alpha \in \mathbb{R}$, the random variable $\alpha\xi$ is $|\alpha|b$-sub-Gaussian. If ξ_1 and ξ_2 are random variables such that ξ_i is b_i-sub-Gaussian, then $\xi_1 + \xi_2$ is $(b_1 + b_2)$-sub-Gaussian.*

Note that in the previous theorem, ξ_1 and ξ_2 are not necessarily independent.

The following theorem gives equivalent conditions for a random variable to be sub-Gaussian.

Theorem 29. *For a centered random variable ξ, the following statements are equivalent:*

1. *Laplace transform condition:* $\exists\, b > 0$, $\forall\, t \in \mathbb{R}$, $\mathbb{E}\left(e^{t\xi}\right) \leq e^{b^2t^2/2}$.
2. *Sub-Gaussian tail estimate:* $\forall\, \lambda > 0$, $\mathbb{P}(|\xi| \geq \lambda) \leq 2\exp\left\{\frac{-\lambda^2}{2b^2}\right\}$.
3. *Orlicz condition:* $\mathbb{E}\left(\exp\left\{\frac{3}{b^2}\xi^2\right)\right\} \leq 2$.
4. *Moments condition:* $\exists\, C > 0$, $\left(\mathbb{E}\left(|\xi|^p\right)\right)^{1/p} \leq Cb\sqrt{p}$ *for all* $p \geq 1$.

Bibliography

[1] Bai, Z., Methodologies in spectral analysis of large-dimensional random matrices, a review, *Statistica Sinica* **9** (1999), 611–677.

[2] Bai, Z., and Silverstein, J., *Spectral Analysis of Large Dimensional Random Matrices*. Mathematics Monograph Series 2, 2006, Science Press, Beijing.

[3] Barrera, G., and Manrique, P., Minimum singular value of circulant random matrices using random polynomials, arXiv:1610.05589 [math], 2016.

[4] Bharucha-Reid, A., and Sambandham, M., *Random Polynomials*, Academic Press, Orlando, FL, 1986.

[5] Bourgain, J., Vu, V., and Wood, P., On the singularity probability of discrete random matrices, *Journal of Functional Analysis* **258** (2010), 559–603.

[6] Bose, A., Subhra, R. & Saha, K., Spectral norm of circulant type matrices with a heavy tailed entries, *Electronic Communications in Probability* **15** (2010), 299–313.

[7] Bose, A., Subhra, R. & Saha, K., Spectral Norm of Circulant-Type Matrices, *Journal of Theoretical Probability* **24** (2011), 479–516.

[8] Bollobás, B., *Random Graphs*, Academic Press, New York, 1985.

[9] Bordenave, C., and Chafaï, D., Around the circular law, *Probability Surveys* **9** (2012), 1–89.

[10] Cook, N., Lower bounds for the smallest singular value of structured random matrices, arXiv:1608.07347 [math], 2016.

[11] Costello, K., Tao, T., and Vu, V., Random symmetric matrices are almost surely non-singular, *Duke Mathematical Journal* **135** (2006), 395–413.

[12] Costello, K., and Vu, V., The rank of random graphs, *Random Structures and Algorithms* **33** (2008), 269–285.

[13] Costello, K., and Vu, V., On the rank of random sparse matrices, *Combinatorics, Probability and Computing* **19** (2010), 321-342.

[14] Dai, A., Kim, C., and Kim, J., Invertibility Probability of Binary Matrices `http://www.yau-awards.org/english/o/invertibility%20probability%20of%20binary%20matrices.pdf`.

[15] Doeblin, W., Sur les sommes d'un grand nombre de variables aléatoires independantes, *Bulletin des Sciences Mathématiques* **63** (1939), 23–64.

[16] Doeblin, W., and Lévy, P., Calcul des probabilités. Sur les sommes de variables aléatoires indéndantes á dispersions bornées inférieurement, *Comptes Rendus de l'Académie des Sciences* **202** (1936), 2027–2029.

[17] Dummit, D., and Foote, R., *Abstract Algebra*, 3rd ed., Wiley India, New Delhi, (2014).

[18] Durrett, R., Oriented Percolation in Two Dimensions, *The Annals of Probability*, **12** (1984), 999–1040.

[19] Eaton, M., *Multivariate Statistics: A Vector Space Approach*, Lecture Notes and Monograph Series, Volume 53, Institute of Mathematical Statistics, 2007.

[20] Erdös, L., Universality of Wigner random matrices: A survey of recent results, *Russian Mathematical Surveys* **66** (2011), 507–626.

[21] Erdös, L., Yan, H., and Yin, J., Bulk universality for generalized Wigner matrices, *Probability Theory and Related Fields* **154** (2012), 341–407.

[22] Erdös, P., On a lemma of Littlewood and Offord, *Bulletin of the American Mathematical Society* **51** (1945), 898–902.

[23] Erdös, P., and Rényi, A., On Random Matrices, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **8** (1964), 455–461.

[24] Esseen, C. G., On the concentration function of a sum of independent random variables, *Z. Wahrscheininlichkeitstheorie verw. Geb.* **9** (1968), 290–308.

[25] Farahmand, K., *Topics in Random Polynomials*, Pitman Research Notes in Mathematics Series, vol. 393, Longman, Harlow, 1998.

[26] Fulman, J., Random matrix theory over finite fields, *Bulletin of the American Mathematical Society* **39** (2002), 51–85.

[27] Ibragimov, I., and Zaporozhets, D., On Distribution of Zeros of Random Polynomials in Complex Plane, in: *Prokhorov and Contemporary Probability Theory*, Proceedings in Mathematics and Statistics, vol. 33. Springer-Verlag, Berlin, 2013, pp. 303–323.

[28] Grimmett, G., *Percolation* (Springer-Verlag, New York), (1991).

[29] Götze, F., and Tikhomirov, A., The circular law for random matrices, *Annals of Probability* **38** (2010), 1444–1491.

[30] Kahane, J., *Some Random Series of Functions*, 2nd Edition, Cambridge University Press, Cambridge, 1985.

[31] Kahn, J., Komlós. J., and Szemerédi, E., On the probability a random ± 1 matrix is singular, *Journal of the American Mathematical Society* **8** (1995), 223–240.

[32] Karapetyan, A., On Minimum Modulus of Trigonometric Polynomials with Random Coefficients, *Mathematical Notes* **61** (1997), 369–373.

[33] Kesten, H., *Percolation Theory for Mathematicians* (Boston: Birkhuser), (1982).

[34] Karapetyan, A., The values of stochastic polynomials in a neighborhood of the unit circle, *Mathematical Notes* **63** (1998), 127–130.

[35] Kesten, H., A sharper form of the Doeblin–Lévy–Kolmogorov–Rogozin inequality for concentration functions, *Mathematica Scandinavica* **25** (1969), 133–144.

[36] Kesten, H., Sums of Independent Random Variables—Without Moment Conditions (The 1971 Riesz Lecture), *Annals of Mathematical Statistics* **43** (1972), 701–732.

[37] Kolmogorov, A., Two uniform limit theorems for sums of independent random variables, *Theory of Probability and Its Applications* **1** (1956), 384–394.

[38] Kolmogorov, A., Sur les propri'et'es des fonctions de concentration de M. P. Lévy, *Annales de I'Institut Henri Poincaré* **16** (1968-60), 27–34.

[39] Komlós, J., On the Determinant of $(0, 1)$ Matrices, *Studia Scientiarum Mathematicarum Hungarica* **2** (1967), 7–21.

[40] Komlós, J., On the determinant of random matrices, *Studia Scientiarum Mathematicarum Hungarica* **3** (1968), 387–399.

[41] Konyagin, S., Minimum of the Absolute Value of Random Trigonometric Polynomials with Coefficients ± 1, *Mathematical Notes* **56** (1994), 931–947.

[42] Konyagin, S., and Schlang, W., Lower Bounds for the Absolute Value of Random Polynomials on a Neighborhood of the Unit Circle, *Transactions of the American Mathematical Society* **351** (1999), 4963–4980.

[43] Lin, Z., and Bai, Z., *Probability Inequalities*, Springer-Verlag, Berlin, 2011.

[44] Littlewood, J., and Offord, A., On the number of real roots of a random algebraic equation III, *Matematicheskii Sbornik* **12** (1943), 277–286.

[45] Manrique, P., Pérez-Abreu, V., and Roy, R., On the universality of the non-singularity of general Ginibre and Wigner random matrices, *Random Matrices: Theory and Applications* **5** 1 (2016), 1650002.

[46] Maples, K., Singularity of Random Matrices over Finite Fields, arXiv:1012.2372 [math], 2010.

[47] Maples, K., Symmetric Random Matrices over Finite Fields, Announcement, `http://user.math.uzh.ch/maples/maples.symma.pdf`, 2013.

[48] Meckes, M., On the spectral norm of a random Toeplitz matrix, *Electronic Journal of Probability* **12** (2007) 315–325.

[49] Meckes, M., Some Results on Random Circulant Matrices, in: *High Dimensional Probability: The Luminy V*, Institute of Mathematical Statistics (2009), pp. 213–223.

[50] Nguyen, H., On the singularity of random combinatorial matrices, *SIAM Journal on Discrete Mathematics* **27** 1 (2013), 447–458.

[51] Nguyen, H., and Vu, V., Small Ball Probability, Inverse Theorems, and Applications, in: *Erdös Centennial*, Springer-Verlag, Berlin. *Bolyai Society Mathematical Studies* **25** (2014), 409–463.

[52] Nguyen, H., and Vu, V., Random Matrices: Law of the Determinant, *Annals of Probability* **42** 1 (2014), 146–167.

[53] Pan, V. Y., Svadlenka, J., and Zhao, L., Estimating the Norms of Circulant Toeplitz Random Matrices and Their Inverses, *Linear Algebra and Its Applications* **468** (2015), 197–210.

[54] Petrov, V., *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*, Clarendon Press, Oxford 1995.

[55] Rahman, Q., and Schmeisser, G., *Analytic Theory of Polynomials: Critical Points, Zeros and Extremal Properties*, Oxford, 2002.

[56] Rivasplata, O., Subgaussian random variables: An expository note, `http://www.stat.cmu.edu/~arinaldo/36788/subgaussians.pdf`, unpublished.

[57] Rogozin, B. A., An estimate for concentration functions, *Theory of Probability and Its Applications* **6** (1961), 94–97.

[58] Roy, R., Sarkar, A., White, D., Backbends in directed percolation, *Journal of Statiscal Physics*, **91** (1998), 889–908.

[59] Rudelson, M., and Vershynin, R., The Littlewood–Offord problem and invertibility of random matrices, *Advances in Mathematics* **218** (2008), 600–633.

[60] Rudelson, M. & Vershynin, R., Non-Asymptotic Theory of Random Matrices: Extreme Singular Values, *Proceedings of the International Congress of Mathematicians Hyderabad Volumen III, India, Editor: Rajendra Bhatia*, 2010, 1576-1602.

[61] Rudelson, M., and Vershynin, R., The Littlewood–Offord Problem and Invertibility of Random Matrices, *Advances in Mathematics* **218** (2008), 600–633.

[62] Sen, A., and Virág, B., The top eigenvalue of the random Toeplitz matrix and the sine kernel, *Annals of Probability* **41** (2013), 4050–4079.

[63] Shmerling, E., and Hochberg, K., Asymptotic behavior of roots of random polynomial equations, *Proceedings of the American Mathematical Society* **130** 9 (2002), 2761–2770.

[64] Spencer, T., Random banded and sparse matrices, in: *The Oxford Handbook on Random Matrix Theory*, Oxford University Press (2011), pp. 471–488.

[65] Slinko, A., A generalization of Komlós theorem on random matrices, *New Zealand Journal of Mathematics* **30** (2001), 81–86.

[66] Spielman, D., and Teng, S., Smoothed analysis of algorithms, *Proceedings of the International Congress of Mathematicians*, Vol. I (Beijing), Higher Ed. Press, (2002), pp. 597–606.

[67] Spielman, D., and Teng, S., Smoothed analysis: Motivation and discrete models, algorithms and data structures, *Lecture Notes in Computer Science* **2748** (2003), pp. 256–270.

[68] Strang, G., *Linear Algebra and Its Applications*, 4th ed., Brooks/Cole, Belmont, MA (2006), p. 65.

[69] Tao, T., and Vu, V., On random ± 1 matrices: Singularity and determinant, *Random Structures and Algorithms Journal* **28** (2006), 1–23.

[70] Tao, T., and Vu, V., On the singularity of random Bernoulli matrices, *Journal of the American Mathematical Society* **20** (2007), 603–628.

[71] Tao, T., and Vu, V., Inverse Littlewood–Offord theorems and the condition number of random matrices, *Annals of Mathematics* **2** (2009), 595–632.

[72] Tao, T., and Vu, V., *Additive Combinatorics*, Cambridge University Press, 2006.

[73] Tikhomirov, K., The smallest singular value of random rectangular matrices with no moment assumptions on entries, *Israel Journal of Mathematics* **212** (2016), 289–314.

[74] Vershynin, R., *Introduction to the Non-Asymptotic Analysis of Random Matrices, Compressed Sensing*, Cambridge University Press, 2012, pp. 210–268.

[75] Vershynin, R., Invertibility of symmetric random matrices, *Random Structures and Algorithms* **44** 2 (2014), 135–182.

[76] Yilmaz, N., Yazlik, Y., and Taskara, N., On the g-Circulant matrix involving the generalized k-Horadam numbers, Preprint, arxiv:1503.00511, 2016.