On the Real Roots and Real Eigenvalues of the Generalized Large Box Model for Random Polynomials and Random Matrices

by

Isabelle Kraus

B.S., University of Central Florida, 2017

M.A., University of Colorado Boulder, 2021

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

2025

Committee Members:

Sean O'Rourke, Chair

Kyle Luh

Judith Packer

Sergei Kuznetsov

Dorin Dutkay

Kraus, Isabelle (Ph.D., Mathematics)

On the Real Roots and Real Eigenvalues of the Generalized Large Box Model for Random Polynomials and Random Matrices

Thesis directed by Professor Sean O'Rourke

Let n and N be in $\mathbb{N}$, and for $0 \leq i \leq n-1$, let $\alpha_i < \beta_i \in \mathbb{R}$. Consider the monic polynomial in a single complex variable of the form $f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ whose coefficients a_i are uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ for each $0 \leq i \leq n-1$ and jointly independent. This random polynomial model is referred to as the **generalized large box model**. When instead $\alpha_{ij} < \beta_{ij}$ for $1 \leq i, j \leq n$ and one considers the n-by-n random matrix whose entries are uniformly distributed on $[\alpha_{ij} N, \beta_{ij} N] \cap \mathbb{Z}$ for each $1 \leq i, j \leq n$ and jointly independent, we say that the matrix is drawn from the **generalized large box model ensemble**.

This thesis is organized into five chapters. Chapter 1 develops and presents the history of the relevant random polynomial and random matrix models, related results, and notation.

Chapters 2 and 3 are concerned with finding the probability that random polynomials whose coefficients obey the generalized large box model have all real roots, as $N \to \infty$. Specifically, in Chapter 2, discriminant and root analysis methods are applied to low degree polynomials, obtaining explicit answers. These methods further find an extremely dominant root, denoted by ξ_n, for all degrees; this is a root whose modulus is not tight as $N \to \infty$, while the moduli of the remaining roots are tight as $N \to \infty$. This expands upon on a discovery made by Dubickas and Sha [Exp. Math., 24(3):312–325, 2015]. As $N \to \infty$, we show that ξ_n is real with probability tending to 1 and that $|\xi_n + a_{n-1}|$ converges in distribution to $|X/Y|$, where X is uniformly distributed on $[\alpha_{n-2}, \beta_{n-2}]$, Y is uniformly distributed on $[\alpha_{n-1}, \beta_{n-1}]$, and X and Y are independent.

In Chapter 3, we consider non-monic degree $n-1$ polynomials whose coefficients are uniformly distributed on $[\alpha_i, \beta_i]$ for $0 \leq i \leq n-1$ and jointly independent, referred to as **generalized bounded height model** polynomials. As $N \to \infty$, the probability that the degree n generalized

large box model polynomial with coefficients uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ for $0 \leq i \leq n-1$ and jointly independent has all real roots converges to the probability that the degree $n-1$ generalized bounded height model polynomial with coefficients uniformly distributed on $[\alpha_i, \beta_i]$ for $0 \leq i \leq n-1$ and jointly independent has all real roots. The methods of Bertók, Hajdu, and Pethö [J. Number Theory, 179:172–184, 2017] are used to express this probability in terms of an integral formula. For the special case when $\alpha_i = 0$ and $\beta_i = 1$ for $0 \leq i \leq n-1$, a relation to the Selberg integral is explored and we show that the probability of such a polynomial having all real roots is positive and monotonically decreasing in n.

In Chapters 4 and 5, the analogous question of the probability that the random matrix whose entries are drawn from the generalized large box model ensemble has all real eigenvalues is considered. In Chapter 4 we begin by letting $\alpha_{ij} < \beta_{ij}$ for $1 \leq i, j \leq n$ and consider the n-by-n random matrix whose entries are uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$ for each $1 \leq i, j \leq n$ and jointly independent. This matrix ensemble is referred to as the **generalized bounded height ensemble**. Using Edelman's method [J. Multivariate Anal., 60(2):203–232, 1997], we factor these matrices into their real Schur decomposition and present an integral formula for the probability that generalized bounded height ensemble matrices have all real roots. In Chapter 5, we show that if A is an n-by-n random matrix with entries that are uniformly distributed on $[\alpha_{ij} N, \beta_{ij} N] \cap \mathbb{Z}$ for $1 \leq i, j \leq n$ and jointly independent and B is an n-by-n random matrix with entries that are uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$ for $1 \leq i, j \leq n$ and jointly independent, then for each $0 \leq k \leq n$, as $N \to \infty$, the probability that A has exactly k real eigenvalues converges to the probability that B has exactly k real eigenvalues. Moreover, the empirical spectral measure of A/N converges weakly in distribution to the empirical spectral measure of B and the joint distribution of the eigenvalues of A/N converges in distribution to the joint distribution of the eigenvalues of B.

Finally, still in Chapter 5, we consider rank one perturbations of the random matrix A whose entries are all independently and identically (iid) uniformly distributed on $[-N, N] \cap \mathbb{Z}$. We say that A is drawn from the **large box model ensemble**. Letting P be the matrix whose entries are all μ_N, we consider the limiting spectral behavior of $A+P$ in the three cases where $\mu_N/N \to \infty$, $\mu_N/N \to 0$, and $\mu_N/N \to c$, for $c \in \mathbb{R}$. If $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, the eigenvalues of $\frac{A+P}{\mu_N}$ converge almost surely to the eigenvalues of P/μ_N, and the largest eigenvalue (in magnitude) is real with probability tending to 1. Additionally, the centered and correctly normalized largest eigenvalue of $A+P$ converges in distribution to a Bates distribution with n^2 parameters. If $\lim_{N\to\infty} \frac{\mu_N}{N} = 0$, the empirical spectral measures of $\frac{A+P}{N}$ and $\frac{A}{N}$ both converge weakly in distribution to the empirical spectral measure of the matrix whose entries are iid and uniformly distributed on $[-1,1]$; the joint distribution of the eigenvalues of $\frac{A}{N}$ and $\frac{A+P}{N}$ both converge in distribution to the joint distribution of eigenvalues of the same random matrix. If $\lim_{N\to\infty} \frac{\mu_N}{N} = c > 0$, the empirical spectral measure of $\frac{A+P}{N}$ converges weakly in distribution to the empirical spectral measure of the matrix whose entries are iid and uniformly distributed on $[-1+c, 1+c]$. In addition, if $c > 12$, we show that as $N \to \infty$, $\frac{A+P}{N}$ has exactly one real outlier eigenvalue and provide a range for its location.

Dedication

To my mom, for being the rain that watered the seeds. To Clint, for being my rock. To Luke, for being the light that brightened the journey.

Acknowledgements

I would like to express my wholehearted gratitude to my advisor, Sean O'Rourke. Thank you for your continuous kindness, support, and patience. Thank you for always providing exactly the right amount of help, while making sure to push me to fight and understand all of the mathematical details deeply. I have learned so much and will always be thankful.

Thank you to Kyle Luh, Judith Packer, Sergei Kuznetsov, and Dorin Dutkay, for serving on the committee and for everything they have ever taught me. I am especially grateful to Dorin, my undergraduate advisor, for opening the door to the world of math research.

Thank you to Kellie Geldreich, for her advocacy and assistance.

Thank you to Andrew Campbell, for his eagerness to discuss (even when halfway across the world), for his advice, and for his friendship.

Thank you to the unknown preschool teacher, who took three-year-old me aside and taught me how to multiply during recess. I think that was the day my love and excitement for math really started. You changed the course of my life, and I wish I could find you.

Contents

Chapter

Figures

Figure

Chapter 1

Introduction

In this thesis, we study random polynomials and random matrices whose entries are jointly independent and uniformly distributed on specific sets of integers, as the cardinality of these sets tend to infinity. In Chapter 1, we present the history of various related random polynomial and random matrix models, including relevant results, and notation. In Chapter 2, we consider the real roots of generalized large box model polynomials. In Chapter 3, we consider instead generalized bounded height model polynomials, allowing us to obtain integral formula answers about the probability that generalized large box model polynomials have all real roots. In Chapter 4, we answer the analogous question for random matrices whose entries are chosen from the generalized bounded height ensemble. In Chapter 5, we apply these results to random matrices whose entries are chosen from the generalized large box model and also investigate rank one perturbations of these matrices.

We consider polynomials with independent and random coefficients. Formally, that is a polynomial of the form

$$f(z) = \sum_{i=0}^{n} c_i \xi_i z^n \tag{1.1}$$

where the $c_i \in \mathbb{R}$ are deterministic constants and the ξ_i are jointly independent real random variables. When $c_i = 1$ for each $0 \leq i \leq n$, these types of polynomials are referred to as a **Kac polynomials**. We are primarily interested in answering questions about integral polynomials, though polynomials with continuous coefficients need to be considered in some arguments. There are generally two types of asymptotic questions that can be asked about Kac polynomials; one can either restrict the coefficients and allow the degree of the polynomials to tend to infinity (known as the **restricted coefficient model**), or one can fix the degree of the polynomial and allow the support of the coefficients $\xi_0, \ldots, \xi_n$ to depend on a parameter N, which tends to infinity (known as the **large box model**). We focus mainly on the second model. For $n, N \in \mathbb{N}$ and letting $\alpha_i < \beta_i \in \mathbb{R}$ for each $0 \leq i \leq n-1$, we consider the polynomial in a single complex variable

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

where each $0 \leq i \leq n-1$, a_i is uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ and $a_0, \ldots, a_{n-1}$ are jointly independent, and ask about that probability that $f(z)$ has all real roots as $N \to \infty$.

We give now a history of the large box model and other results relating to Kac polynomials. This history shows that for the large box model, authors were generally concerned with results pertaining to the reducibility of the polynomials over $\mathbb{Q}$. The following definitions can be found in standard algebra textbooks, see for instance [47]. A polynomial of degree $n \geq 1$ with coefficients in $\mathbb{Q}$ is **reducible** over $\mathbb{Q}$ if it can be factored into two non-constant polynomials of strictly smaller degree, each with coefficients still in $\mathbb{Q}$. If a polynomial is not reducible, we say that it is **irreducible**. The **splitting field** of a polynomial f is the smallest field containing $\mathbb{Q}$ that also contains all of the roots of f. If F is the splitting field of a polynomial f, let G_f be the group of automorphisms of F which leave $\mathbb{Q}$ fixed. Then G_f is the **Galois group** of the polynomial f.

A few non-reducibility results are discussed as well. When the degree of the Kac polynomials tends to infinity instead, there is a rich history of authors bounding the expected number of real roots. Our results lie somewhere in the middle of these theoretical paradigms. While we are mostly interested in the probability that (generalized) large box polynomials have all real roots, the degrees of our polynomials remain fixed.

Asymptotic notation is used throughout the thesis, so we mention it now. Asymptotic notation may be used under the assumption that $N \to \infty$ or $n \to \infty$, and should be clear from the context. For functions $f(x)$ and $g(x)$, we say that $f(x) = O(g(x))$ if there exists some positive constant M such that $|f(x)| \leq M|g(x)|$ for all x sufficiently large; $f(x) = \Omega(g(x))$ if there exists some positive constant M such that $|f(x)| \geq M|g(x)|$ for all x sufficiently large. Similarly, we say that

$f(x) \ll g(x)$ if $f(x) = O(g(x))$, and $g(x) \ll f(x)$ if $f(x) = \Omega(g(x))$. We say that $f(x) = o(g(x))$ if for all $\epsilon > 0$, $|f(x)| < \epsilon|g(x)|$ for all x sufficiently large. To indicate that the constants M or ϵ depend on another constant, subscripts are used, e.g., $f(x) = O_n(g(x))$. The subscript may be omitted if the context is clear. For example, $f(N) = O(1/N)$ means that $|f(N)| \leq M_1/N$ for some positive constant M_1 for all N sufficiently large and $f(n) = O(\sqrt{n})$ means that $|f(n)| \leq M_2\sqrt{n}$ for some positive constant M_2 for all n sufficiently large.

1.1 The large box model

In 1934, van der Waerden [164] considered integer polynomials of fixed degree

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

where the coefficients $a_0, \ldots, a_n$ lie inside the *Schranke* $[-N, N]$. He proved, without relying on Hilbert's irreducibility theorem, that if one forms all possible equations $f(z) = 0$, then the proportion of these equations whose Galois group is the symmetric group of n elements, called S_n, tends to 1 as $N \to \infty$. In 1936, van der Waerden [165] wanted to determine how quickly the proportion of polynomials $f(z)$ whose Galois group is not S_n (called polynomials *mit Affekt*) tends to zero as $N \to \infty$, and proved that the ratio of polynomials *mit Affekt* to all polynomials with coefficients in $[-N, N]$ is no more than $N^{-\frac{1}{6(n-2)\ln\ln N}}$. He went on to say that he "suspects that the frequency of polynomials *mit Affekt* is essentially the same as the frequency of the reducible polynomials, even for $n \geq 2$." For non-monic polynomials, this conjecture is that the number of polynomials *mit Affekt* is $O(N^n)$, or bounded above by CN^n for some constant $C > 0$ and all N sufficiently large.

Authors following van der Waerden considered the monic polynomial analog to his question. This model of polynomials became known as the large box model or bounded degree model.

Definition 1.1.1 (Large Box Model)**.** Let $N \in \mathbb{N}$ and let $a_0, \ldots, a_{n-1}$ be independently and identically distributed on $[-N, N]$. Then the monic integral polynomial

$$f(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

is a random polynomial coming from the **large box model** ensemble.

Advances towards van der Waerden's conjecture were made by various authors [4, 32, 39, 68, 96, 97, 139, 175], and the case for degrees $n \leq 4$ was proven true by Chow and Dietmann in 2012 [33]. In 2023, the conjecture was finally proven to be true for all degrees n by Bhargava [19], who stated that his method applies to both monic and non-monic polynomials.

Aside from just counting the number of polynomials *mit Affekt*, there are several papers that pertain to the reducibility of polynomials whose coefficients are given by the large box model. We now give a brief history.

1.1.1 Reduciblity large box model results

Many authors following van der Waerden were similarly interested in reducibility of polynomials whose coefficients are described by the large box model. All reducibility of polynomials in this section is considered to be over $\mathbb{Q}$, unless stated otherwise. In 1963, Chela [31] showed that for monic large box model polynomials of degree n, letting $\rho(n, N)$ denote the total number of reducible polynomials with $n > 2$,

$$\lim_{N \to \infty} \frac{\rho(n, N)}{N^{n-1}} = 2^n \left(\zeta(n-1) - \frac{1}{2} + \frac{k_n}{2^{n-1}} \right), \tag{1.2}$$

where $\zeta(z)$ is the Riemann zeta function in the complex variable z and $k_n = \int_R dx_1 \dots dx_{n-1}$ where the region R is given by

$$R = \left\{ (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} \;\middle|\; |x_i| \leq 1 \text{ for } 1 \leq i \leq n-1 \text{ and } \left| \sum_{i=1}^{n-1} x_i \right| \leq 1 \right\}.$$

When $n = 2$, the limit above is not valid. Letting $\tilde{\rho}(n, N)$ denote the number of monic large box model polynomials whose coefficients $a_0, \dots, a_{n-1}$ also satisfy $\sum_{i=0}^{n-1} a_i^2 < N^2$, Chela shows that

$$\lim_{N \to \infty} \frac{\rho(2, N)}{\tilde{\rho}(2, N)} = 1.$$

From here, by using an asymptotic formula established by Specht [148] for $\tilde{\rho}(n, N)$, Chela concludes that

$$\lim_{N \to \infty} \frac{\rho(2, N)}{2N \log N} = 1$$

as well.

In 1973, Gallagher [68] showed that the number of monic large box model polynomials of degree n whose Galois group is not S_n is asymptotically $N^{n-1/2}\log^{1-\gamma} N$, with $\gamma = \gamma_n > 0$ and the asymptotic constant depending only on the degree of the polynomial. In 1974, while investigating Hilbert's irreducibility theorem, Fried [66] shows that the number of reducible large box model polynomials of degree n which have $n-1$ fixed coefficients is $O(N^{1/2})$, where the asymptotic constant does not depend on the fixed coefficients. In 1979, Cohen [34] considered a more general restricted large box model, where $n-s$ coefficients of the monic degree n polynomial $f(z)$ are fixed. Excluding the exceptional cases where the constant term is zero or where all of the polynomials are members of $\mathbb{Z}[X^r]$ for some $r > 1$, Cohen shows that when $s \geq 2$, the number of polynomials whose Galois group is not S_n is $O(N^{s-1/2}\log N)$. When $s = 1$, Cohen shows that the number of polynomials whose Galois group is not S_n is $O(N^{-1/2}\log N)$, where the asymptotic constant does depend on the fixed coefficient of the polynomial, though the justification of this case relies upon the generalized Riemann hypothesis.

In 2009, Kuba [101] considered non-monic reducible large box model polynomials and gave several asymptotic results as $N \to \infty$. Let $p^*(n,N)$ denote the total number of degree n non-monic reducible large box model polynomials with height at most N. Kuba first demonstrated the correct order of magnitude of $p^*(n,N)$, improving the result over those of Polya and Szegö [132] and Dörge [43]. Kuba's result is summarized in the next theorem.

Theorem 1.1.2 (Theorems 1 and 2 in [101])**.** *As $N \to \infty$,*

$$N^2 \log N \ll p^*(2,N) \ll N^2 \log N.$$

For every $n \geq 3$, there is a constant $C_n > 0$ such that

$$N^n \leq p^*(n,N) \leq C_n N^n \quad \textit{for all } N \geq 1.$$

Moreover, letting $p_s^*(n,N)$ denote the total number of degree n polynomials $f(z) = a_n z^n + \cdots + a_1 z + a_0$ over $\mathbb{Z}$ with height $H(f) := \max\{|a_i| \mid i = 0,1,\ldots,n\}$ at most N that contribute to

$p^*(n,N)$ and split completely into linear factors over $\mathbb{Z}$, Kuba shows next that splitting completely is highly unlikely for the non-monic large box model.

Theorem 1.1.3 (Theorem 3 in [101])**.** *For every fixed* $n \geq 2$*,*

$$N^2 (\log N)^{n-1} \ll p_s^*(n,N) \ll N^2 (\log N)^{n-1}.$$

Kuba also shows the following.

Theorem 1.1.4 (Theorem 4 in [101])**.** *For* $n/2 < k < n$*, let* $p_k^*(n,N)$ *denote the number of degree* n *polynomials of height at most* N *in* $\mathbb{Z}[X]$ *that have an irreducible factor of degree* k *in* $\mathbb{Z}[X]$*. Then for* $1 < n/2 < k < n$*,*

$$N^{k+1} \ll p_k^*(n,N) \ll N^{k+1}.$$

This theorem shows that for $n \geq 3$, one has that $N^n \ll p_{n-1}^*(n,N) \ll N^n$; comparing with Theorem 1.1.2, we see that almost all elements that contribute to $p^*(n,N)$ are polynomials that split into an irreducible degree $n-1$ factor and a linear factor.

In 2018, Borst et al. [26] formulated a universality heuristic for integral polynomials, claiming that certain reducibility properties of the polynomials do not depend on the exact distribution of the coefficients. Their heuristic is the following.

Heuristic 1.1.5 (Universality, see [26]). Let $f(x)$ be a random polynomial with sufficiently well-behaved integer coefficients. Then the probability that $f(x)$ is reducible over the rationals divided by the probability that f has a linear (or lowest possible degree) factor approaches a constant C in the limit as the degree goes to infinity, or in the limit as the support of the random coefficients goes to infinity, or both. Furthermore, in the limit as the degree goes to infinity, whether or not the support goes to infinity, the constant C should equal 1.

For the case of the large box model, by rescaling Chela's result (1.2) and dividing by the probability that the constant coefficient is zero, the authors are able make a connection to their heuristic. They show that

$$\lim_{N\to\infty}\frac{\mathbb{P}(f(z)\text{ is reducible })}{\mathbb{P}(a_0=0)}=\lim_{N\to\infty}\frac{\left(\frac{p(n,N)}{(2N+1)^n}\right)}{1/(2N+1)}=\lim_{N\to\infty}\frac{p(n,N)}{(2N+1)^{n-1}}=\frac{2^n\left(\zeta(n-1)-\frac{1}{2}+\frac{k_n}{2^{n-1}}\right)}{2^{n-1}}$$
$$=2\zeta(n-1)-1+\frac{k_n}{2^{n-2}},$$

where we have added in the details of the application of (1.2).

Similarly, for non-monic large box model polynomials, the authors reinterpret Theorem 1.1.2 from Kuba, to get that for $n\geq 3$,

$$\frac{\mathbb{P}(f(z)\text{ is reducible })}{\mathbb{P}(a_0=0)}=\frac{\left(\frac{p^*(n,N)}{(2N+1)^{n+1}}\right)}{1/(2N+1)}\leq\frac{C_n}{2^n}\qquad\text{for all }N\geq 1,$$

where C_n is the constant appearing in Theorem 1.1.2.

For other fixed degree models, such as monic polynomials with iid coefficients uniform in $[0,N]$ or non-monic polynomials with binomial coefficients from the interval $[0,N]$, the authors show images from computer simulations that suggest that the heuristic holds. This is also the case for the growing degree monic zero-one polynomials or Rademacher ± 1 polynomials.

In 2019, O'Rourke and Wood [126] were also interested in whether the irreducibility of integral polynomials is a universal phenomenon that occurs across different coefficient models. They consider monic integral polynomials, and show that such a polynomial f has an irreducible factor of degree k over $\mathbb{Q}$ if and only if f has a root that is an algebraic number of the same degree. Their main theorem is the following bound on the probability of algebraic roots of degree at most k for monic integral polynomials.

Theorem 1.1.6 (Theorem 1.7 in [126])**.** *Let f be a degree n random monic polynomial with integer coefficients. Let $M>0$ and $2\leq k\leq n$. Take $\Omega\subseteq\{z\in\mathbb{C}\mid |z|\leq M\}$, and suppose there exists $p\in[0,1]$ such that*

$$\sup_{z\in\Omega}\mathbb{P}(f(z)=0)\leq p.$$

Then, the probability that f has an algebraic root of degree at most k in Ω is at most

$$p(eM)^{k^2}+\mathbb{P}(|\xi_i(f)|>M\textit{ for some }i),$$

where $\xi_1(f), \ldots, \xi_n(f)$ *are the roots of* f*. If* $k = 1$*, the result holds if* $p(eM)^{k^2}$ *is replaced wth* $p(3M)$*.*

They also prove the following.

Theorem 1.1.7 (Theorem 2.1 in [126])**.** *For each* $n \geq 1$*, let* $f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$*, where* $a_0, a_1, \ldots$ *are iid Rademacher random variables, which take the values* $+1$ *or* -1 *with equal probability. Then the probability that* $f_n(z)$ *has an algebraic root of degree at most* $\frac{n^{1/3}}{\log^3(n)}$ *is at most* $O\left(\frac{1}{\sqrt{n}}\right)$*.*

The authors claim that their approach also works for other distributions on the coefficients, as long as they are still iid, integer valued, and satisfy certain additional (but unspecified) assumptions.

In 2021, Pham and Xu [131] were able to show that other polynomial models, such as ones with coefficients supported in $[-N, N]$ but not necessarily uniformly or independently, are also irreducible with probability tending to 1 as $N \to \infty$. See Theorem 1.1 in [131] for complete details.

In 2023, Anderson et al. [4] give upper bounds for both monic and non-monic degree n polynomials whose Galois groups are contained in the alternating group A_n, the group of even permutations of a set of n elements. These results are given in asymptotic terms depending on the height of the polynomial.

In 2024, Bary-Soroker, Ben-Porath, and Matei [11] considered the probability that random polynomials with coefficients coming from the large box model have Galois group A_n. Bhargava [19, 20] recently showed that that

$$\mathbb{P}(G_f = A_n) = O(N^{-1}),$$

and the authors of this recent paper conjecture that for $\epsilon > 0$, the stronger bound

$$\mathbb{P}(G_f = A_n) = O(N^{-n/2+\epsilon})$$

holds. They prove the following two main theorems.

Theorem 1.1.8 (Theorem 1.2 in [11])**.** *Let* $n \geq 4$ *and put* $k = \frac{1}{2}\lfloor\frac{n+1}{2}\rfloor$*. Let* f *be monic polynomial of degree* n *drawn from the large box model. Then* $\mathbb{P}(G_f = A_n) = O(N^{k-n})$*.*

Theorem 1.1.9 (Theorem 1.3 in [11]). *Let $n \geq 6$ be even and let f be monic polynomial of degree n drawn from the large box model. Then the probability that the discriminant of f is a square is at least $\Omega\left(N^{-\frac{n+1}{2}}\right)$.*

Theorem 1.1.9 is relevant to their question because the probability of the discriminant of a separable polynomial being a square provides an upper bound for the probability that its Galois group is A_n, which can be seen in the discussion on page 2 of [11].

1.1.2 Other large box model results

We now review results for large box model polynomials that are not concerned with the reducibility of the polynomials. Instead, these results are more concerned with properties of the roots of the polynomials and are more closely related to the results of this thesis.

In 2015, Dubickas and Sha [45] define a dominant polynomial as a polynomial having one root whose modulus is greater than the moduli of all remaining roots. Letting $A_n(N)$ denote the number of dominant monic integer polynomials of degree n and height at most N, they show that

$$\lim_{N\to\infty} \frac{A_n(N)}{(2N)^n} = 1.$$

The authors state that they are surprised by this result, because they would expect a random polynomial (at least with iid Gaussian coefficients) to have a large number of non-real roots, which of course occur in complex pairs. In Chapter 2, we are able to explicitly find this dominant root for generalized large box model polynomials. In fact, we call this root extremely dominant, we show that its modulus is not tight as $N \to \infty$, while the moduli of all other roots of generalized large box model polynomials are tight as $N \to \infty$. We also determine the fluctuations of the extremely dominant root.

Letting $A_n^*(N)$ denote the number of dominant non-monic integral polynomials of degree n and height at most N, Dubickas and Sha are only able to obtain an explicit formula for $A_2^*(N)$, which is given by

$$\lim_{N\to\infty} A_2^*(N)/(2N)^3 = \frac{41 + 6\log 2}{72} \approx 0.6262.$$

They conjecture that the limit

$$\lim_{N\to\infty} A_n^*(N)/(2N)^{n+1}$$

does not exist, but are able to provide upper and lower bounds which show that the proportion of dominant polynomials is strictly in $(0,1)$. They also prove the following.

Theorem 1.1.10 (Theorem 3.2 in [45]). *Let $H(f) = \max_{0\le j\le n} |a_j|$. Let $f(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_0 \in \mathbb{C}[z]$. If $|a_{n-1}| > n(n+1)^{1/4}|a_n|^{1/2}H(f)^{1/2}$, then f is dominant.*

From Theorem 1.1.10, setting $a_n = 1$, we see since as $N \to \infty$

$$|a_{n-1}| \geq n(n+1)^{1/4}N^{1/2}$$

with probability tending to 1, monic integer polynomials with coefficients iid on $[-N, N]$ have a dominant root with probability tending to 1 as $N \to \infty$. The authors prove their results using Sturm's theorem and the Bistriz stability criterion, but we take a different approach and use Rouché's theorem on the extremely dominant root for quadratic polynomials to locate the extremely dominant roots of higher degrees. In a related 2016 paper, Dubickas and Sha [46] show that the proportion of non-monic large box model integer polynomials with exactly one or two dominant roots tends to 1 as $N \to \infty$. Their result is the following.

Theorem 1.1.11 (Theorem 1.1 in [46]). *Assume that the roots $\xi_1, \ldots, \xi_n$ (listed with multiplicities) of a polynomial f are labeled so that $|\xi_1| \geq |\xi_2| \geq \cdots \geq |\xi_n|$. In case $|\xi_1| = \cdots = |\xi_k| > |\xi_{k+1}|$, we say that f has exactly k roots with maximal modulus. Let $A_n^*(k, N)$ denote the cardinality of the set of integer polynomials of degree n and height at most N with exactly k roots of maximal modulus. For any integer $n \geq 2$,*

$$\lim_{N\to\infty} \frac{A_n^*(1,N) + A_n^*(2,N)}{2N(2N+1)^n} = 1.$$

They also prove the following, where $D_n^*(s, N)$ denotes the number of non-monic integral polynomials f of degree n and height at most N such that f has exactly r real roots and $2s$ non-real roots. We emphasize now that that these are no longer dominant integral polynomials, and simply integral polynomials.

Theorem 1.1.12 (Theorem 1.2 in [46]). *For any integers $n, N \geq 1$ and any non-negative integers r, s such that $n = r + 2s$, we have*

$$N^{n+1} \ll D_n^*(s, N) \ll N^{n+1}.$$

The authors state their desire for an asymptotic formula in this case, which is answered by Bertók, Hajdu, and Pethő in 2017 [18] with the following theorem. Denote by $\mathcal{H}_n(s, N)$ the set of $(n+1)$-dimensional vectors $(a_0, \ldots, a_n)$ satisfying $|a_i| \leq N$ for $0 \leq i \leq n, a_n \neq 0$, with exactly s pairs of complex conjugate roots. Let λ^n denote the n-th dimensional Lebesgue measure.

Theorem 1.1.13 (Theorem 2.1 in [18]). *We have*

$$D_n^*(s, N) = \lambda^{n+1}(\mathcal{H}_n(s, 1))N^{n+1} + O(N^n).$$

In addition, letting $D_n(s, N)$ denote the number of monic integral polynomials f of degree n and height at most N such that f has exactly r real roots and $2s$ non-real roots, they are able to provide the following bound.

Theorem 1.1.14 (Theorem 2.2 in [18]). *We have*

$$N^n \ll D_n(s, B) \ll N^n$$

where the implied constants depend on n only.

Unfortunately, Theorem 1.1.13 is incorrect for $s = n/2$, and is corrected by Dubickas in a subsequent paper [44]. The correct statement in this case is the following.

Theorem 1.1.15 (Theorem 1.1 in [44]). *Let n be an even integer. Then*

$$N^{n-1/2} \ll D_n(n/2, N) \ll N^{n-1/2},$$

where the implied constants in $\ll$ depend on n only.

Bertók, Hajdu, and Pethő [18] also left open the question of whether the limit

$$\lim_{N \to \infty} \frac{D_n(s, N)}{N^n}$$

exists. Dubickas responds, proving that the limit does exist and finding its value in terms of the n-th dimensional Lebesgue measure of the set $\mathcal{H}_{n-1}(s,1)$.

Theorem 1.1.16 (Theorem 1.2 in [44])**.** *Let* $n \geq 1$ *and let* s *be two integers satisfying* $0 \leq s < n/2$. *Then*

$$\lim_{N\to\infty} \frac{D_n(s,N)}{N^n} = \lambda^n(\mathcal{H}_{n-1}(s,1)).$$

The methods of Bertók, Hajdu, and Pethő were very similar to techniques used by Akiyama and Pethő [2], who considered bounded roots (rather than bounded coefficients) of the polynomials. In Chapter 3, we begin by showing that as $N \to \infty$, the probability that a monic degree n generalized larged box model polynomial, with coefficients uniformly distributed on $[\alpha_j N, \beta_j N] \cap \mathbb{Z}$ for $0 \leq j \leq n-1$ and jointly independent, has all real roots converges to the probability that the non-monic generalized bounded height polynomial of degree $n-1$ with coefficients uniformly distributed on $[\alpha_j, \beta_j]$ for $\alpha_j < \beta_j, 0 \leq j \leq n-1$ has all real roots. We then give integral formulas for the latter probability, generalizing the work of Bertók, Hajdu, and Pethő by using the same techniques. We give some additional results for nice cases of coefficients, such as when the coefficients are all iid on $[0,1]$, or when every coefficient is allowed to take on the value zero.

1.2 Other independent coefficient polynomial models

Waring, in 1782, is credited [162,167] as being the first mathematician to apply probabilistic ideas to answer questions about the zeros of polynomials. He was followed by Sylvester [151] in 1864. Both of these investigations were in response to a proposition presented by Newton (without proof) bounding the minimum number of complex pairs of roots of an algebraic equation, and can be considered the earliest studies of random polynomials. An excellent account of this history is provided by Holgate [78].

These days, there are several models of random polynomials to consider. The most common model involves assigning probability distributions to the coefficients of the polynomial, often independently, and letting the degree n tend towards infinity. The most complete surveys of independent coefficient random polynomial results we have found are provided by Thangaraj and Sambandham [159] and Nguyen, Nguyen, and Vu [122]. We elaborate on some fundamental results here.

The independent coefficient model was popularized in the 1930s by Bloch and Pólya [25], who considered polynomials of the form

$$f(z) = a_n z^n + \cdots + a_1 z + 1$$

where each coefficient a_i takes one of the values in the discrete set $\{-1, 0, 1\}$. They found that the average number of real roots of such a polynomial is at most $O(\sqrt{n})$. They then asked about the maximum number of real zeros in the interval $(0, 1)$ that $f(z)$ can take. Letting P_n denote this maximum, they prove that there exists a positive constant A, such that for $n > 3$,

$$\frac{1}{A}\left(\frac{n^{1/4}}{(\log n)^{1/2}}\right) < P_n < A\left(\frac{n \log\log n}{\log n}\right).$$

In the 1930s and 1940s, Littlewood and Offord [105, 106, 107] also ask about the expected number of real roots of a polynomial with real coefficients. In 1939 [106], they show that for polynomials with iid coefficients uniform on the interval $(-1, 1)$, for $n > 2000$, the probability that such a polynomial has more than $25\log(n)^2$ real roots is at most $(12\log n)/n$ and that the expected number of real roots is at most $25\log(n)^2 + (12\log n)/n$. Littlewood and Offord also show that when the polynomial coefficients are iid on the discrete set $\{-1, 0, 1\}$, the expected number of real roots is bounded above by $O(\sqrt{n \log n})$, markedly improving Bloch and Pólya's earlier result; when the polynomial coefficients are iid on $\{-1, 1\}$ instead, the upper bound $O(\sqrt{n})$ is provided. Other related Littlewood and Offord papers from this time period include [108, 109].

After these initial explorations, authors began considering polynomials of the form

$$\epsilon_n a_n z^n + \cdots + \epsilon_1 a_1 z + a_0,$$

where $a_0, \ldots, a_n$ are fixed complex numbers and $\epsilon_1, \ldots, \epsilon_n$ are the iid random variables which take the values ± 1 with equal probability. Letting $M = |a_0| + \ldots + |a_n|$, Schur [145] first shows in 1933 that such an equation can have at most

$$4\sqrt{n \log\left(\frac{M}{\sqrt{|a_0 a_n|}}\right)}$$

real roots.

Ten years later, Littlewood and Offord [107] show that except a set of size $A(\log\log n)/(\log n)$, where A is an absolute constant, the remaining equations have at most

$$10 \log n \left(\log\left(\frac{M}{\sqrt{|a_0 a_n|}}\right) + 2(\log n)^5\right)$$

real roots.

Beginning in 1943, Kac [88, 89, 90] also considers the real roots of random polynomials. He begins by presenting a formula for the number of real roots of a continuous function in an interval, now known as the Kac-Rice formula.

Theorem 1.2.1 (Lemma 1 in [89])**.** *Let $f(x)$ be a continuous function in $(-\infty, \infty)$ having a continuous first derivative $f'(x)$ and only a finite number of turning points (that is, a point where $f(x)$ changes between being increasing and decreasing) in each finite interval. Let ψ_ϵ be 1 if $-\epsilon < x < \epsilon$ and 0 otherwise. If neither a nor b is a zero of $f(x)$, then for sufficiently small ϵ's,*

$$\frac{1}{2\epsilon}\int_a^b \psi_\epsilon(f(x))|f'(x)|\ dx$$

is equal to the number of zeros of $f(x)$ inside the interval (a, b). (Multiple roots are counted only once.)

Evaluating this formula in the case where the coefficients of $f(z) = a_n z^n + \cdots + a_1 z + a_0$ are iid Gaussians, with mean zero and variance 1/2, Kac computes the average number of real roots as

$$\frac{4}{\pi}\int_0^1 \frac{\left(1 - n^2\left(x^2\left(1-x^2\right)/\left(1-x^{2n}\right)\right)^2\right)^{1/2}}{1-x^2}\ dx,$$

which is asymptotically

$$\frac{2\log n}{\pi} \tag{1.3}$$

and bounded above by

$$\frac{2\log n + 14}{\pi},$$

giving a sharp result. In that paper, Kac stated that the asymptotic bound (1.3) and his proof method were also valid for other iid random variables with variance 1 (not necessarily only standard normals) via the central limit theorem, but this was not quite true. It took him six more years, until 1949 [90], to be able to find a proof for uniformly distributed coefficients on $(-1, 1)$; the proof idea and result is the same as for the normally distributed coefficients, but the actual calculations are much harder. The method also ends up not applying to the discrete $\{-1, 1\}$ case.

In 1950, Erdös and Turán [58] give a bound on the number of roots of a polynomial in an interval.

Theorem 1.2.2 (Theorem 1 in [58]). *If the roots of the polynomial*

$$f(z) = a_n z^n + \cdots + a_1 z + a_0$$

are denoted by

$$z_\nu = r_\nu e^{i\psi_\nu}, \ \nu = 1, 2, \ldots, n,$$

where $\psi_\nu \in [0, 2\pi)$*, then for every* $0 \leq \alpha < \beta \leq 2\pi$*, we have*

$$\left| \sum_{\alpha \leq \psi_\nu \leq \beta} 1 - \frac{\beta - \alpha}{2\pi} n \right| < 16 \sqrt{n \log \frac{|a_n| + \cdots + |a_0|}{\sqrt{|a_n a_0|}}}.$$

In other words, this theorem implies an angular equidistribution of the roots of a polynomial whenever the coefficients are sufficiently nice.

In 1956, Erdös and Offord [59] are finally able to tackle the case when the polynomial coefficients are iid on $\{-1, 1\}$. They improve the earlier result of the second author by proving the following theorem.

Theorem 1.2.3 (Theorem in [59]). *Let* $\epsilon_i, i = 1, 2, \ldots, n$ *be* $+1$ *or* -1 *with equal probability. The number of real roots of most of the equations*

$$f_n(x) = \sum_{i=0}^{n} \epsilon_i x^i = 0$$

is

$$\frac{2\log n}{\pi} + o((\log n)^{2/3}\log\log n),$$

where the exceptional set does not exceed a proportion $o\left((\log\log n)^{-1/2}\right)$ *of the total number of equations.*

Their proof method is to essentially divide the interval $(1/2, 1)$ in a clever way and to count the number of sign changes of the polynomial, which is then used to estimate the number of zeros.

In 1962, Samal [142] considers non-monic degree n polynomials with iid coefficients which have mean zero, finite non-zero variance, and finite non-zero third absolute moment. He provides both asymptotic upper and lower bounds on the number of real roots of such polynomials, along with the sizes of the exceptional sets. His result is stated as follows.

Theorem 1.2.4 (Theorems 1 and 2 in [142]). *Let* $f(z) = \sum_{i=0}^{n} \xi_i z^i$ *be a polynomial such that the* ξ_i *are iid with mean zero, and finite and non-zero variance and third absolute moment. Then, there exists* n_0 *sufficiently large such that for* $n \geq n_0$*, the number of real rots of most of the equations* $f(z) = 0$ *is at most* $\alpha(\log n)^2$*, where* $\alpha > 0$ *is a constant; the measure of the exceptional set tends to zero as n tends to infinity. Let* $\{\epsilon_n\}$ *be a sequence of numbers tending to zero but such that* $\epsilon_n \log n$ *tends to infinity. Then, for* $n \geq n_0$*, the number of real roots of most of the equations* $f(z) = 0$ *is at least* $\epsilon_n \log n$*; the measure of the exceptional set tends to zero as n tends to infinity.*

In 1965, Evans [60] considers the collection of random polynomials

$$f_n(x,t) = \sum_{i=0}^{n} h_i(t)x^i = 0.$$

The coefficients $h_i(t)$ $(0 \leq i \leq n)$ are mutually independent and normally distributed random variables with mean zero and unit variance defined on a common probability space and indexed by the parameter $t \in [0, 1]$. Evans finds both asymptotic upper and lower bounds on the number of real roots, outside of some exceptional sets. Unlike previous similar results, the exceptional sets found by Evans in this case do not depend the degree of the polynomial.

Theorem 1.2.5 (Theorems 1 and 2 in [60]). *Consider the collection of random polynomials $f_n(x,t) = \sum_{i=0}^{n} h_i(t)x^i$. The coefficients $h_i(t)$ $(0 \leq i \leq n)$ are mutually independent and normally distributed random variables with mean zero and unit variance defined on a common probability space and indexed by the parameter $t \in [0,1]$. There exists an integer n_0 and a set E of measure at most*

$$\frac{A}{\log n_0 - \log\log\log n_0}$$

such that, for each $n > n_0$ and all t not belonging to E, the equations $f_n(x,t) = 0$ have at most $\alpha(\log\log n)^2 \log n$ real roots, where α and A are constants. Moreover, there exists an integer n_1 and a set F of measure at most

$$\frac{B \log\log n_1}{\log n_1}$$

such that, for each $n > n_1$ and all t not belonging to F, the equations $f_n(x,t) = 0$ have at least $\frac{\beta \log n}{\log\log n}$ real roots, where β and B are constants.

In 1968, Ibragimov and Maslova [80] show that for non-monic iid real random variables $a_1, \ldots, a_n$ with mean zero and non-zero variance, the expected value of the number of real roots of $f(z) = \sum_{i=1}^{n} a_j z^j$, given that $f(z)$ is non-zero, is asymptotically $\frac{2}{\pi}\log(n)$ as $n \to \infty$.

In 1969, Stevens [150], also using the Kac-Rice formula, was able to generalize the previous results of Kac to many more continuous random variables. This result is the following.

Theorem 1.2.6 (Theorem in [150]). *If in $f(z) = a_n z^n + \cdots + a_1 z + a_0$, the a_i are independent random variables with*

$$E(a_i) = 0, \quad E(a_i^2) = 1, \quad E(a_0^4) < B,$$

$$\frac{d}{dy}\mathbb{P}(a_0 < y) < \frac{B}{1 + y^{1+1/B}}, \quad \frac{d}{dy}\mathbb{P}(a_n < y) < \frac{B}{1 + y^{1+1/B}}$$

for some finite B and for all $y \in \mathbb{R}$, then the average number of real zeros of $f(z)$ is asymptotic to $(2/\pi)\log(n+1)$ as $n \to \infty$.

For the next improvement by Ibragimov and Maslova, we need the following definition.

Definition 1.2.7 (Domain of Attraction, [130])**.** Let $\{X_n\}$ be sequence of independent random variables $\{X_n\}$ all having the same cumulative distribution function $F(x)$. If there exist sequences of constants $\{a_n\}$ and $\{b_n\}$ such that $a_n > 0$, and the cumulative distributions of the sums

$$Z_n = \frac{1}{a_n}\sum_{k=1}^{n} X_k - b_n$$

converge weakly to some cumulative distribution function $G(x)$, then we say that $F(x)$ is **attracted to** $G(x)$. The set of all cumulative distribution functions that are attracted to $G(x)$ is said to make up the **domain of attraction** of $G(x)$.

In 1971, Ibragimov and Maslova [82] show that when the coefficients $a_0, \ldots, a_n$ are iid random variables belonging to the domain of attraction of the normal law, with $\mathbb{E}(a_j) = 0$, the expected number of real zeros of $f(z) = \sum_{i=0}^{n} a_i x^i$ is asymptotically $\frac{2}{\pi}\log(n)$ as $n \to \infty$. Under the same assumptions, but now with $\mathbb{E}(a_j) \neq 0$ instead, the expected number of real zeros is asymptotically $\frac{1}{\pi}\log(n)$ as $n \to \infty$, see [83]. The **characteristic function** of a random variable X is given by $\varphi_X(t) = \mathbb{E}[e^{itX}]$, and determines the cumulative distribution of the random variable, see Section 3.3 and Theorem 3.3.11 in [48]. In a third paper of the same year, Ibragimov and Maslova [81] show that if all of the iid variables $a_0, \ldots, a_n$,with $\mathbb{P}(a_0 = 0) = 0$, have cumulative distribution function F, where F belongs to the domain of attraction of a non-degenerate stable distribution with characteristic function

$$\varphi(t) = e^{\left(it - c|t|^{\alpha}\left(1 - i\beta\frac{t}{|t|}w(t,\alpha)\right)\right)}$$

where $d \in \mathbb{R}$, $c > 0$, $\alpha \in (0, 2], \beta \in [-1, 1]$, and

$$w(t,\alpha) = \begin{cases} \frac{\pi\alpha}{2}, & \text{if } \alpha \neq 1, \\ \frac{2}{\pi}\ln|t|, & \text{if } \alpha = 1, \end{cases}$$

then the expected number of real zeros of $f(z)$ is asymptotically $B\log n$, for some constant B depending on the parameters of F. The reader is referred to two additional related results due to Maslova [116, 117] in 1974.

In 1972, Samal and Mishra [143] consider polynomials of the form $\sum_{i=0}^{n} \xi_i z^i$, where the characteristic function of the iid coefficients ξ_i are given by $e^{-C|t|^\alpha}$, where $\alpha \geq 1$ and $C > 0$ is a constant. They note that when $1 \leq \alpha < 2$, the coefficients of such a polynomial have infinite variance. Samal and Mishra establish that except for a set whose measure tends to zero as $n \to \infty$, the number of real roots of most such polynomials is at least $(\mu \log n)/(\log \log n)$ whenever n is sufficiently large and where $\mu > 0$. In a follow-up 1972 paper [144], under the same hypothesis as before but now restricting $\alpha > 1$, Samal and Mishra show that there exists $n_0 > 0$ such that the exceptional set from their previous paper has measure at most $\frac{\mu'}{(\log((\log n_0)/(\log \log n_0)))^{\alpha-1}}$, where μ' is a positive constant.

In 1988, Wilkins [173] algebraically manipulated the Kac-Rice formula to show that the expected number of real zeros of $f(z) = \sum_{i=1}^{n} a_i z^i$, where $a_0, \ldots, a_n$ are iid standard normals, is given by

$$\frac{2}{\pi} \log(n+1) + \sum_{p=0}^{\infty} A_p (n+1)^{-p},$$

where the rounded values are given by $A_0 = 0.625735818, A_1 = 0, A_2 = -0.24261274, A_3 = 0, A_4 = -0.08794067, A_5 = 0$.

In 1995, Edelman and Kostlan [55] gave a geometrical derivation of the Kac-Rice formula, by arguing that the expected number of real zeros for iid Gaussian coefficients is given by $|\gamma|/\pi$, where

$$\gamma = (1, t, t^2, \ldots, t^n)/\|(1, t, t^2, \ldots, t^n)\|,$$

and $|\gamma|$ denotes the arclength of γ and $\|(1, t, \ldots, t^n)\| = \sqrt{\sum_{i=1}^{n} |t^i|^2}$ is the Euclidean norm. They are easily able to recover the previous results of Kac and Wilkins, and also find that in the case of iid standard Gaussians, the expected number of real zeros is given by $\frac{2}{\pi} \log n + C_1 + \frac{2}{n\pi} + O(1/n^2)$ as $n \to \infty$, where $C_1 = 0.6257358072....$ Here, the authors state that the $\frac{2}{\pi n}$ and C_1 terms are new improvements. Their proof uses the probability density function of the real zeros, which was obtained with their rectifiable curve argument.

In 2002, Dembo, Poonen, Shao, and Zeitouni [38] prove the following.

Theorem 1.2.8 (Theorem 1.1 and Theorem 1.2 in [38])**.** *Suppose that* $\{a_i\}$ *is a sequence of zero-mean, unit variance, iid random variables possessing finite moments of all orders. Let* $f_n(x) = \sum_{i=0}^{n-1} a_i x^i$ *and for* n *odd,*

$$P_n = \mathbb{P}(f_n(x) > 0 \ \forall x \in \mathbb{R}).$$

Then

$$\lim_{n\to\infty} \frac{\log P_{2n+1}}{\log n} = -b,$$

where b *is a certain positive finite constant. Moreover, the probability that the random polynomial* $f_{n+1}(x)$ *of degree* n *has* $o(\log n/\log\log n)$ *real zeros is* $n^{-b+o(1)}$ *as* $n \to \infty$*. For any fixed* k*, the probability* $p_{n,k}$ *that* f_{n+1} *has exactly* k *real zeros, all of which are simple, satisfies*

$$\lim_{n\to\infty} \frac{\log p_{2n+k,k}}{\log n} = -b.$$

If instead $\mathbb{E}(a_i) = \mu \neq 0$*, letting* $P_n^\mu = \mathbb{P}(f_n(x) \neq 0 \ \forall x \in \mathbb{R})$*, then*

$$\lim_{n\to\infty} \frac{\log P^\mu_{2n+1}}{\log n} = -b/2.$$

To obtain this result, the authors first studied the Gaussian case, and used the approximation techniques developed in [98] to extend to other distributions. This work was recently generalized by Ghosal and Mukherjee [70], who provide a new result assuming only finite mean and finite variance assumptions on the iid random variables.

In 2008, Erdélyi [57] proves the following theorem, improving upon the previously discussed lower bound of Bloch and Póyla regarding the maximum number of real zeros that polynomials with coefficients in $\{-1, 0, 1\}$ can take.

Theorem 1.2.9 (Theorem 2.1 in [57])**.** *There are absolute constants* $c_1 > 0$ *and* $c_2 > 0$ *such that for every* $\{a_0, \dots, a_n\} \in [1, M]$ *with* $1 \leq M \leq e^{c_1 n^{1/4}}$, *there are* $b_0, \dots, b_n \in \{-1, 0, 1\}$ *such that the polynomial*

$$b_n a_n z^n + \cdots + b_1 a_1 z_1 + b_0 a_0$$

has at least $c_2 n^{1/4}$ *distinct sign changes in* $(0, 1)$.

In 2015, Do, Nguyen, and Vu [41] find bounds for the probability of double real roots of certain Kac polynomials. They further find an asymptotic bound for the expected number of real roots, as $n \to \infty$, of many types of Kac polynomials. This result is contained in the following theorem, and their main proof method uses Inverse Littlewood-Offord theory, see for instance [123].

Theorem 1.2.10 (Theorem 1.5 in [41])**.** *For any positive integer N, we say that ξ has a uniform distribution with parameter N (or type I) if $\mathbb{P}(\xi = i) = 1/(2N)$ independently, $i \in \{\pm 1, \pm 2, \ldots, \pm N\}$. Let $\epsilon_0 > 0$ and $p > 1$. We say that a random variable ξ of mean zero has type II distribution with parameter (p, ϵ_0) is it has a p-integrable density function and its $(2 + \epsilon_0)$-moment is bounded.*

Let ξ be a random variable with either type I or type II fixed parameters. Then the expected number of real roots is $\frac{2}{\pi} \log n + C + o(1)$, where C is an absolute constant depending on ξ.

In 2016, Nguyen, Nguyen, and Vu [122] show that the expected number of real roots of certain Kac polynomials is given by the following theorem. Their new method uses universality from [41] and the lack of double roots in intervals.

Theorem 1.2.11 (Theorem 1 in [122])**.** *For any random variable ξ with mean 0 and variance 1 and bounded $(2 + \epsilon)$-moment, the expected number of real roots of $\sum_{i=0}^{n} \xi_i x^i$ (where the ξ_i are iid copies of ξ) is given by*

$$\frac{2}{\pi} \log n + O_{\epsilon,\xi}(1).$$

Also in 2015, Tao and Vu [157] demonstrate a universality result for the zeros of many types of random polynomials with independent coefficients, taking the form given in (1.1), including Kac polynomials. Roughly, the authors show that for polynomials f and $\tilde{f}$ of the same degree with iid coefficients ξ and $\tilde{\xi}$ which have the same mean and variance, as the degree of the polynomials tend to infinity, the correlation functions of the zeros of the polynomials are very similar to each other. This implies that the zeros of the polynomials behave similarly, as $n \to \infty$, even though ξ and $\tilde{\xi}$ can be quite different; this is known as the universality phenomenon. The authors prove this by using the replacement principle, which is Theorem 2.1 in their paper. With universality established, the authors give the expected number of real zeros for different types of polynomials, such as flat and elliptical.

Many other polynomial models are possible, and the coefficients do not necessarily need to be taken to be independent. We mention a few now, though this list is not exhaustive.

Letting the coefficients in (1.1) be given by $c_i = \sqrt{\frac{1}{i!}}$ corresponds to Flat or Weyl polynomials, letting $c_i = \sqrt{\binom{n}{i}}$ corresponds to elliptical or binomial polynomials, and letting $c_i = \sqrt{\frac{L(L+1)\cdots(L+i-1)}{i!}}$ for some $L > 0$ corresponds to hyperbolic polynomials; Tao and Vu [157] provide a good summary. For the behavior of the number of real roots non-Kac polynomial ensembles when the degree tends to infinity, we refer the reader to Kostlan [99], Wang [166] Logan and Shepp [110, 111], Shiffman and Zelditch [146], Reid and Sambandham [21] and Farahmand [61], and Kabluchko and Zaporozhets [87].

One can also consider random polynomials in a single complex variable z of the form

$$f(z) = \prod_{i=1}^{n} (z - \xi_i),$$

where the roots $\xi_1, \ldots, \xi_n$ are independent random variables. For more information on this topic, see for instance [100, 124] and references therein.

One can study the random trigonometric polynomials, which are of the form

$$\sum_{k=1}^{n} a_k \cos(kt) + b_k \sin(kt), t \in [0, 2\pi]$$

where a_k, b_k are independent random variables; see for instance [5, 42]

1.3 Random matrix ensembles

Letting $n, N \in \mathbb{N}$ and $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for $1 \leq i, j \leq n$, we say that n-by-n random matrices whose entries are uniformly distributed on $[\alpha_{ij} N, \beta_{ij} N]$ for $1 \leq i, j \leq n$ and jointly independent are drawn from the **generalized large box model ensemble**. We wish to find the probability that such random matrices have all real eigenvalues. As was the case with polynomials, many authors first considered random matrices whose entries were iid real or complex Gaussians, and obtained asymptotic results as the dimension of the matrix tended to infinity. We begin with a history of these ensembles and present interesting and pertinent results.

We refer the reader to a complete account of the history of random matrices given by Forrester, Snaith, and Verbaarschot [63], which details the introduction of random matrix theory by Wishart [174] in 1928 and the early pioneering work by Wigner [169,170,171], Dyson [49,50,51,52,53], Porter and Rosenzweig [133], Gaudin [69], and Mehta [118] in the mid-20th century. Early symmetry conditions, such as constraining the matrices to be Hermitian, were also common, see for instance Wigner [172] and Füredi and J. Komlós [67].

One of the main motivations of the work in Chapters 4 and 5 of this thesis is the Ginibre ensemble, which we define now.

Definition 1.3.1 (Ginibre Ensemble)**.** Recall that a real standard Gaussian random variable is a normal random variable which has mean zero and variance one. A complex standard Gaussian random variable has real and imaginary parts that are independent normal random variables, each with mean zero and variance 1/2. When the entries of an n-by-n random matrix A are iid real standard Gaussian random variables, we say that A is drawn from the **real Ginibre ensemble**. When the entries are complex standard Gaussian variables instead, we say that A is drawn from the **complex Ginibre ensemble**.

In 1965, Ginibre [71] first inquired about the eigenvalue distributions of this matrix ensemble, allowing both complex Gaussian and real Gaussian entries. He discovers the joint probability density function of the eigenvalues for matrices drawn from the complex Ginibire ensemble. For the real Ginibire ensemble, Ginibire is only able to calculate the joint probability density function for the eigenvalues in the case where all are assumed to be real. Ginibre's methods and results are detailed in Mehta's [118] book, and we outline them here.

Let $\mathbb{P}(A)\,dA$ denote the joint probability density function for a random matrix A on the space of n-by-n real (or complex) matrices with respect to the Lebesgue measure on $\mathbb{R}^{n^2}$ (or $\mathbb{R}^{2n^2}$, respectively). For the real Ginibre ensemble, this density is given by $\mathbb{P}(A)\ dA = \frac{1}{(2\pi)^{n^2/2}} e^{-\operatorname{tr}(A^T A)}\ dA$, where A^T denotes the transpose of the real matrix A and $dA = \prod_{1\le i,j\le n} dA_{ij}$. For the complex Ginibre ensemble, this density is given by $\mathbb{P}(A)\, dA = \frac{1}{\pi^{n^2}} e^{-\operatorname{tr}(A^*A)}\, dA$, where A^* denotes the conjugate transpose of the complex matrix A and $dA = \prod_{1\le i,j\le n} dA_{ij} dA^*_{ij}$; here $dzdz^*$ means $2dxdy$ if $z = x + iy$ and z^* denotes the complex conjugate of z.

For the complex Ginibre ensemble, Ginibre finds that the joint probability density function for the eigenvalues is given by

$$P_n(z_1, \dots, z_n) = \frac{1}{\pi^n \prod_{1\le k\le n} k!} e^{-\sum_{i=1}^n |z_i|^2} \prod_{1\le i<j\le n} |z_i - z_j|^2.$$

Ginibre then determines the k-point correlation functions. Using Mehta's notation, this is given by

$$R_k(z_1, \dots, z_k) = \frac{1}{\pi^k} e^{-\sum_{i=1}^k |z_i|^2} \det[K_n(z_i, z_j)]_{i,j=1,\dots,k}$$

where

$$K_n(z_i, z_j) = \sum_{p=0}^{n-1} \frac{(z_i z_j^*)^p}{p!}.$$

Considering the one-point correlation function, corresponding to the eigenvalue density in the complex plane, Ginibre and Mehta both show that

$$R_1(z) = \frac{1}{\pi} e^{-|z|^2} \sum_{l=0}^{n-1} \frac{|z|^{2l}}{l!}.$$

Letting $n \to \infty$, this density is $\frac{1}{\pi}$ whenever $|z| < \sqrt{n}$ and 0 otherwise. Scaling by $\sqrt{n}$, this establishes the circular law for the complex Ginibire ensemble.

For the real Ginibire ensemble, Ginibre is unable to proceed as in the complex case due to difficulties in carrying out the necessary integrations. However, assuming that all eigenvalues are real, he is able to show that the joint probability density of the real eigenvalues $\lambda_1, \dots, \lambda_n$ is given by

$$P_n(\lambda_1, \dots, \lambda_n) = \frac{1}{2^{n/2} \prod_{i=1}^n \Gamma(i/2)} e^{-\sum_{i=1}^n |\lambda_i|^2/2} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|,$$

where for $\mathrm{Re}(s) > 0$, the gamma function is given by $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt$.

In 1997, Edelman [54] continues the investigation into the real Ginibre ensemble. He finds the joint density of the eigenvalues of an n-by-n random matrix drawn from the real Ginibre ensemble and even obtains a non-asymptotic matrix result: computing the probability that a matrix from this ensemble has exactly k real eigenvalues. Letting $k = n$, he finds that the probability that an n-by-n real random matrix with entries that are iid standard Gaussians has all real eigenvalues, which comes out to $2^{-n(n-1)/4}$. He also establishes a version of the circular law in this case, given by the following theorem. A stronger version of the circular law is given by Girko [72], with an additional proof given by Bai [7]. Mehta [119] again publishes the details in an updated book.

Theorem 1.3.2 (Theorem 6.3 in [54]). *Let A be a real n-by-n matrix whose elements are independent random variables with standard normal distributions. Let $\hat{x} = x/\sqrt{n}$ and $\hat{y} = y/\sqrt{n}$, where $z = x + iy$ denotes an eigenvalue of A. Let $\hat{p}_n(\hat{x}, \hat{y})/n$ denote the density of a randomly chosen normalized eigenvalue in the upper half plane. Then as $n \to \infty$,*

$$\lim_{n\to\infty} \frac{\hat{p}_n(\hat{x}, \hat{y})}{n} = \begin{cases} \pi^{-1} & \text{if } \hat{x}^2 + \hat{y}^2 < 1 \\ 0 & \text{if } \hat{x}^2 + \hat{y}^2 > 1 \end{cases}.$$

Furthermore, a randomly chosen normalized eigenvalue of A converges in distribution to the uniform distribution in the unit disk.

In Chapter 4, we use Edelman's method of considering the real Schur decomposition to write the probability that an n-by-n random matrix drawn from the generalized bounded height ensemble (that is, entries uniformly drawn from $[\alpha_{ij}, \beta_{ij}]$ and jointly independent for $1 \leq i, j \leq n$, with $\alpha_{ij} < \beta_{ij}$) has all real eigenvalues via an integral formula. In Chapter 5, we show how this probability relates to the probability that a matrix drawn from the generalized large box model ensemble has all real eigenvalues.

Around the same time, Edelman, Kostlan, and Shub [56] prove the following.

Theorem 1.3.3 (Corollary 5.2 in [56]). *Let E_n denote the expected number of real eigenvalues of an n-by-n random matrix whose elements are independent random variables from a real standard normal distribution. Then*

$$\lim_{n\to\infty} \frac{E_n}{\sqrt{n}} = \sqrt{\frac{2}{\pi}}$$

and

$$E_n = \sqrt{\frac{2n}{\pi}}\left(1 - \frac{3}{8n} - \frac{3}{128n^2} + \frac{27}{1024n^3} + \frac{499}{32768n^4} + O\left(\frac{1}{n^5}\right)\right) + \frac{1}{2}$$

as $n \to \infty$.

In 2010, Bourgain, Vu, and Matchett Wood [27] study complex random matrices with independently distributed discrete entries, as the matrix size n tends to infinity. For $p \in (0, 1)$ and $r \in \mathbb{N}$, they establish the upper bound of $(p^{1/r} + o(1))^n$ for the probability that such a matrix is singular, given that its entries satisfy a property known as being p-bounded of exponents r. This relates to eigenvalues because a matrix is singular if and only if it has zero as an eigenvalue; equivalent criteria for the singularity of a matrix can be found on page 14 in [79]. As an example, when the matrix entries are Bernoulli random variables with parameter 1/2, the authors show that their upper bound equals $(1/\sqrt{2} + o(1))^n$, although they mention a long-standing conjecture which suggests that the correct probability is $(1/2 + o(1))^n$. Recently, in 2020, Tikhomirov [160] shows that the probability that an n-by-n matrix whose entries are Bernoulli random variables with parameter $p \in (0, 1/2]$ is singular equals $(1 - p + o(1))^n$; when $p = 1/2$, this confirms the conjecture. As a corollary to their main theorem, Bourgain, Vu, and Matchett Wood provide the following upper bound on the probability that an integer matrix has a rational eigenvalue.

Corollary 1.3.4 (Corollary 3.7 in [27])**.** *Fix a positive integer* N*, and let* $M_{n,N}$ *be an* n*-by-*n *random integer matrix with independent entries, each of which takes values in the set* $[-N, N] \cap \mathbb{Z}$*. Let* c *be a constant such that for every matrix entry* m*, we have* $\max_{-N \leq x \leq N} \mathbb{P}(m = x) \leq c/N$*. Then the probability that* $M_{n,N}$ *has a rational eigenvalue is at most* $(c/N + o(1))^{n/2}$*, where the* $o(1)$ *term goes to zero as* $n \to \infty$*.*

In 2010, Tao and Vu [156] make an exciting breakthrough by establishing universality and proving that the circular law holds under only a finite variance assumption. The authors reference several prior results on this topic, which prove the circular law under stronger hypothesis, such as [8, 73, 75, 128, 155, 156].

Theorem 1.3.5 (Theorem 1.10 (Circular law) in [156]). *Given an n-by-n complex matrix X_n, let*

$$\mu_{X_n}(x, y) := \frac{1}{n}|\{1 \leq i \leq n, \mathrm{Re}(\lambda_i) \leq x, \mathrm{Im}(\lambda_i) \leq y\}|$$

be the empirical spectral distribution (ESD) of its eigenvalues $\lambda_i \in \mathbb{C}, i = 1, \ldots, n$. Let X_n be the n-by-n random matrix whose entries are iid complex random variables with mean zero and variance one. Then, the ESD of $\frac{1}{\sqrt{n}}X_n$ converges (both in probability and in the almost sure sense) to the uniform distribution on the unit disk.

In 2005, Akemann and Kanzieper [91] deliver a formula for the exact probability that an n-by-n matrix with iid real entries from $\mathcal{N}(0, 1)$ has exactly k real eigenvalues, whenever $n - k$ is even, expanding on Edelman's result.

Theorem 1.3.6 ((I) in [91]). *Let A be an n-by-n random real matrix whose entries are statistically independent random variables drawn from the normal distribution $\mathcal{N}(0, 1)$. For $n - k = 2l$ even, the probability $p_{n,k}$ of exactly k real eigenvalues occurring is*

$$p_{n,k} = p_{n,n-2l} = \frac{p_{n,n}}{l!} Z_{(1^l)}(p_1, \ldots, p_l),$$

where $p_{n,n} = 2^{-n(n-1)/4}$ due to Edelman, $Z_{(1^0)} = 1$ and

$$Z_{(1^l)}(p_1, \ldots, p_l) = (l - 1)! \sum_{r=0}^{l-1} \frac{(-1)^{l-r-1}}{r!} p_{l-r} Z_{(1^r)}(p_1, \ldots, p_r).$$

The $Z_{(1^l)}$ are called zonal polynomials.

In 2007, Akemann and Kanzieper [1] provide the details of their proof, which is primarily based on a new Pfaffian integration technique. In 2016, Kazieper et. al [92] consider the real Ginibre ensemble, and find the probability that a $2n$-by-$2n$ random matrix has exactly $2k$ real eigenvalues, denoted $p_{2n,2k}$, as $n \to \infty$. They show

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k} = -\frac{1}{\sqrt{2\pi}}\zeta(3/2),$$

where ζ is the Riemann zeta-function.

In 2015, Tao and Vu [158] obtain universality results for non-Hermitian random matrices. They prove what is called a "Four Moment Theorem" for both real and complex matrices, essentially showing that matrices which match the respective Ginibre ensemble to four moments have correlation functions which are very similar to one another. As with their analogous universality result for polynomials, as a corollary, they are able to obtain an asymptotic bound on the expected number of real eigenvalues of such matrices. A precise statement is provided by the following theorem.

Theorem 1.3.7 (Corollary 17 in [158]). *Let $M_n = (\xi_{ij})_{1\leq i,j\leq n}$, where the ξ_{ij} are independent and complex random variables, which match moments with $\mathcal{N}(0,1)$ to the fourth order. Moreover, assume that for all real $t \geq 0$,*

$$\mathbb{P}(|\xi_{ij}| \geq t) \leq Ce^{-t^c}$$

for some fixed $C, c > 0$ (independent of n) and all i, j. Assume n is even. Then the expected number of real eigenvalues is

$$\sqrt{\frac{2n}{\pi}} + O(n^{1/2-c'})$$

for some fixed $c' > 0$.

In 2018, Luh [112] provides an upper bound for certain complex random matrices having a real eigenvalue.

Theorem 1.3.8 (Theorem 3.2 in [112]). *Let $\zeta = \xi + i\xi'$ be a random variable, where ξ and ξ' are iid, mean zero, variance 1, and subgaussian with moment B. Let M_n be an n-by-n random matrix populated with independent copies of ζ. Then there exists a $c \in (0,1)$ depending only on B such that*

$$\mathbb{P}(M_n \textit{ has a real eigenvalue }) < c^n.$$

This bound is the best possible up to the constant c.

1.3.1 Large box model ensemble

The following results are related specifically to the large box model ensemble, where entries are uniformly iid on $[-N, N] \cap \mathbb{Z}$.

In 2008, Rivin [138] considers large box model random matrices. Let $SL(n, \mathbb{Z})$ denote the special linear group of degree n over $\mathbb{Z}$. Rivin shows the following.

Theorem 1.3.9 (Theorem 5.2 in [138]). *The probability that a matrix in $SL(n, \mathbb{Z})$ with coefficients bounded by B has a reducible characteristic polynomial goes to zero as B goes to infinity.*

In 2021, Karingula and Lovett [93] find an upper bound for the probability that an n-by-n random matrix with entries uniformly distributed in $[-N, N] \cap \mathbb{Z}$ is singular, when both n and N are large. They show that for $n, N \geq 1$, the probability of singularity of such a matrix is bounded above by N^{-cn}, where $c > 0$ is an absolute constant. They mention that Katznelson [94, 95] considered a similar problem with large N and constant n, obtaining the bound of $c_n N^{-n}$, where c_n is a constant depending on the dimension of the matrix.

In 2007, Hetzel, Liew, and Morrison [77] show that the probability that a 2-by-2 random matrix with entries uniformly iid on $[-1, 1]$ has all real eigenvalues is 49/72, and pose a question regarding the diagonalizability of integer matrices over $\mathbb{Q}$. In 2008, Martin and Wong [114, 115] show that for $C = \frac{7\sqrt{2}+4+3\log(\sqrt{2}+1)}{3\pi^2}$, the probability as $N \to \infty$ that a 2-by-2 random matrix with entries uniformly iid on $[-N, N] \cap \mathbb{Z}$ has integer eigenvalues is asymptotically $C \log N/N$. They also answer the question of the previous authors, showing dimension $n \geq 2$, it is highly unlikely for integer matrices to have integer eigenvalues, as seen in the following theorem.

Theorem 1.3.10 (Theorem in [115]). *Let $M_n(N)$ denote the set of all n-by-n matrices whose entries are uniformly chosen in $[-N, N] \cap \mathbb{Z}$. Given any integer $n \geq 2$ and any real number $\epsilon > 0$, the probability that a randomly chosen matrix in $M_n(N)$ has an integer eigenvalues is $O_{n,\epsilon}(1/N^{1-\epsilon})$. In particular, the probability that a randomly chosen matrix in $M_n(N)$ is diagonalizable over the rational numbers is $O_{n,\epsilon}(1/N^{1-\epsilon})$.*

1.3.2 Finite rank perturbations of random matrices

In the second half of Chapter 5, we restrict ourselves to just the large box model ensemble (where matrix entries are uniformly iid on $[-N, N] \cap \mathbb{Z}$), and ask about what happens to the eigenvalues of such matrices when perturbed by another matrix whose entries are all μ_N. The three cases where $\lim_{N\to\infty} \frac{\mu_N}{N}$ tends to 0, ∞, or some real finite positive constant c are all considered.

We review some of the relevant literature on finite rank perturbations now. In the following definition adapted from Tao [153], we emphasize that what Tao refers to as the empirical spectral distribution, we call the empirical spectral measure.

Definition 1.3.11 (Definition 1.2 in [153])**.** Let A_n be an n-by-n complex matrix with eigenvalues $\lambda_1^N, \ldots, \lambda_n^N \in \mathbb{C}$ counted with algebraic multiplicity. The **empirical spectral measure** μ_{A_n} of A_n is defined to be the probability measure

$$\mu_{A_N} := \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j^N}.$$

Here, δ_x denotes the Dirac measure centered at a point x. If A_n is a random n-by-n complex matrix, we say that μ_{A_N} **converges weakly almost surely** to another (Borel) probability measure μ on the complex plane $\mathbb{C}$ if for every bounded and continuous function $F : \mathbb{C} \to \mathbb{C}$, $\int_{\mathbb{C}} F \, d\mu_{A_N}$ converges almost surely to $\int_{\mathbb{C}} F \, d\mu$ as $n \to \infty$.

In 2005, Baik, Ben Arous, and Péché [9] analyze the sample covariance matrix created by considering m iid complex Gaussian sample vectors, each in n variables. Fixing all but r eigenvalues of the covariance matrix, and letting $n, m \to \infty$, they show that the largest eigenvalue of the covariance matrix is described by a generalization of the Tracy-Widom distribution, and they describe a phase transition that occurs. In 2006, Péché [129] also considers the distribution of the largest eigenvalue of Hermitian random matrices with iid Gausian entries as the matrix size tends to infinity.

In 2010, Tao and Vu [156] establish the circular law for additive perturbations, demonstrating that the normalized empirical spectral measure of an iid complex random matrix with entries which have mean zero and unit variance, when additively perturbed by a deterministic matrix M_n of rank $o(n)$ and such that $\sup_n \frac{1}{n^2}(\text{ tr } M_n M_n^*) < \infty$, converges (both in probability and in the almost sure sense) to the uniform distribution on the unit disk.

In 2011, Benaych-Georges and Nadakuditi [16] extend the result of Baik, Ben Arous, and Péché, as the authors investigate both additive and multiplicative perturbations of random matrices. For symmetric or Hermitian random matrices, their main theorem for additive perturbations is the following.

Theorem 1.3.12 (Eigenvalue phase transition, Theorem 2.1 in [16])**.** *Let X_n be an n-by-n symmetric (or Hermitian) random matrix whose ordered eigenvalues are denoted by $\lambda_1(X_n) \geq \cdots \geq \lambda_n(X_n)$. Suppose that the probability measure $\mu_{X_n} = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j(X_n)}$ converges almost surely weakly as $n \to \infty$ to a non-random compactly supported probability measure μ_X. Let a and b be, respectively, the infimum and supremum of the support of μ_X. We suppose the smallest and largest eigenvalue of X_n converge almost surely to a and b. For a given $r \geq 1$, let $\theta_1, \ldots, \theta_r$ be deterministic non-zero real numbers, chosen independently of n. For every n, let P_n be an $n \times n$ symmetric (or Hermitian) random matrix having rank r with its r non-zero eigenvalues equal to $\theta_1, \ldots, \theta_r$. Let the index $s \in \{0, \ldots, r\}$ be defined such that $\theta_1 \geq \cdots \geq \theta_s > 0 > \theta_{s+1} \geq \cdots \geq \theta_r$. We suppose that X_n and P_n are independent and that either X_n or P_n is orthogonally (or unitarily) invariant.*

Consider the rank r additive perturbation of the random matrix X_n given by $\tilde{X}_n = X_n + P_n$.

The extreme eigenvalues of $\tilde{X}_n$ exhibit the following behavior as $n \to \infty$. We have that for each $1 \leq i \leq s$,

$$\lambda_i(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} G_{\mu_X}^{-1}(1/\theta_i) & \text{if } \theta_i > 1/G_{\mu_X}(b^+), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $i > s$, $\lambda_i(\tilde{X}_n) \xrightarrow{a.s.} b$.

Similarly, for the smallest eigenvalues, we have that for each $0 \leq j < r - s$,

$$\lambda_{n-j}(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} G_{\mu_X}^{-1}(1/\theta_{r-j}) & \text{if } \theta_j < 1/G_{\mu_X}(a^-), \\ a & \text{otherwise,} \end{cases}$$

while for each fixed $j \geq r - s$, $\lambda_{n-j}(\tilde{X}_n) \xrightarrow{a.s.} a$.

Here,

$$G_{\mu_X}(z) = \int \frac{1}{z-t} d\mu_X(t) \quad \textit{for } z \notin \textit{supp}(\mu_X),$$

is the Cauchy transform of μ_X, *and* $G_{\mu_X}^{-1}(\cdot)$ *is its functional inverse.*

The reader is referred to [14, 15, 17] for more papers by Benaych-Georges and co-authors concerning eigenvalue fluctuations after low rank perturbations.

In 2013, Tao [152, 153] characterizes the limiting behavior of outlier eigenvalues created by additively perturbing certain iid random matrices by deterministic matrices which have both rank and operator norm $O(1)$. As the matrix size tends to infinity, Tao shows that the number and location of the resulting outlier eigenvalues depend on the number and location of the eigenvalues of the perturbation matrix. We state Tao's theorem now.

Theorem 1.3.13 (Theorem 1.7 in [153])**.** *Let* X_n *be an* n*-by-*n *iid random matrix with complex entries normalized to have mean zero, variance one, and finite fourth moment. For each* n, *let* P_n *be a deterministic matrix with rank* $O(1)$ *and operator norm* $O(1)$. *Let* $\epsilon > 0$, *and suppose that for all sufficiently large* n, *there are no eigenvalues of* P_n *in the band* $\{z \in \mathbb{C} \mid 1 + \epsilon < |z| < 1 + 3\epsilon\}$, *and there are* j *eigenvalues* $\lambda_1(P_n), \ldots, \lambda_j(P_n)$ *for some* $j = O(1)$ *in the region* $\{z \in \mathbb{C} \mid |z| \geq 1 + 3\epsilon\}$. *Then, almost surely, for sufficiently large* n, *there are precisely* j *eigenvalues* $\lambda_1\left(\frac{1}{\sqrt{n}}X_n + P_n\right), \ldots, \lambda_j\left(\frac{1}{\sqrt{n}}X_n + P_n\right)$ *of* $\frac{1}{\sqrt{n}}X_n + P_n$ *in the region* $\{z \in \mathbb{C} \mid |z| \geq 1 + 2\epsilon\}$, *and after labeling these eigenvalues properly,* $\lambda_i\left(\frac{1}{\sqrt{n}}X_n + P_n\right) = \lambda_i(P_n) + o(1)$ *as* $n \to \infty$ *for each* $1 \leq i \leq j$.

In 2024, Banerjee, Mukherjee, and Pal [10] consider n-by-n symmetric random matrices X_n with standard Gaussian entries. For any $\epsilon > 0$, the decay condition $\sup_{(i,j) \neq (i',j')} \left|\mathbb{E}[X_{ij} X_{i'j'}]\right| = O(n^{-(1+\epsilon)})$ on the correlation between matrix entries is also imposed. Under these assumptions, the authors are able to show that as the matrix size tends to infinity, the largest eigenvalue of the normalized matrix converges to 2 almost surely. Let $\mathbb{1}$ be the n-dimensional column vector of all ones. When the matrix X_n is additively perturbed by the matrix $\frac{\lambda}{\sqrt{n}} \mathbb{1}\mathbb{1}^T$ and one takes $\epsilon \geq 1$ and

$\lambda \gg n^{1/4}$, the authors show that as $n \to \infty$, the fluctuations of the largest eigenvalue behave like a standard Gaussian.

There are many more results regarding the outlier eigenvalues and the fluctuations which result after additive perturbations. Coston [36], who studied both additive and multiplicative perturbations of products of iid random matrices, provides a detailed review of the random matrix literature in her thesis. O'Rourke and Renfrew [125], who studied perturbations of elliptic random matrices, also include an informative history on the topic. Both works discuss several of the sources also mentioned here. We additionally refer the reader to Belinschi, Bordenave, Capitaine, and Cébron [13], Coston, O'Rourke, and Matchett Wood [35], Capitaine, Donati-Martin, Féral, and Février [29,30], Ding and Hong, [40], and Féral and Péché [62], as well as the references mentioned within.

1.4 Notation

We let n and N be in $\mathbb{N}$. For $x \in \mathbb{R}$, we let $\lfloor x \rfloor$ denote the greatest integer less than or equal to x and $\lceil x \rceil$ the least integer greater than or equal to x. For $z \in \mathbb{C}$ and $r > 0$, we let $B(z, r)$ be the open disk of radius r centered at z in the complex plane, and we let $\mathcal{C}_{z,r}$ be the circle of radius r centered at z in the complex plane. Let $\log(\cdot)$ denote the natural logarithm. For a real number x, let

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \cdot \\ 1 & \text{if } x > 0 \end{cases}$$

For a permutation σ, let $\operatorname{sgn}(\sigma) = 1$ if the permutation is even, and $\operatorname{sgn}(\sigma) = -1$ if the permutation is odd. Let λ denote the 1-dimensional Lebesgue measure and λ^n denote the n dimensional Lebesgue measure. For $\alpha < \beta \in \mathbb{R}$, let $U(\alpha, \beta)$ denote the continuous uniform distribution on the interval $[\alpha, \beta]$. All random variables in this thesis are real real, unless otherwise stated. For a real-valued random variable X, we let f_X denote the probability density function of X and F_X denote the cumulative distribution function of X. For any probability distribution P, we say $X \sim P$ to indicate that the random variable X has distribution P. For a sequence of real-valued distribution functions F_n, let $F_n \overset{\mathrm{d}}{\Rightarrow} F$ denote convergence in distribution to the limit F. For a sequence of real-valued random variables X_n, let $X_n \overset{\mathrm{d}}{\Rightarrow} X$ denote convergence in distribution of the distribution functions of X_n to the limit X as $n \to \infty$. Let $X_n \overset{\mathrm{as}}{\longrightarrow} X$ denote convergence almost surely, and let $X_n \overset{\mathrm{p}}{\to} X$ denote convergence in probability.

Throughout the thesis, we let a_i denote coefficients for random polynomials and a_{ij} denote entries for random matrices. Each a_i is contained in the interval $[\alpha_i, \beta_i]$ or $[\alpha_i N, \beta_i N]$ depending on the polynomial; each a_{ij} is contained in the interval $[\alpha_{ij}, \beta_{ij}]$ or $[\alpha_{ij} N, \beta_{ij} N]$ depending on the random matrix model. We always assume $\alpha_i < \beta_i$ (resp. $a_{ij} < b_{ij}$), and do not consider degenerate intervals. We denote the roots of the random polynomials by ξ_i and denote the eigenvalues of the random matrices by λ_i. Sometimes extra subscripts or tildes are used to clarify to which polynomial or matrix the roots or eigenvalues belong. Let x be a random variable and suppose that A is an n-by-n random matrix with iid entries that each have distribution x. Then we say that A has **atom distribution** x.

Denote the n-by-n identity matrix by I_n; I may also be used when the dimension is clear. Denote the i-th row of a matrix A by $A_{i,}$ and the i-th column by $A_{,i}$. We let $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ denote the sets of all n-by-n matrices over the fields $\mathbb{R}$ and $\mathbb{C}$, respectively. A principal submatrix of size k of an n-by-n matrix is formed by deleting $n-k$ rows and the corresponding $n-k$ columns. The determinant of a principal submatrix is called a principal minor. For an n-by-n matrix A, $E_k(A)$ denotes the sum of the principal minors of A of size k. $S_k(\xi_1, \ldots, \xi_n)$ denotes the k-th elementary symmetric polynomials in the variables $\xi_1, \ldots, \xi_n$, i.e.,

$$S_k(\xi_1, \ldots, \xi_n) = \sum_{1 \le l_1 < l_2 < \ldots < l_k \le n} \left(\prod_{j=1}^{k} \xi_{l_j} \right), \quad \text{for } 1 \le k \le n,$$

where

$$S_0(\xi_1, \ldots, \xi_n) = 1.$$

S_k denotes the symmetric group of size k.

For $\mathrm{Re}(\alpha) > 0, \mathrm{Re}(\beta) > 0, \mathrm{Re}(\gamma) > -\min\{1/n, \mathrm{Re}(\alpha)/(n-1), \mathrm{Re}(\beta)/(n-1)\}$, the n-dimensional Selberg integral is given by

$$\mathcal{S}_n(\alpha,\beta,\gamma) = \int_0^1 \cdots \int_0^1 \prod_{j=1}^{n} \xi_j^{\alpha-1}(1-\xi_j)^{\beta-1} \prod_{1\le j<k\le n} |\xi_j - \xi_k|^{2\gamma}\, d\xi_1 \ldots d\xi_n.$$

We let A^T denote the transpose of A, and A^* denotes the conjugate transpose. The trace of A is given by $\operatorname{tr} A$, and the determinant by $\det A$. The spectral radius of a matrix A is denoted by $\rho(A)$ and the set of all eigenvalues of A is denoted by $\sigma(A)$. J_n is the n-by-n matrix of all ones. The symbol $\mathbb{1}_A$ denotes the indicator function on a set A. For a vector $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$, the Euclidean norm of v is given by $\|v\| = \sqrt{\sum_{i=1}^n |v_i|^2}$. For a matrix $A \in M_n(\mathbb{C})$, the maximum column sum norm of A is given by $\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|$ and the spectral norm (or operator norm) of A, denoted $\|A\|_2$, is the square root of the largest eigenvalue of A^*A. We let $\|\cdot\|_{\max}$ denote the maximum magnitude of the matrix entries.

Let δ_{ij} denote the Kronecker delta, where

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}.$$

Let X be a set and let $\mathcal{M}$ be any σ-algebra of subsets of X. Suppose that $x \in X$. For $A \in \mathcal{M}$, define the Dirac measure δ_x by

$$\delta_x(A) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases}.$$

Chapter 2

Monic generalized large box model polynomials

2.1 Introduction and main results

Let n and N be in $\mathbb{N}$. Let $\alpha_i < \beta_i \in \mathbb{R}$ for $0 \leq i \leq n-1$ and consider the polynomial in a single complex variable

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

where for each $0 \leq i \leq n-1$, a_i is uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ and $a_0, \ldots, a_{n-1}$ are jointly independent. We wish to find the probability that $f_n(z)$ has all real roots as $N \to \infty$. We refer to this random polynomial model, with fixed degree and letting $N \to \infty$, as the **generalized large box** model; this contrasts with the usual large box model, where the coefficients are contained in $[-N, N] \cap \mathbb{Z}$ as $N \to \infty$.

Consider the following example.

Example 2.1.1. Let $n = 2$ and consider the case where $\alpha_0, \alpha_1 = 0$ and $\beta_0, \beta_1 = 1$ so that a_0 and a_1 are both identically, independently, and uniformly distributed on $[0, N] \cap \mathbb{Z}$. Then the polynomial $f_2(z) = z^2 + a_1 z + a_0$ has two real roots with probability tending to 1 as $N \to \infty$. To see this, notice that $f_2(z)$ will have two real roots if and only if the discriminant $a_1^2 - 4a_0$ is non-negative. Therefore, we compute

$$\mathbb{P}\left(a_1^2 - 4a_0 \geq 0\right) = \mathbb{P}\left(2\sqrt{a_0} \leq a_1\right),$$

where the equality holds by the non-negativity of the coefficients. Then

$$1 \geq \mathbb{P}\left(2\sqrt{a_0} \leq a_1\right) \geq \mathbb{P}\left(2\sqrt{N} \leq a_1\right) = \frac{N + 1 - \lceil 2\sqrt{N} \rceil}{N+1}.$$

By the squeeze theorem, $\lim_{N\to\infty} \mathbb{P}\left(2\sqrt{a_0} \leq a_1\right) = 1$, so the probability that $f_2(z)$ has two real roots as $N \to \infty$ tends to 1.

For degree 2 and 3 polynomials whose coefficients obey the generalized large box model, we are able to tackle the discriminants of the polynomials, allowing us to obtain precise results for the probability of having all real roots; these probabilities are simply in terms of the distributions of the coefficients of the polynomials. This results in the following two main theorems.

Theorem 2.1.2. *Let $\alpha_0 < \beta_0 \in \mathbb{R}$ and $\alpha_1 < \beta_1 \in \mathbb{R}$. Suppose that a_0 is uniformly distributed on $[\alpha_0 N, \beta_0 N]$ and that a_1 is uniformly distributed on $[\alpha_1 N, \beta_1 N]$, with a_0 and a_1 independent. Then the polynomial $f_2(z) = z^2 + a_1 z + a_0$ has two real roots with probability tending to 1 as $N \to \infty$.*

In [45], Dubickas and Sha define a dominant polynomial as a polynomial having one root whose modulus is greater than the moduli of all remaining roots. In terms of the notation of this thesis, they consider monic large box model polynomials, where each coefficient is uniformly distributed on $[-N, N] \cap \mathbb{Z}$ and jointly independent, and find that the probability that such a polynomial is dominant tends to 1 as $N \to \infty$. Since complex roots of polynomials with real coefficients occur in pairs, this means that the root with the largest modulus of a dominant polynomial must be real. For a quadratic polynomial, this means that the polynomial must have two real roots with probability tending to 1 as $N \to \infty$. Thus, Theorem 2.1.2 can be be interpreted as an extension of Dubickas and Sha's result about the dominance of quadratic polynomials. Dubickas and Sha prove their result that a degree n large box model polynomial is dominant by counting the number of polynomials that satisfy the hypothesis of Theorem 1.1.10. For the quadratic polynomial in Theorem 2.1.2, we simply compute the probability that the discriminant is positive.

Theorem 2.1.3. *For $0 \leq i \leq 2$, let $\alpha_i < \beta_i \in \mathbb{R}$ and let a_i be uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with a_0, a_1, a_2 jointly independent. As $N \to \infty$, the probability that the cubic polynomial*

$$f_3(z) = z^3 + a_2 z^2 + a_1 z + a_0$$

has three real roots converges to

$$\mathbb{P}\left(1/4 > \frac{XY}{Z^2}\right),$$

where $X \sim U(\alpha_0, \beta_0), Y \sim U(\alpha_2, \beta_2),$ *and* $Z \sim U(\alpha_1, \beta_1)$ *are independent.*

For $n = 4$, the discriminant method is already intractable, and for $n \geq 5$, the discriminant no longer provides complete information on the number of real roots of a polynomial, see for instance Chapter 10 in [84]. Chapter 3 offers the solution to these problems.

We also establish that the limiting distribution of the distance between one of the roots of the polynomial $f_2(z) = z^2 + a_1 z + a_0$ and the negative of the coefficient a_1 is simply the ratio distribution between the coefficients a_0 and a_1. This result is contained in the following theorem.

Theorem 2.1.4. *For* $0 \leq i \leq 1$, *let* $\alpha_i < \beta_i \in \mathbb{R}$. *Let* $f_2(z) = z^2 + a_1 z + a_0$, *where for each* $0 \leq i \leq 1$, a_i *is uniformly distributed on* $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, *with* a_0 *and* a_1 *independent. Consider the root*

$$\xi_2 = \begin{cases} \frac{-a_1 - \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} & \text{if } a_1 \neq 0 \\ \sqrt{-a_0} & \text{if } a_1 = 0 \end{cases} \tag{2.1}$$

of $f_2(z)$, *which may be real or imaginary. Let* Z_N *be the Euclidean distance between* $-a_1$ *and* ξ_2. *As* $N \to \infty$, $Z_N \stackrel{d}{\Rightarrow} |X/Y|$, *where* $X \sim U(\alpha_0, \beta_0)$, $Y \sim U(\alpha_1, \beta_1)$ *and* X *and* Y *are independent.*

Theorem 2.1.4 is illustrated in Figure 2.1

The root ξ_2 defined in Theorem 2.1.4 is described as extremely dominant. This is because its modulus is not tight as $N \to \infty$, while the modulus of the other root of $f_2(z)$ is tight as $N \to \infty$; complete details are contained in Section 2.6.1.

Using Rouché's theorem, we show in Section 2.6.2 that any monic polynomial $f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ whose coefficients obey the generalized large box model also has an extremely dominant root, which is well approximated by the extremely dominant root of the corresponding quadratic polynomial $f_2(z) = z^2 + a_{n-1}z + a_{n-2}$. The limiting distribution of the distance between this extremely dominant root and the negative of the coefficient a_{n-1} is surprisingly still only in terms of the ratio distribution between a_{n-2} and a_{n-1}. These results are given in

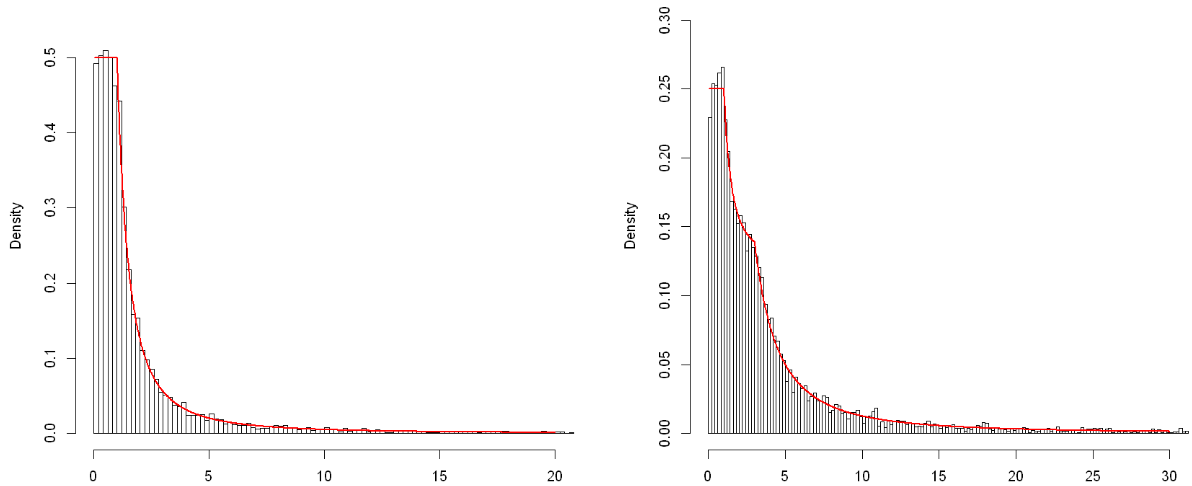

Figure 2.1: This figure illustrates Theorem 2.1.4. We generated 10000 random polynomials of the form $z^2 + a_1 z + a_0$ for both images. In the left image, the coefficients a_1 and a_0 are independently and uniformly distributed on $[0, 10^{15}] \cap \mathbb{Z}$. In the right image, a_1 is uniformly distributed on $[0, 10^{15}] \cap \mathbb{Z}$ and a_0 is independently and uniformly distributed on $[-3 \cdot 10^{15}, 10^{15}] \cap \mathbb{Z}$. The distance between the largest root (in magnitude) and the negative of the coefficient a_1 for each polynomial was calculated and plotted on the histogram. The red lines denote the density of $|X/Y|$, where $X \sim U(0,1)$ and $Y \sim U(0,1)$ are independent for the left image and where $X \sim U(-3,1)$ and $Y \sim U(0,1)$ are independent for the right image. Both plot windows were cropped.

Theorem 2.1.5 and Corollary 2.1.6. In [45], Dubickas and Sha test the dominance of a polynomial, without finding its roots, by giving an algorithm based on Sturm's theorem (see Theorem 1.4.3 in [134]), which gives the number of unique real roots of a polynomial in an interval by constructing a sequence and counting the number of sign changes of the sequence at the end points of the interval. For the case of the monic generalized large box model polynomials, the use of Rouché's theorem provides both the location of the dominant root and allows us to investigate its fluctuations. However, no additional information is provided for non-monic generalized large box model polynomials.

Theorem 2.1.5. *For* $0 \leq i \leq n-1$, *let* $\alpha_i < \beta_i \in \mathbb{R}$ *and let* a_i *be uniformly distributed on*

$[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with $a_0, \ldots, a_{n-1}$ jointly independent. For $n \geq 2$, define

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

and

$$f(z) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2}.$$

Define

$$\xi = \begin{cases} \frac{-a_{n-1} - \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} & \text{if } a_{n-1} \neq 0 \\ \sqrt{-a_{n-2}} & \text{if } a_{n-1} = 0 \end{cases}.$$

When $n = 2$, ξ is a root of both $f_n(z)$ and $f(z)$, which is real with probability tending to 1 as $N \to \infty$.

Suppose $n \geq 3$. If $0 \notin [\alpha_{n-1}N, \beta_{n-1}N]$, let $\epsilon = C_{\alpha,\beta,n}/N$, where $C_{\alpha,\beta,n}$ is the positive finite constant

$$C_{\alpha,\beta,n} = \frac{\sum_{i=0}^{n-3} \max\{|\alpha_i|, |\beta_i|\}}{(\min\{|\alpha_{n-1}|, |\beta_{n-1}|\})^2}.$$

Then the polynomial $f_n(z)$ has exactly one root, denoted ξ_n, in $B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. If $0 \in [\alpha_{n-1}N, \beta_{n-1}N]$, then for any fixed $0 < \epsilon < 1$, $f_n(z)$ has exactly one root, denoted ξ_n, in $B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. In both cases, ξ_n is a real root of $f_n(z)$ with probability tending to 1 as $N \to \infty$.

Corollary 2.1.6. *For $0 \leq i \leq n-1$, let $\alpha_i < \beta_i \in \mathbb{R}$ and let a_i be uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with $a_0, \ldots, a_{n-1}$ jointly independent. Define*

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0.$$

Let ξ_n be defined as in Theorem 2.1.5. Let Z_N be the Euclidean distance between $-a_{n-1}$ and ξ_n, i.e.,

$$Z_N = |\xi_n + a_{n-1}|.$$

As $N \to \infty$,

$$Z_N \overset{d}{\Rightarrow} \left| \frac{X}{Y} \right|,$$

where $X \sim U(\alpha_{n-2}, \beta_{n-2}), Y \sim U(\alpha_{n-1}, \beta_{n-1})$, and X and Y are independent.

Corollary 2.1.6 is illustrated in Figure 2.2.

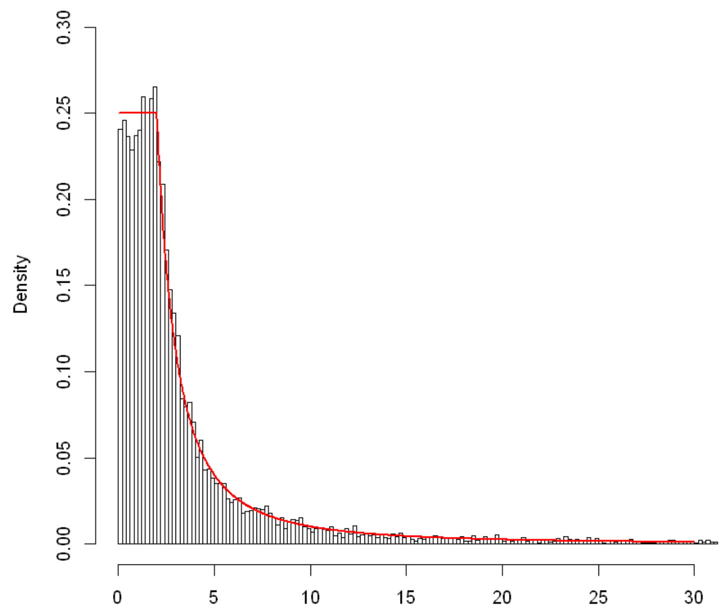

Figure 2.2: This figure illustrates Corollary 2.1.6. We generated 10000 random polynomials of the form $z^4 + a_3z^3 + a_2z^2 + a_1z + a_0$, where all coefficients are independent and a_3 is uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$, a_2 is uniformly distributed on $[-2 \cdot 10^{15}, 2 \cdot 10^{15}] \cap \mathbb{Z}$, a_1 is uniformly distributed on $[-3 \cdot 10^{15}, 3 \cdot 10^{15}] \cap \mathbb{Z}$, and a_0 is uniformly distributed on $[-4 \cdot 10^{15}, 4 \cdot 10^{15}] \cap \mathbb{Z}$. The distance between the largest root (in magnitude) and the negative of the coefficient a_3 for each polynomial was calculated and plotted on the histogram. The red line denotes the density of $|X/Y|$, where $X \sim U(-2, 2)$ and $Y \sim U(-1, 1)$ are independent. The plot window is cropped.

2.2 Tools and preliminary results

This section introduces several well-known probabilistic tools that will be necessary to obtain convergence results. We also prove some simple lemmas that will need to be referred to multiple times throughout the chapter.

Lemma 2.2.1. *Let $\alpha < \beta \in \mathbb{R}$ and let the random variable X_N be uniformly distributed on $[\alpha N, \beta N] \cap \mathbb{Z}$. Then X_N/N converges in distribution to $U(\alpha, \beta)$ as $N \to \infty$. If the random variable X_N is uniformly distributed on $([\alpha N, \beta N] \cap \mathbb{Z}) \setminus \{0\}$ instead, then X_N/N still converges* *in distribution to* $U(\alpha,\beta)$ *as* $N\to\infty$.

Proof. The cumulative distribution function for $U(\alpha,\beta)$ is given by

$$F_{U(\alpha,\beta)}(x)=\begin{cases}0 & \text{for } x<\alpha\\ \frac{x-\alpha}{\beta-\alpha} & \text{for } \alpha\le x\le\beta\\ 1 & \text{for } x>\beta\end{cases},$$

which is continuous for all $x\in\mathbb{R}$. Let $y\in\mathbb{R}$. If $y<\alpha$, then $\mathbb{P}(X_N/N\le y)=0$ since $X_N/N\in[\alpha,\beta]$. Similarly, if $y>\beta$, then $\mathbb{P}(X_N/N\le y)=1$. If $y\in[\alpha,\beta]$, then

$$\lim_{N\to\infty}\mathbb{P}(X_N/N\le y)=\lim_{N\to\infty}\mathbb{P}(X_N\le Ny)=\lim_{N\to\infty}\frac{\lfloor Ny\rfloor-\lceil\alpha N\rceil+1}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil+1}=\frac{y-\alpha}{\beta-\alpha},$$

so $X_N/N\overset{\text{d}}{\Rightarrow}U(\alpha,\beta)$.

Suppose now that X_N is uniformly distributed on $([\alpha N,\beta N]\cap\mathbb{Z})\setminus\{0\}$. If $0\notin[\alpha N,\beta N]\cap\mathbb{Z}$, then the statement and proof are the same as above. Therefore, suppose $0\in[\alpha N,\beta N]\cap\mathbb{Z}$. As before, if $y<\alpha$, then $\mathbb{P}(X_N/N\le y)=0$ since $X_N/N\in[\alpha,\beta]$ and if $y>\beta$, then $\mathbb{P}(X_N/N\le y)=1$. If $y\in[\alpha,0)$, we now have

$$\lim_{N\to\infty}\mathbb{P}(X_N/N\le y)=\lim_{N\to\infty}\mathbb{P}(X_N\le Ny)=\lim_{N\to\infty}\frac{\lfloor Ny\rfloor-\lceil\alpha N\rceil+1}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil}=\frac{y-\alpha}{\beta-\alpha}.$$

If $y\in[0,\beta]$, we now have

$$\lim_{N\to\infty}\mathbb{P}(X_N/N\le y)=\lim_{N\to\infty}\mathbb{P}(X_N\le Ny)=\lim_{N\to\infty}\frac{\lfloor Ny\rfloor-\lceil\alpha N\rceil}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil}=\frac{y-\alpha}{\beta-\alpha},$$

so $X_N/N\overset{\text{d}}{\Rightarrow}U(\alpha,\beta)$. Altogether, we see that $X_N/N\overset{\text{d}}{\Rightarrow}U(\alpha,\beta)$. □

Remark 2.2.2. Suppose that X_N is uniformly distributed on $[\alpha N,\beta N]\cap\mathbb{Z}$. If $0\in[\alpha N,\beta N]$, we will occasionally condition on the event that $X_N\ne 0$. Observe that for any $k\in([\alpha N,\beta N]\cap\mathbb{Z})\setminus\{0\}$, we have that

$$\mathbb{P}(X_N=k\mid X_N\ne 0)=\frac{\mathbb{P}(X_N=k)}{\mathbb{P}(X_N\ne 0)}=\frac{\left(\frac{1}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil+1}\right)}{\left(\frac{\lfloor\beta N\rfloor-\lceil\alpha N\rceil}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil+1}\right)}=\frac{1}{\lfloor\beta N\rfloor-\lceil\alpha N\rceil}.$$

Therefore, the conditional distribution of X_N, given that $X_N \neq 0$, is uniform on $([\alpha N, \beta N] \cap \mathbb{Z}) \setminus \{0\}$. Lemma 2.2.1 tells us that for any $z \in \mathbb{R}$

$$\lim_{N\to\infty} \mathbb{P}\left(X_N/N \leq z \mid X_N \neq 0\right) = \mathbb{P}(U(\alpha,\beta) \leq z).$$

Theorem 2.2.3 (Continuous Mapping Theorem, Theorem 3.2.10 from [48])**.** *Let g be a measurable function and $D_g = \{x : g \text{ is discontinuous at } x\}$. If $X_n \overset{d}{\Rightarrow} X$ as $n \to \infty$ and $\mathbb{P}(X \in D_g) = 0$, then $g(X_n) \overset{d}{\Rightarrow} g(X)$ as $n \to \infty$.*

Lemma 2.2.4 (Example 3.2 from [23])**.** *Suppose that $X_n \overset{d}{\Rightarrow} X$ and $Y_n \overset{d}{\Rightarrow} Y$ as $n \to \infty$, with X_n independent of Y_n for all $n \geq 1$ and X independent of Y. Then $X_nY_n \overset{d}{\Rightarrow} XY$ and $X_n+Y_n \overset{d}{\Rightarrow} X+Y$ as $n \to \infty$.*

Theorem 2.2.5 (Slutsky's Theorem, Theorem 11.4 in [76])**.** *Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be sequences of random variables. Suppose that X_n converges in distribution to X and Y_n converges in probability to a constant a as $n \to \infty$. Then*

$$X_n + Y_n \overset{d}{\Rightarrow} X + a,$$

$$X_n - Y_n \overset{d}{\Rightarrow} X - a,$$

$$X_n \cdot Y_n \overset{d}{\Rightarrow} X \cdot a,$$

and

$$\frac{X_n}{Y_n} \overset{d}{\Rightarrow} \frac{X}{a}, \text{ for } a \neq 0,$$

as $n \to \infty$.

Theorem 2.2.6 (Theorem 2.7, [163])**.** *Let X_n, X and Y_n be random vectors. Then*

(1) $X_n \overset{as}{\longrightarrow} X$ implies $X_n \overset{p}{\to} X$;

(2) $X_n \overset{p}{\to} X$ implies $X_n \overset{d}{\Rightarrow} X$;

(3) $X_n \overset{p}{\to} c$ for a constant c if and only if $X_n \overset{d}{\Rightarrow} c$;

(4) if $X_n \overset{d}{\Rightarrow} X$ *and* $d(X_n, Y_n) \overset{p}{\to} 0$*, then* $Y_n \overset{d}{\Rightarrow} X$*, where* $d(x, y)$ *is a distance function of* $\mathbb{R}^k$ *that generates the usual topology;*

(5) if $X_n \overset{d}{\Rightarrow} X$ *and* $Y_n \overset{p}{\to} c$ *for a constant* c*, then* $(X_n, Y_n) \overset{d}{\Rightarrow} (X, c)$*;*

(6) if $X_n \overset{p}{\to} X$ *and* $Y_n \overset{p}{\to} Y$*, then* $(X_n, Y_n) \overset{p}{\to} (X, Y)$*;*

Remark 2.2.7 (Division by Zero)*.* There are several places throughout the thesis where we are dividing by a discrete random variable that may take the value zero. In all of these instances, the probability of this happening tends to zero as $N \to \infty$. Since we are only concerned with the limiting distributions in these cases, and since the probability of the continuous limiting distributions taking the value zero is zero, this will not be an issue. The problem is avoided by either conditioning on the denominator of the discrete distribution being non-zero or using convergence theorems that rely on the set of discontinuities of the limiting distribution having measure zero.

The following fact is needed several times throughout the thesis.

Lemma 2.2.8. *Let* $\alpha_1 < \beta_1 \in \mathbb{R}$ *and* $\alpha_2 < \beta_2 \in \mathbb{R}$*. Let* a_1 *be uniformly distributed on* $[\alpha_1 N, \beta_1 N] \cap \mathbb{Z}$ *and let* a_2 *be uniformly distributed on* $[\alpha_2 N, \beta_2 N] \cap \mathbb{Z}$*, with* a_1 *and* a_2 *independent. Then for any* $\epsilon > 0$*,*

$$\lim_{N\to\infty} \mathbb{P}\left(|a_1| > \epsilon |a_2^2|\right) = 0.$$

Proof. If $0 \notin [\alpha_2 N, \beta_2 N]$, we may consider the limit

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{a_1}{a_2^2}\right| > \epsilon\right). \tag{2.2}$$

By Lemma 2.2.1, $a_1/N \overset{\mathrm{d}}{\Rightarrow} U(\alpha_1, \beta_1)$ and $a_2/N \overset{\mathrm{d}}{\Rightarrow} U(\alpha_2, \beta_2)$. By Theorem 2.2.3, $(a_2/N)^2 \overset{\mathrm{d}}{\Rightarrow} (U(\alpha_2, \beta_2))^2$. By Theorem 2.2.5, since $1/N$ converges in probability to the constant 0, it follows that $(a_1/N)(1/N) = a_1/N^2 \overset{\mathrm{d}}{\Rightarrow} 0$. Therefore, by Theorem 2.2.3 and Lemma 2.2.4,

$$\left|\frac{a_1}{a_2^2}\right| = \left|\frac{a_1/N^2}{(a_2/N)^2}\right| \overset{\mathrm{d}}{\Rightarrow} 0.$$

Since convergence in distribution to a constant implies convergence in probability to the same constant, (2.2) tends to zero.

If $0 \in [\alpha_2 N, \beta_2 N]$, we condition on the value of a_2. By the law of total probability, we have

$$\begin{aligned}\mathbb{P}(|a_1| > \epsilon|a_2^2|) &= \mathbb{P}(|a_1| > \epsilon|a_2^2| \mid a_2 \neq 0)\mathbb{P}(a_2 \neq 0) + \mathbb{P}(|a_1| > \epsilon|a_2^2| \mid a_2 = 0)\mathbb{P}(a_2 = 0) \\ &= \mathbb{P}(|a_1| > \epsilon|a_2^2| \mid a_2 \neq 0)\mathbb{P}(a_2 \neq 0) + \mathbb{P}(|a_1| > 0)\mathbb{P}(a_2 = 0) \\ &= \mathbb{P}\left(\left|\frac{a_1}{a_2^2}\right| > \epsilon \mid a_2 \neq 0\right)\left(\frac{\lceil \beta_2 N\rceil - \lfloor \alpha_2 N\rfloor}{\lceil \beta_2 N\rceil - \lfloor \alpha_2 N\rfloor + 1}\right) \\ &\quad + \mathbb{P}\left(|a_1| > 0\right)\left(\frac{1}{\lceil \beta_2 N\rceil - \lfloor \alpha_2 N\rfloor + 1}\right).\end{aligned} \tag{2.3}$$

Clearly $0 \leq \mathbb{P}\left(|a_1| > 0\right) \leq 1$. By Remark 2.2.2, we see that the conditional distribution of a_2, given that $a_2 \neq 0$, is uniform on $([\alpha_2, \beta_2] \cap \mathbb{Z}) \setminus \{0\}$ and that for any $z \in \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{P}\left(a_2/N \leq z \mid a_2 \neq 0\right) = \mathbb{P}(U(\alpha_2, \beta_2) \leq z).$$

Arguing similarly to the previous case, we see that

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{a_1}{a_2^2}\right| > \epsilon \mid a_2 \neq 0\right) = 0.$$

Therefore, letting $N \to \infty$ in (2.3),

$$\lim_{N\to\infty} \mathbb{P}(|a_1| > \epsilon|a_2^2|) = 0$$

as well. □

2.3 Probability of all real roots for generalized large box model polynomials of degree two

We now show that any monic degree two integral polynomial whose coefficients obey the generalized large box model has all real roots with probability tending to 1 as N tends to infinity; we have already seen this for a special case in Example 2.1.1.

Proof of Theorem 2.1.2. Notice that $f_2(z)$ will have two real roots if and only if the discriminant $a_1^2 - 4a_0$ is non-negative. By Lemma 2.2.1, Theorem 2.2.3, and Theorem 2.2.5,

$$\frac{a_1^2}{N^2} - \frac{4a_0}{N^2} \overset{\text{d}}{\Rightarrow} X^2,$$

where $X \sim U(\alpha_1, \beta_1)$, which has a cumulative distribution function that is continuous everywhere. Therefore, as $N \to \infty$, we have

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(a_1^2 - 4a_0 \geq 0\right) &= \lim_{N\to\infty} \mathbb{P}\left(\frac{a_1^2}{N^2} - \frac{4a_0}{N^2} \geq 0\right) \\ &= \mathbb{P}\left(X^2 \geq 0\right) \\ &= 1.\end{aligned}$$

□

2.4 Probability of all real roots for generalized large box model polynomials of degree three

In this section, we compute the probability that a degree three random polynomial whose coefficients obey the generalized large box model has all real roots; the result is given in terms of the coefficients of the polynomial. We begin with a technical detail that allows us to omit the possibility of such a polynomial having a repeated root.

Theorem 2.4.1 (See Pages 8 and 9 in [136]). *Suppose $a_n \neq 0$. For a polynomial $f_n(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ with roots $\xi_1, \ldots, \xi_n$, counted with algebraic multiplicity, the discriminant is*

$$D(f_n) = a_n^{2n-2} \prod_{1\leq j<k\leq n} (\xi_j - \xi_k)^2 .$$

The discriminant vanishes if and only if f_n has a multiple zero.

For two polynomials

$$f_n(z) = \sum_{i=0}^{n} a_i z^i = a_n \prod_{i=1}^{n} (z - \xi_i) \quad \text{and} \quad g_m(z) = \sum_{j=0}^{m} b_j z^j = b_m \prod_{j=1}^{m} (z - \tilde{\xi}_j),$$

*with $a_n, b_m \neq 0$, their **resultant** is defined by*

$$R(f_n, g_m) := a_n^m b_m^n \prod_{j=1}^{m} \prod_{i=1}^{n} (\xi_i - \tilde{\xi}_j).$$

The resultant is also given by the following determinant, which has m rows containing the coefficients of f_n and n rows with those of g_m, while empty positions indicate vanishing entries:

$$R(f_n, g_m) = \det \begin{pmatrix} a_0 & a_1 & \cdots & a_n & & & \\ & a_0 & a_1 & \cdots & a_n & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_m & & & \\ & b_0 & b_1 & \cdots & b_m & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_m \end{pmatrix}. \quad (2.4)$$

Finally, the resultant of f_n and its derivative f_n' relates to the discriminant of f_n as follows:

$$R(f_n, f_n') = a_n(-1)^{n(n-1)/2} D(f_n). \quad (2.5)$$

Lemma 2.4.2. *For $0 \leq i \leq 2$, let $\alpha_i < \beta_i \in \mathbb{R}$ and let a_i be uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with a_0, a_1, a_2 jointly independent. For $f_3(z) = z^3 + a_2 z^2 + a_1 z + a_0$, the probability of $f_3(z)$ having a repeated root as $N \to \infty$ tends to 0.*

Proof. Combining (2.4) and (2.5), one obtains the familiar formula which expresses the discriminant of $f_3(z)$ in terms of its coefficients. This is[1]

$$D(f_3) = a_2^2 a_1^2 - 4a_1^3 - 4a_2^3 a_0 - 27a_0^2 + 18a_2 a_1 a_0.$$

By Theorem 2.4.1, $D(f_3)$ is zero if and only if $f_3(z)$ has a repeated root. Suppose that a_0 and a_1 are fixed. If a_0 and a_1 are both zero, then $D(f_3)$ is the zero polynomial, but this only happens with probability $\frac{1}{(\lfloor \beta_0 N \rfloor - \lceil \alpha_0 N \rceil + 1)(\lfloor \beta_1 N \rfloor - \lceil \alpha_1 N \rceil + 1)}$, which tends to zero as $N \to \infty$. Otherwise, $D(f_3)$ is a degree one, two, or three polynomial in the variable a_2 and has at most three roots in $[\alpha_2 N, \beta_2 N] \cap \mathbb{Z}$. Therefore, as a_0 varies over $[\alpha_0 N, \beta_0 N]$ and a_1 varies over $[\alpha_1 N, \beta_1 N] \cap \mathbb{Z}$, the total possibilities for zeros of $D(f_3)$ is bounded above by

$$\frac{3(\lfloor \beta_0 N \rfloor - \lceil \alpha_0 N \rceil + 1)(\lfloor \beta_1 N \rfloor - \lceil \alpha_1 N \rceil + 1)}{(\lfloor \beta_0 N \rfloor - \lceil \alpha_0 N \rceil + 1)(\lfloor \beta_1 N \rfloor - \lceil \alpha_1 N \rceil + 1)(\lfloor \beta_2 N \rfloor - \lceil \alpha_2 N \rceil + 1)},$$

[1] We obtained the formulas for the discriminants of cubic and quartic polynomials from Wikipedia: https://en.wikipedia.org/wiki/Discriminant. These were verified using Mathematica.

which tends to zero as $N \to \infty$. □

Remark 2.4.3. By an analogous argument, $f_4(z) = z^4 + a_3z^3 + a_2z^2 + a_1z + a_0$ also does not have a repeated root with probability tending to 1 as $N \to \infty$. Indeed, again combining (2.4) and (2.5), we have

$$\begin{aligned} D(f_4) = 256a_0^3 &- 192a_0^2a_1a_3 - 128a_0^2a_2^2 + 144a_0^2a_2a_3^2 \\ &- 27a_0^2a_3^4 + 144a_0a_1^2a_2 - 6a_0a_1^2a_3^2 - 80a_0a_1a_2^2a_3 \\ &+ 18a_0a_1a_2a_3^3 + 16a_0a_2^4 - 4a_0a_2^3a_3^2 - 27a_1^4 \\ &+ 18a_1^3a_2a_3 - 4a_1^3a_3^3 - 4a_1^2a_2^3 + a_1^2a_2^2a_3^2, \end{aligned}$$

so we see again that if a_0 and a_1 are both fixed to be zero, $D(f_4)$ is the zero polynomial; this happens only with probability tending to zero as $N \to \infty$. Otherwise, fixing a_0, a_1 and a_2, we see that $D(f_4)$ is a degree one, two, three, or four polynomial in the variable a_3, so counting similarly to Lemma 2.4.2 proves the result.

We are now ready to prove Theorem 2.1.3.

Proof of Theorem 2.1.3. The cubic polynomial $z^3 + a_2z^2 + a_1z + a_0$ has the discriminant

$$a_2^2a_1^2 - 4a_1^3 - 4a_2^3a_0 - 27a_0^2 + 18a_2a_1a_0 \tag{2.6}$$

which can also be written in the form

$$(\xi_1 - \xi_2)^2(\xi_1 - \xi_3)^2(\xi_2 - \xi_3)^2,$$

where ξ_1, ξ_2, and ξ_3 are the roots of the cubic polynomial $f_3(z)$. Since the coefficients of $f_3(z)$ are all real numbers, the discriminant is positive if and only if $f_3(z)$ has three distinct real roots. To see this, observe that the discriminant is zero if and only if a root is repeated. Assuming all roots are real and distinct, clearly the expression is positive. On the other hand, if there is one real root and two complex roots, without loss of generality, suppose that $\xi_1 = r$ is the real root and

$\xi_2, \xi_3 = a \pm bi$ are the complex roots, which occur in pairs because the coefficients of the polynomial are real. Then

$$(\xi_1 - \xi_2)^2(\xi_1 - \xi_3)^2(\xi_2 - \xi_3)^2 = (r - a - bi)^2(r - a + bi)^2(2bi)^2$$
$$= -4b^2(a^2 + b^2 - 2ar + r^2)^2,$$

which is negative. By Lemma 2.4.2, the probability of a repeated root tends to 0 as $N \to \infty$, so it suffices to calculate the probability of the discriminant being positive. Dividing (2.6) by N^4, notice that $-4a_1^3/N^4, -27a_0^2/N^4$, and $18a_2a_1a_0/N^4$ converge in probability to 0 as $N \to \infty$. By Theorem 2.2.5,

$$(a_2/N)^2(a_1/N)^2 - 4a_1^3/N^4 - 4(a_2/N)^3(a_0/N) - 27a_0^2/N^4 + 18a_2a_1a_0/N^4$$

and

$$(a_2/N)^2(a_1/N)^2 - 4(a_2/N)^3(a_0/N)$$

converge in distribution to the same limit.

If $0 \notin [\alpha_1 N, \beta_1 N]$, we have that

$$\mathbb{P}\left((a_2/N)^2(a_1/N)^2 - 4(a_2/N)^3(a_0/N) > 0\right) = \mathbb{P}\left(1/4 > \frac{(a_2/N)(a_0/N)}{(a_1/N)^2}\right).$$

By the independence of a_0, a_1, and a_2, and since $a_0/N \overset{\mathrm{d}}{\Rightarrow} U(\alpha_0, \beta_0), a_1/N \overset{\mathrm{d}}{\Rightarrow} U(\alpha_1, \beta_1)$, and $a_2/N \overset{\mathrm{d}}{\Rightarrow} U(\alpha_2, \beta_2)$, by Theorem 2.2.3 and Lemma 2.2.4,

$$\lim_{N\to\infty} \mathbb{P}\left(1/4 > \frac{(a_2/N)(a_0/N)}{(a_1/N)^2}\right) = \mathbb{P}\left(1/4 > \frac{XY}{Z^2}\right),$$

where $X \sim U(\alpha_0, \beta_0), Y \sim U(\alpha_2, \beta_2)$ and $Z \sim U(\alpha_1, \beta_1)$ are independent.

If $0 \in [\alpha_1 N, \beta_1 N]$, we condition on a_1 and see that

$$
\begin{aligned}
&\mathbb{P}\left((a_2/N)^2(a_1/N)^2 - 4(a_2/N)^3(a_0/N) > 0\right)\\
=&\mathbb{P}\left(1/4 > \frac{(a_2/N)(a_0/N)}{(a_1/N)^2} \mid a_1 \neq 0\right)\mathbb{P}(a_1 \neq 0)\\
&+ \mathbb{P}\left(-4(a_2/N)^3(a_0/N) > 0\right)\mathbb{P}(a_1 = 0)\\
=&\mathbb{P}\left(1/4 > \frac{(a_2/N)(a_0/N)}{(a_1/N)^2} \mid a_1 \neq 0\right)\left(\frac{\lfloor \beta_1 N\rfloor - \lceil \alpha_1 N\rceil}{\lfloor \beta_1 N\rfloor - \lceil \alpha_1 N\rceil + 1}\right)\\
&+ \mathbb{P}\left(-4(a_2/N)^3(a_0/N) > 0\right)\left(\frac{1}{\lfloor \beta_1 N\rfloor - \lceil \alpha_1 N\rceil + 1}\right).
\end{aligned} \tag{2.7}
$$

Clearly $0 \leq \mathbb{P}\left(-4(a_2/N)^3(a_0/N) > 0\right) \leq 1$. Using Remark 2.2.2 and arguing similarly to previous case, we see that

$$\lim_{N\to\infty} \mathbb{P}\left(1/4 > \frac{(a_2/N)(a_0/N)}{(a_1/N)^2} \mid a_1 \neq 0\right) = \mathbb{P}\left(1/4 > \frac{XY}{Z^2}\right),$$

where $X \sim U(\alpha_0, \beta_0), Y \sim U(\alpha_2, \beta_2)$ and $Z \sim U(\alpha_1, \beta_1)$ are independent. Thus, letting $N \to \infty$ in (2.7) proves the theorem. $\square$

Our next example obtains a numerical value for $\mathbb{P}\left(1/4 > \frac{XY}{Z^2}\right)$, when $X, Y, Z \sim U(0,1)$ are jointly independent. This corresponds to precisely the probability that the polynomial $z^3 + a_2 z^2 + a_1 z + a_0$, with a_0, a_1, a_2 uniformly distributed on $[0, N] \cap \mathbb{Z}$ and jointly independent has all real roots, as $N \to \infty$.

Example 2.4.4. We now compute $\mathbb{P}\left(1/4 \geq XY/Z^2\right)$ in the case that $X, Y, Z \sim U(0,1)$ are jointly independent. Springer [149] shows how to compute both XY and X/Z in the special case that $X, Y, Z \sim U(0,1)$ are jointly independent; Curtiss [37] also finds the distribution of a quotient using advanced techniques. We take a simple geometric approach.

We start by considering $\mathbb{P}(XY \leq z)$. Observe that if $z \leq 0$, $\mathbb{P}(XY \leq z) = 0$; if $z \geq 1$,

$\mathbb{P}(XY \leq z) = 1$. Otherwise, for $z \in (0,1)$, we have

$$\begin{aligned}\mathbb{P}(XY \leq z) &= \mathbb{P}(Y \leq z/X) \\ &= \int_0^z 1 \, dx + \int_z^1 \frac{z}{x} \, dx \\ &= z - z\log(z),\end{aligned}$$

where we have used the visual aid in Figure 2.3 to identify the correct region of integration.

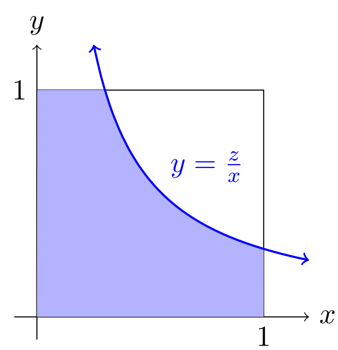

Figure 2.3: This image shows the area below the curve $y = z/x$ for some $z \in (0,1)$ that is also contained in the rectangle $[0,1] \times [0,1]$.

Differentiating yields the probability density function

$$f_{XY}(z) = \begin{cases} -\log(z), & z \in (0,1) \\ 0, & \text{otherwise} \end{cases}.$$

Moreover, for $Z \sim U(0,1)$, observe that if $z \leq 0$, $\mathbb{P}\left(Z^2 \leq z\right) = 0$; if $z \geq 1$, $\mathbb{P}\left(Z^2 \leq z\right) = 1$.

For $z \in (0,1)$, we have that

$$\begin{aligned}\mathbb{P}(Z^2 \leq z) &= \mathbb{P}(Z \leq \sqrt{z}) \\ &= \sqrt{z}.\end{aligned}$$

Differentiating yields the probability density function

$$f_{Z^2}(z) = \begin{cases} \frac{1}{2\sqrt{z}}, & z \in (0,1) \\ 0, & \text{otherwise} \end{cases}.$$

Finally, let $S = XY$ and let $T = Z^2$, and let $f(s,t)$ be the joint probability density function of S and T. Then

$$\begin{aligned}
\mathbb{P}\left(1/4 > \frac{S}{T}\right) &= \int_0^1 \int_0^{1/4t} f(s,t)\ ds\ dt \\
&= \int_0^1 \int_0^{1/4t} \frac{-\log s}{2\sqrt{t}}\ ds\ dt \\
&= \frac{1}{36}(5 + \log(64)) \\
&\approx 0.254413.
\end{aligned}$$

This shows that the probability of $f_3(z)$ having three real roots as $N \to \infty$ is a little more than a fourth.

2.5 An example for degree n

The following theorem due to Kurtz [102] can be used to find some non-trivial examples of integral polynomials of any degree that have positive probability of having all real roots.

Theorem 2.5.1 (Theorem 1 in [102]). *Let $f_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ be a polynomial of degree $n \geq 2$ with positive coefficients. If*

$$a_i^2 - 4a_{i-1}a_{i+1} > 0, \quad i = 1, 2, \ldots, n-1 \tag{2.8}$$

then all the roots of $f_n(z)$ are real and distinct.

We apply the above theorem with $a_n = 1$ to find examples that pertain to monic polynomials.

Example 2.5.2. Let the degree of our monic polynomial be n. Then rearranging (2.8), to create a polynomial with probability one of having all real roots, we need the system of inequalities

$$\begin{cases} 0 < & a_{n-1} \\ 0 < & a_{n-2} < \frac{a_{n-1}^2}{4} \\ 0 < & a_{n-3} < \frac{a_{n-2}^2}{4a_{n-1}} \\ & \vdots \\ 0 < & a_1 < \frac{a_2^2}{4a_3} \\ 0 < & a_0 < \frac{a_1^2}{4a_2} \end{cases}$$

to be satisfied. Let us use the notation $a_k \in [\alpha_k, \beta_k]$ for each $0 \leq k \leq n-1$, and assume that each α_k is positive. Then observe that to satisfy the above inequalities, we may use the recursive formula

$$a_k \in \left[\frac{\alpha_{k+1}^2}{8\beta_{k+2}}, \frac{\alpha_{k+1}^2}{4\beta_{k+2}} \right].$$

Starting with $\alpha_{n-1} = 1$ and $\beta_{n-1} = 2$, this means that $a_{n-1} \in [1, 2]$, $a_{n-2} \in [1/8, 1/4], a_{n-3} \in [1/1024, 1/512]$, and so forth. Notice that we may also multiply all of these coefficients by a large constant $C > 0$. Then

$$0 < Ca_{n-1},$$

$$0 < Ca_{n-2} < \frac{C^2 a_{n-1}^2}{4}$$

and for $0 < k < n-2$,

$$0 < Ca_k < \frac{C^2 a_{k+1}^2}{Ca_{k+2}}.$$

This is useful because multiplying by C allows us to create examples of integral polynomials with probability one of having all real roots, without zero being contained in any of the intervals for the coefficients. This contrasts with Theorem 3.5.18, that shows if zero is in all of the intervals, we always have positive probability of having all real roots. For example, letting $n = 4$ and $C = 2097152$, we see that the monic polynomial

$$z^4 + a_3 z^3 + a_2 z^2 + a_1 z + a_0$$

with independent coefficients

$$a_3 \sim U([2097152, 4194304] \cap \mathbb{Z}), \quad a_2 \sim U([262144, 524288] \cap \mathbb{Z}),$$
$$a_1 \sim U([2048, 4096] \cap \mathbb{Z}), \quad a_0 \sim U([1, 2] \cap \mathbb{Z}),$$

has probability one of having all real roots.

2.6 Discovering an extremely dominant root

In this section, we show that as $N \to \infty$, the random polynomial

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

whose coefficients obey the generalized large box model has an extremely dominant root, which we will call ξ_n and define shortly. We first find this root for the quadratic polynomial, and then use Rouché's theorem to locate the root for higher degree polynomials. Rouché's theorem is a common approach for locating polynomial roots, see for instance [24, 35, 141, 153, 168]. We then describe the distribution of the distance between the coefficient $-a_{n-1}$ and the root ξ_n, and compute the probability density function of this distribution explicitly in certain cases. We note that the root ξ_n is the dominant root of the integer polynomials discovered by Dubickas and Sha in [45], but the location and fluctuations of ξ_n is a new discovery to the best of our knowledge.

2.6.1 Extremely dominant root for degree two generalized large box model polynomials

For $0 \leq i \leq 1$, let $\alpha_i < \beta_i \in \mathbb{R}$. Let $f_2(z) = z^2 + a_1 z + a_0$, where for each $0 \leq i \leq 1$, a_i is uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ with a_0 and a_1 independent. From Theorem 2.1.2, we know that this polynomial has two real roots with probability tending to 1 as $N \to \infty$. Consider the root

$$\xi_2 := \begin{cases} \frac{-a_1 - \operatorname{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} & \text{if } a_1 \neq 0 \\ \sqrt{-a_0} & \text{if } a_1 = 0 \end{cases},$$

which may be real or imaginary. We call a root of a random polynomial **extremely dominant** if its modulus is not tight, while the moduli of all of the other roots are tight. Details are given in Lemma 2.6.3 showing that ξ_2 is the extremely dominant root of $f_2(z)$.

Let Z_N be the distance between the coefficient $-a_1$ and the root ξ_2. We now have the tools necessary to compute the distribution of Z_N in the quadratic case. This is the content of Theorem 2.1.4, whose proof is given below.

Proof of Theorem 2.1.4. We wish to find the distribution of the Euclidean distance between ξ_2 and $-a_1$, which is given by

$$Z_N = \begin{cases} \left|\frac{-a_1 - \mathrm{sgn}(a_1)\sqrt{a_1^2-4a_0}}{2} + a_1\right| & \text{if } a_1 \neq 0 \\ \sqrt{|a_0|} & \text{if } a_1 = 0 \end{cases}.$$

To show a sequence of distribution functions F_n converge in distribution to the limiting distribution function F, we must show that the convergence holds at all continuity points of F. In Appendix A, we find the pdf of the random variable X/Y, which implies that its cumulative distribution function (and also that of $|X/Y|$) is continuous.

If $0 \notin [\alpha_1 N, \beta_1 N]$, we have that

$$\begin{aligned}
&\left|\frac{-a_1 - \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} + a_1\right| \\
&= \left|\frac{a_1 - \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2}\right| \\
&= \left|\frac{\left(a_1 - \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}\right)}{2} \cdot \frac{\left(a_1 + \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}\right)}{\left(a_1 + \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}\right)}\right| \\
&= \left|\frac{a_1^2 - (a_1^2 - 4a_0)}{2\left(a_1 + \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}\right)}\right| \\
&= \left|\frac{2a_0}{a_1 + \mathrm{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}\right| \\
&= \left|\frac{a_0}{a_1\left(1/2\left(1 + \sqrt{1 - 4a_0/a_1^2}\right)\right)}\right|.
\end{aligned}$$

Multiplying by N/N gives that

$$Z_N = \left| \frac{a_0/N}{(a_1/N)\left(1/2\left(1+\sqrt{1-4a_0/a_1^2}\right)\right)} \right|.$$

We want to show $Z_N \overset{\text{d}}{\Rightarrow} |X/Y|$, where $X \sim U(\alpha_0, \beta_0)$, $Y \sim U(\alpha_1, \beta_1)$ and X and Y are independent. By Lemma 2.2.8, a_0/a_1^2 converges in probability to 0, so by Theorem 2.2.3,

$$\frac{1}{1/2\left|1+\sqrt{1-4a_0/a_1^2}\right|} \tag{2.9}$$

converges in probability to 1. By Lemma 2.2.1, $a_0/N \overset{\text{d}}{\Rightarrow} U(\alpha_0, \beta_0)$ and $a_1/N \overset{\text{d}}{\Rightarrow} U(\alpha_1, \beta_1)$. By Theorem 2.2.3, $\left|\frac{1}{a_1/N}\right| \overset{\text{d}}{\Rightarrow} |1/U(\alpha_1, \beta_1)|$. By Lemma 2.2.4 and the independence of a_0 and a_1,

$$\left|\frac{a_0/N}{a_1/N}\right| \overset{\text{d}}{\Rightarrow} |X/Y|, \quad X \sim U(\alpha_0, \beta_0), Y \sim U(\alpha_1, \beta_1),$$

where X and Y are independent. Theorem 2.2.5 allows us to multiply by (2.9) and conclude that $Z_N \overset{\text{d}}{\Rightarrow} |X/Y|$.

If $0 \in [\alpha_1 N, \beta_1 N]$, we condition on a_1. For any $z \in \mathbb{R}$, we have by the law of total probability that

$$\begin{aligned}
\mathbb{P}(Z_N \le z) &= \mathbb{P}(Z_N \le z \mid a_1 \ne 0)\mathbb{P}(a_1 \ne 0) + \mathbb{P}(Z_N \le z \mid a_1 = 0)\mathbb{P}(a_1 = 0) \\
&= \mathbb{P}\left(\left|\frac{-a_1 - \operatorname{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} + a_1\right| \le z \mid a_1 \ne 0\right)\left(\frac{\lceil \beta_1 N\rceil - \lfloor \alpha_1 N \rfloor}{\lceil \beta_1 N\rceil - \lfloor \alpha_1 N \rfloor + 1}\right) \\
&\quad + \mathbb{P}\left(\sqrt{|a_0|} \le z\right)\left(\frac{1}{\lceil \beta_1 N\rceil - \lfloor \alpha_1 N \rfloor + 1}\right).
\end{aligned} \tag{2.10}$$

For any $z \in \mathbb{R}$, clearly $0 \le \mathbb{P}\left(\sqrt{|a_0|} \le z\right) \le 1$. Using Remark 2.2.2 and arguing similarly to the previous case, we find that for any $z \in \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{-a_1 - \operatorname{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} + a_1\right| \le z \mid a_1 \ne 0\right) = \mathbb{P}\left(\left|\frac{X}{Y}\right| \le z\right),$$

where $X \sim U(\alpha_0, \beta_0)$, $Y \sim U(\alpha_1, \beta_1)$ and X and Y are independent. Taking the limit as $N \to \infty$ in (2.10) proves the theorem. □

We now define tightness and introduce Prohorov's theorem, which will allow us to quickly prove that the root ξ_2 from Theorem 2.1.4 is extremely dominant.

Definition 2.6.1 (Chapter 2 in [163]). A set of random vectors $\{X_\alpha : \alpha \in A\}$ is called **tight** if for every $\epsilon > 0$, there exists a constant M such that

$$\sup_{\alpha} \mathbb{P}(\|X_a\| > M) < \epsilon.$$

Theorem 2.6.2 (Prohorov's theorem, Theorem 2.4 in [163]). *Let X_n be random vectors in $\mathbb{R}^k$.*

- *If $X_n \overset{d}{\Rightarrow} X$ for some X, then $\{X_n : n \in \mathbb{N}\}$ is tight;*
- *If X_n is tight, then there exists a subsequence with $X_{n_j} \overset{d}{\Rightarrow} X$ as $j \to \infty$, for some X.*

We now prove that the root ξ_2 is extremely dominant.

Lemma 2.6.3. *For $0 \le i \le 1$, let $\alpha_i < \beta_i \in \mathbb{R}$. Let $f_2(z) = z^2 + a_1 z + a_0$, where for each $0 \le i \le 1$, a_i is uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with a_0 and a_1 independent. Denoting the roots of $f_2(z)$ by ξ_2 as defined in Theorem 2.1.4 and $\tilde{\xi_2}$ for the conjugate root, we have that the modulus of ξ_2 is not tight as $N \to \infty$, whereas the modulus of $\tilde{\xi_2}$ is tight as $N \to \infty$.*

Proof. Modifying the proof of Theorem 2.1.4 slightly, we see that the modulus of the conjugate root

$$\tilde{\xi_2} = \begin{cases} \frac{-a_1 + \operatorname{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} & \text{if } a_1 \neq 0 \\ -\sqrt{-a_0} & \text{if } a_1 = 0 \end{cases}$$

converges in distribution to $|X/Y|$, where $X \sim U(\alpha_0, \beta_0)$, $Y \sim U(\alpha_1, \beta_1)$. Theorem 2.6.2 immediately implies that $\left\{\left|\tilde{\xi_2}\right| : N \in \mathbb{N}\right\}$ is tight.

On the other hand, if $a_1 \neq 0$,

$$\begin{aligned} |\xi_2| &= \left| \frac{-a_1 - \operatorname{sgn}(a_1)\sqrt{a_1^2 - 4a_0}}{2} \right| \\ &= |a_1| \left| \frac{1 + \sqrt{1 - \frac{4a_0}{a_1^2}}}{2} \right| \\ &\ge \frac{|a_1|}{2}. \end{aligned}$$

Therefore, if $0 \notin [\alpha_1 N, \beta_1 N]$, we have that for any positive $M \in \mathbb{R}$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}(|\xi_2| > M) &\geq \lim_{N\to\infty} \mathbb{P}\left(\frac{|a_1|}{2} > M\right)\\ &= 1.\end{aligned}$$

If $0 \in [\alpha_1 N, \beta_1 N]$, we have that for any positive $M \in \mathbb{R}$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(|\xi_2| > M\right) &\geq \lim_{N\to\infty} \mathbb{P}\left(|\xi_2| > M \mid a_1 \neq 0\right)\mathbb{P}(a_1 \neq 0)\\ &\geq \lim_{N\to\infty} \mathbb{P}\left(\frac{|a_1|}{2} > M\right)\left(\frac{\lceil \beta_1 N\rceil - \lfloor \alpha_1 N\rfloor}{\lceil \beta_1 N\rceil - \lfloor \alpha_1 N\rfloor + 1}\right)\\ &= 1.\end{aligned}$$

In either case, this shows that $\{|\xi_2| : N \in \mathbb{N}\}$ is not tight. □

2.6.2 Extremely dominant root for degree n polynomials

Interestingly, a degree n polynomial whose coefficients obey the generalized large box model also has an extremely dominant root. This root is well approximated by the analogous dominant root from the degree two generalized large box model polynomial. The distribution of this root is the content of this section.

Lemma 2.6.4. *Let $\epsilon > 0$ and $z_0 \in \mathbb{C}$. Then for any $z \in C_{z_0,\epsilon}$,*

$$|z_0| - \epsilon \leq |z| \leq |z_0| + \epsilon.$$

Proof. This follows from the triangle inequality. On one hand, we have that

$$|z| = |z - z_0 + z_0| \leq |z - z_0| + |z_0| = \epsilon + |z_0|.$$

On the other hand,

$$|z_0| = |z_0 - z + z| \leq |z_0 - z| + |z| = \epsilon + |z|$$

so

$$|z_0| - \epsilon \leq |z|.$$

□

Theorem 2.6.5 (Rouché's Theorem, 3.42 in [161])**.** *If $f(z)$ and $g(z)$ are analytic inside and on a closed contour C and*

$$|f(z) - g(z)| < |g(z)|, \quad \textit{for all } z \in C,$$

then $f(z)$ and $g(z)$ have the same number of zeros inside C.

Theorem 2.1.5 states that random polynomials whose coefficients obey the generalized large box model have a real root very close to the extremely dominant root for $f_2(z)$ described in Section 2.6.1, as $N \to \infty$. We will also refer to this root as extremely dominant, but the proof of this is delayed until Chapter 3; the result is given in Theorem 3.3.4.

We now present the proof of Theorem 2.1.5.

Proof of Theorem 2.1.5. Let us begin with the case where $0 \notin [\alpha_{n-1}N, \beta_{n-1}N]$ and let $\epsilon = \frac{C_{\alpha,\beta,n}}{N}$, where

$$C_{\alpha,\beta,n} = \frac{\sum_{i=0}^{n-3} \max\{|\alpha_i|, |\beta_i|\}}{(\min\{|\alpha_{n-1}|, |\beta_{n-1}|\})^2}.$$

By the quadratic formula, the two non-zero roots of $f(z)$ can be written as

$$\frac{-a_{n-1} \pm \sqrt{a_{n-1}^2 - 4a_{n-2}}}{2}.$$

From the statement of the theorem, since $0 \notin [\alpha_{n-1}N, \beta_{n-1}N]$, we have that

$$\xi = \frac{-a_{n-1} - \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2}.$$

Therefore, let the second non-zero root of $f(z)$ be given by

$$\zeta = \frac{-a_{n-1} + \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2}.$$

Observe that the roots of $f(z)$ are zero with multiplicity $n-2$, ξ with multiplicity one, and ζ with multiplicity one. We begin by showing that the ball $B(\xi, \epsilon)$ contains only the root ξ of $f(z)$ with probability tending to 1 as $N \to \infty$. We have

$$|\xi - \zeta| = |a_{n-1}| \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right|. \tag{2.11}$$

and therefore

$$\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} = \frac{|a_{n-1}|}{N}\left|\sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right| - \frac{\epsilon}{N}.$$

By Lemma 2.2.1 and Theorem 2.2.3, $\frac{|a_{n-1}|}{N}$ converges in distribution to $|U(\alpha_{n-1}, \beta_{n-1})|$. By Lemma 2.2.8, a_{n-2}/a_{n-1}^2 converges in probability to zero, and so by Theorem 2.2.3, $\left|\sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right|$ converges in probability to one. Finally, in both cases, $\frac{\epsilon}{N}$ converges to zero. Then, by two applications of Theorem 2.2.5, we have that

$$\frac{|a_{n-1}|}{N}\left|\sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right| - \frac{\epsilon}{N} \stackrel{d}{\Longrightarrow} |U(\alpha_{n-1}, \beta_{n-1})|.$$

Altogether, this shows that

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(|\xi - \zeta| > \epsilon\right) &= \lim_{N\to\infty} \mathbb{P}\left(\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} > 0\right)\\ &= \mathbb{P}\left(|U(\alpha_{n-1}, \beta_{n-1})| > 0\right)\\ &= 1,\end{aligned}$$

since 0 is a continuity point of the cdf of $|U(\alpha_{n-1}, \beta_{n-1})|$. Thus, $\zeta \notin B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$,

By a similar argument, we can show that

$$\lim_{N\to\infty} \mathbb{P}\left(|\xi| > \epsilon\right) = \mathbb{P}\left(|U(\alpha_{n-1}, \beta_{n-1})| > 0\right) = 1, \tag{2.12}$$

implying that $0 \notin B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. We conclude that $f(z)$ has exactly one root in $B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$.

When $n = 2$, the polynomials $f_2(z)$ and $f(z)$ are equal, so have the same roots. In particular, Theorem 2.1.2 shows that both roots of these polynomials are real with probability tending to 1 as $N \to \infty$. Suppose now that $n \geq 3$. We wish to show that the inequality

$$|f_n(z) - f(z)| = |a_{n-3}z^{n-3} + \cdots + a_1 z + a_0| < |z|^{n-2}|z^2 + a_{n-1}z + a_{n-2}| = |f(z)| \tag{2.13}$$

holds for all $z \in \mathcal{C}_{\xi,\epsilon}$ with probability tending to 1 as $N \to \infty$. If (2.13) holds, Rouché's theorem tells us that $f_n(z)$ and $f(z)$ both have exactly one root in $B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. By Theorem 2.1.2, the root ξ is a real root of $f(z)$ with probability tending to 1 as $N \to \infty$. In particular, this means the root ξ_n of $f_n(z)$ captured by $B(\xi, \epsilon)$ must be real with probability tending to 1 as $N \to \infty$ as well. To see this, recall that complex roots of polynomials with real coefficients occur in complex conjugate pairs. Since the ball $B(\xi, \epsilon)$ is centered on the real axis (with probability tending to 1 as $N \to \infty$) and captures exactly one root ξ_n of $f_n(z)$, then if ξ_n is complex, $B(\xi, \epsilon)$ would also capture the complex conjugate of ξ_n. This is a contradiction to the number of roots of $f_n(z)$ in $B(\xi, \epsilon)$.

For $z \neq 0$, we may divide both sides of (2.13) by $|z|^{n-3}$ to obtain

$$\left| a_{n-3} + \frac{a_{n-4}}{z} + \cdots + \frac{a_1}{z^{n-2}} + \frac{a_0}{z^{n-3}} \right| < |z||z^2 + a_{n-1}z + a_{n-2}|. \tag{2.14}$$

Since $0 \notin \mathcal{C}_{\xi,\epsilon}$ with probability tending to 1 as $N \to \infty$, it suffices to show that (2.14) holds for all $z \in \mathcal{C}_{\xi,\epsilon}$ with probability tending to 1 as $N \to \infty$. Factoring the right-hand side of (2.14) gives

$$|z| \left| z - \frac{-a_{n-1} - \sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} \right| \left| z - \frac{-a_{n-1} + \sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} \right| \tag{2.15}$$

which is equal to

$$|z| \, |z - \xi| \, |z - \zeta| \, .$$

We will bound each of the terms on the right-hand side of (2.15) from below. Let Ω be the event that ξ is a real root of $f(z)$. Let $z \in \mathcal{C}_{\xi,\epsilon}$. By Lemma 2.6.4, on Ω,

$$|z| \geq |\xi| - \epsilon = \frac{|a_{n-1}| \left| 1 + \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right|}{2} - \epsilon.$$

We also have that

$$|z - \xi| = \epsilon.$$

Finally, on Ω, the conjugate root ζ of $f(z)$ is also real. By the reverse triangle inequality,

$$\begin{aligned} |z - \zeta| &= |(z - \xi) + (\xi - \zeta)| \\ &\geq ||z - \xi| - |\xi - \zeta|| \\ &= \left| \epsilon - |a_{n-1}| \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|, \end{aligned}$$

where we have used the computation of $|\xi - \zeta|$ given by (2.11). Hence, a lower bound for (2.15) is given by

$$\left(\frac{|a_{n-1}| \left|1 + \sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right|}{2} - \epsilon \right) \epsilon \left| \epsilon - |a_{n-1}| \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|$$

which we may factor into

$$a_{n-1}^2 \left(\frac{\left|1 + \sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \epsilon \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|. \tag{2.16}$$

Hence, in order to show that (2.14) holds, we wish to show that

$$\left| \sum_{i=0}^{n-3} \frac{a_i}{z^{n-3-i}} \right| \tag{2.17}$$

is strictly less than (2.16) for all $z \in \mathcal{C}_{\xi,\epsilon}$ with probability tending to 1 as $N \to \infty$. Similarly to (2.12), we can show that $|z| \geq 1$ with probability tending to 1 as $N \to \infty$, so an upper bound for the left hand side of (2.17) on Ω is $\sum_{i=0}^{n-3} |a_i|$. Then dividing by everything in (2.16) except for ϵ, it suffices to show

$$\lim_{N \to \infty} \mathbb{P} \left(\frac{\sum_{i=0}^{n-3} |a_i|}{a_{n-1}^2 \left(\frac{\left|1 + \sqrt{1 - \frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|} < \epsilon \right) = 1. \tag{2.18}$$

Multiplying both sides by N, we consider

$$\mathbb{P} \left(\frac{N \sum_{i=0}^{n-3} |a_i|}{a_{n-1}^2 \left(\frac{\left|1 + \sqrt{1 - \frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|} < C_{\alpha,\beta,n} \right). \tag{2.19}$$

Clearly

$$\frac{\epsilon}{|a_{n-1}|} = \frac{C_{\alpha,\beta,n}}{N|a_{n-1}|}$$

converges in probability to zero as $N \to \infty$, so by Lemma 2.6 and Theorem 2.2.3,

$$\left(\frac{\left|1 + \sqrt{1 - \frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|$$

converges in probability to one as $N \to \infty$.

Since $a_0, \ldots, a_{n-3}, a_{n-1}$ are independent, by Lemma 2.2.1, Theorem 2.2.3, and repeated application of Lemma 2.2.4 (which is really just a multivariate version of the continuous mapping theorem), we have that

$$\frac{N \sum_{i=0}^{n-3} |a_i|}{a_{n-1}^2} = \frac{N^2 \sum_{i=0}^{n-3} |a_i|}{N a_{n-1}^2} \stackrel{\mathrm{d}}{\Longrightarrow} \frac{\sum_{i=0}^{n-3} |X_i|}{X_{n-1}^2},$$

where each X_i is independently and uniformly distributed on $[\alpha_i, \beta_i]$ for $0 \le i \le n-3$ and $i = n-1$. By Theorem 2.2.5,

$$\frac{\sum_{i=0}^{n-3} |a_i|}{a_{n-1}^2 \left(\frac{\left|1 + \sqrt{1 - \frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|}$$

also converges in distribution to

$$\frac{\sum_{i=0}^{n-3} |X_i|}{X_{n-1}^2}, \tag{2.20}$$

so (2.19) tends to 1 as $N \to \infty$, since $C_{\alpha,\beta,n}$ was chosen to be the maximum value of the limiting distribution.

If $0 \in [\alpha_{n-1}N, \beta_{n-1}N]$, the proof is very similar. The key differences are that we fix $0 < \epsilon < 1$ since the limiting distribution given by (2.20) is no longer bounded and that we condition on the value of a_{n-1}. We elaborate on a few steps of the proof now. As before, let us denote the non-zero roots of $f(z)$ by

$$\xi = \begin{cases} \frac{-a_{n-1} - \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} & \text{if } a_{n-1} \neq 0 \\ \sqrt{-a_{n-2}} & \text{if } a_{n-1} = 0. \end{cases}$$

and

$$\zeta = \begin{cases} \frac{-a_{n-1} + \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} & \text{if } a_{n-1} \neq 0 \\ -\sqrt{-a_{n-2}} & \text{if } a_{n-1} = 0. \end{cases}$$

We have

$$|\xi - \zeta| = \begin{cases} |a_{n-1}| \left|\sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right| & \text{if } a_{n-1} \neq 0 \\ 2\sqrt{|a_{n-2}|} & \text{if } a_{n-1} = 0 \end{cases}. \tag{2.21}$$

Fix $0 < \epsilon < 1$. Conditioning on a_{n-1},

$$\begin{aligned} \mathbb{P}\left(|\xi - \zeta| > \epsilon\right) &= \mathbb{P}\left(\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} > 0\right) \\ &\geq \mathbb{P}\left(\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} > 0 \mid a_{n-1} \neq 0\right) \mathbb{P}(a_{n-1} \neq 0) \\ &= \mathbb{P}\left(\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} > 0 \mid a_{n-1} \neq 0\right) \left(\frac{\lfloor \beta_{n-1} N \rfloor - \lceil \alpha_{n-1} N \rceil}{\lfloor \beta_{n-1} N \rfloor - \lceil \alpha_{n-1} N \rceil + 1}\right). \end{aligned} \tag{2.22}$$

Using Remark 2.2.2 and arguing similarly to the previous case,

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{|\xi - \zeta|}{N} - \frac{\epsilon}{N} > 0 \mid a_{n-1} \neq 0\right) = \mathbb{P}\left(|U(\alpha_{n-1}, \beta_{n-1})| > 0\right) = 1.$$

Letting $N \to \infty$ in (2.22) shows that

$$\lim_{N\to\infty} \mathbb{P}(|\xi - \zeta| > \epsilon) = 1,$$

so $\zeta \notin B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. Similarly, and again conditioning on a_{n-1}, we can show

$$\lim_{N\to\infty} \mathbb{P}\left(|\xi| > \epsilon\right) = \mathbb{P}\left(|U(\alpha_{n-1}, \beta_{n-1})| > 0\right) = 1, \tag{2.23}$$

implying that $0 \notin B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. We conclude that $f(z)$ has exactly one root in $B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$.

Arguing as in the previous case, we wish to show that (2.14) holds for all $z \in \mathcal{C}_{\xi,\epsilon}$ with probability tending to 1 as $N \to \infty$. Conditioning on a_{n-1}, which holds with probability tending to 1 as $N \to \infty$, it suffices to show that

$$\mathbb{P}\left(\frac{\sum_{i=0}^{n-3} |a_i|}{a_{n-1}^2 \left(\frac{\left|1+\sqrt{1-\frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|}\right) \left|\frac{\epsilon}{|a_{n-1}|} - \left|\sqrt{1 - 4a_{n-2}/a_{n-1}^2}\right|\right|} < \epsilon \mid a_{n-1} \neq 0\right) \tag{2.24}$$

tends to 1 as $N \to \infty$.

Multiplying the inside of (2.24) by N^2/N^2, using Remark 2.2.2, and making a similar argument to the previous case, we now see that

$$\frac{N^2 \sum_{i=0}^{n-3} |a_i|}{N^2 a_{n-1}^2 \left(\frac{\left|1+\sqrt{1-\frac{4a_{n-2}}{a_{n-1}^2}}\right|}{2} - \frac{\epsilon}{|a_{n-1}|} \right) \left| \frac{\epsilon}{|a_{n-1}|} - \left| \sqrt{1 - 4a_{n-2}/a_{n-1}^2} \right| \right|} \overset{\mathrm{d}}{\Rightarrow} 0.$$

Since convergence in distribution to a constant implies convergence in probability to the same constant, we see that (2.24) tends to 1 as $N \to \infty$. □

Next, we show that the extremely dominant root of degree n also deviates from $-a_{n-1}$ in a manner that converges in distribution to the ratio of two continuous uniform random variables. This is the content of Corollary 2.1.6, which we prove now.

Proof of Corollary 2.1.6. Let ξ be defined as in Theorem 2.1.5. Theorem 2.1.5 says that if $0 \in [\alpha_{n-1}N, \beta_{n-1}N]$, then for any $\epsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(|\xi - \xi_n| > \epsilon\right) = 0.$$

When $0 \notin [\alpha_{n-1}N, \beta_{n-1}N]$, the proof of Theorem 2.6.5 considered only $\epsilon = O(1/N)$. Fixing $0 < \epsilon < 1$ instead and carrying out the proof as in the other case, we see that for any $0 < \epsilon < 1$,

$$\lim_{N\to\infty} \mathbb{P}\left(|\xi - \xi_n| > \epsilon\right) = 0.$$

Therefore, in either case, for any $\epsilon > 0$, we have that

$$\begin{aligned} \lim_{N\to\infty} \mathbb{P}\left(|(\xi + a_{n-1}) - (\xi_n + a_{n-1})| > \epsilon\right) &= \lim_{N\to\infty} \mathbb{P}\left(|\xi - \xi_n| > \epsilon\right) \\ &= 0. \end{aligned}$$

In other words, $|(\xi + a_{n-1}) - (\xi_n + a_{n-1})| \overset{p}{\to} 0$ as $N \to \infty$. In Theorem 2.1.4 (with a relabeling of the coefficients), we showed that $|\xi + a_{n-1}| \overset{d}{\Rightarrow} |X/Y|$. Therefore, invoking part (4) of Theorem 2.2.6, we conclude that $|\xi_n + a_{n-1}| \overset{d}{\Rightarrow} |X/Y|$ as well. □

In Appendix A, we investigate the probability density function of the limiting distribution of $|Z_N|$.

Chapter 3

Generalized bounded height polynomials

3.1 Introduction and main results

In this chapter, we show that $n-1$ of the roots of monic integral polynomials whose coefficients are uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$ for $0 \le i \le n-1$ and jointly independent can be approximated by the roots of degree $n-1$ random polynomials whose coefficients are uniformly distributed on $[\alpha_i, \beta_i]$ for $0 \le i \le n-1$ and jointly independent. (The remaining root is described by the extremely dominant root discovered in Chapter 2.) We will call this random polynomial model the **generalized bounded height model**; the usual bounded height model considers all coefficients in the intervals $[-N, N]$ for some fixed $N > 0$.

We begin by showing the equivalence of these cases, arguing that as $N \to \infty$, for a fixed degree, the probability that the generalized large box model polynomials has all real roots converges to the probability that generalized bounded height polynomials have all real roots. This is summarized by the following theorem.

Theorem 3.1.1. *Let $\alpha_j < \beta_j \in \mathbb{R}$ for each $0 \le j \le n-1$. Let $N \in \mathbb{N}$ and let a_j be uniformly distributed on $[\alpha_j N, \beta_j N] \cap \mathbb{Z}$ for each $0 \le j \le n-1$ with $a_0, \dots, a_{n-1}$ jointly independent. Let b_j be uniformly distributed on $[\alpha_j, \beta_j]$ for each $0 \le j \le n-1$ with $b_0, \dots, b_{n-1}$ jointly independent. Define*

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

and

$$g(z) = b_{n-1}z^{n-1} + \cdots + b_1 z + b_0.$$

Then as $N \to \infty$, the probability that f has all real roots converges to the probability that g has all real roots.

Coupling the coefficients of $f(z)$ and $g(z)$ as in Lemma 3.3.2 and Remark 3.3.3, we refer to the coupled polynomials as $f'(z)$ and $g'(z)$. Order the roots $\xi_1', \ldots, \xi_{n-1}'$ of $g'(z)$ by increasing magnitude and then argument in $[0, 2\pi)$. As $N \to \infty$, when labeled properly, there exists a unique root $\tilde{\xi}_i'$ of $f'(z)$ that converges almost surely to ξ_i', for each $1 \leq i \leq n-1$. The remaining root of $f'(z)$ is the extremely dominant root described in Theorem 2.1.5.

To prove the above theorem, we use a technique called coupling, which is defined in Section 3.3.1. Lemma 3.3.2 and Remark 3.3.3 describe the coupling used between the coefficients of generalized large box model and generalized bounded height polynomials. After coupling the polynomials, we use a Rouché's theorem argument to pair up the roots of the coupled polynomials and count how many are real. An advantage to our argument is that we are able to prove the almost sure convergence of the polynomial roots under the coupling and argue that the extra root of the integral polynomial must be the extremely dominant root discussed in Chapter 2.

Dubickas [44] already discovered this equivalence between the polynomial models when all of the coefficients taking integral values in $[-N, N]$. More precisely, he showed the following.

Theorem 3.1.2 (Theorem 1.2 in [44])**.** *Let $n \geq 1$ and let s be two integers satisfying $0 \leq s < n/2$. Let $D_n(s, N)$ denote the number of monic integral polynomials of degree n with coefficients in $[-N, N] \cap \mathbb{Z}$ with exactly r real roots and $2s$ non-real roots. Let λ^n be the n-th dimensional Lebesgue measure, and let $\mathcal{H}_n(s, N)$ denote the set of $(n+1)$-dimensional vectors $(a_0, \ldots, a_n)$ satisfying $|a_i| \leq N$ for $0 \leq i \leq n$, $a_n \neq 0$, such that the polynomial $a_n z^n + \cdots + a_1 z + a_0$ has exactly s pairs of complex conjugate roots. Then*

$$\lim_{N \to \infty} \frac{D_n(s, N)}{N^n} = \lambda^n(\mathcal{H}_{n-1}(s, 1)).$$

Dubickas proves this by arguing that the polynomial $z^n + b_{n-1}z^{n-1} + \cdots + b_0$ with coefficients $|b_0|, \ldots, |b_{n-1}| \leq N$ has $2s$ non-real roots if and only if $(b_0/N, \ldots, b_{n-1}/N) \in \mathcal{H}_{n-1}(s, 1/N)$. Hence, when restricted to the integers in $[-N, N]$, the set of coefficients $(b_0, \ldots, b_{n-1})$ satisfying the polynomial and giving $2s$ non-real roots is asymptotic to $N^n \lambda^n(\mathcal{H}_{n-1}(s, 1/N))$ as $N \to \infty$. From here, for any $\epsilon > 0$, he bounds $|D_n(s, N)/N^n - \lambda^n(\mathcal{H}_{n-1}(s, 1))| < 2n\epsilon$. To obtain the lower bound, he considers the polynomials $U(z) = a_{n-1}z^{n-1} + \cdots a_1 z + a_0$ where $|a_j| \leq 1$ for each $0 \leq j \leq n-1$ and $U(z)$ has exactly $2s$ non-real roots and the polynomial $V(z) = \frac{1}{N}z^n + U(z)$. Using the Taylor expansion of the polynomials, he is able to show with Rouché's theorem that each have exactly one real root in certain disjoint discs, showing that for any $\epsilon > 0$ and N sufficiently large, $\lambda^n(\mathcal{H}_{n-1}(s, 1/N)) > \lambda^n(\mathcal{H}_{n-1}(s, 1)) - \epsilon$. The upper bound is obtained by counting. An advantage to this proof is the additional knowledge granted by the inequality $\lambda^n(\mathcal{H}_{n-1}(s, 1/N)) > \lambda^n(\mathcal{H}_{n-1}(s, 1)) - \epsilon$.

Lemma 3.3.4 then shows that root ξ_n of $f(z)$ is actually extremely dominant, fulfilling a promise of Chapter 2.

The remainder of the chapter investigates properties of generalized bounded height polynomials. In Section 3.4, we find integral formulas for the probability that a generalized bounded height polynomial has all real roots. These are given by Theorem 3.4.1 for monic generalized bounded height polynomials and by Theorem 3.4.3 for non-monic generalized bounded height polynomials. The proofs of these theorems follow the same outline as those of the analogous results of Bertók, Hadju, and Pethő in [18], who established them for the case where each coefficient is restricted to the interval $[-\ell, \ell]$ for some $\ell \in \mathbb{R}$. The only modifications to our proof arise from allowing the coefficients to vary over different (and non-symmetric) intervals.

In Section 3.5, special cases of coefficients are considered. One such case is when all coefficients of the generalized bounded height polynomial are iid on $[0, 1]$. Theorem 3.5.1 states that the probability of a non-monic generalized bounded height having all real roots is positive, for every degree n. By the equivalence between these polynomials and the generalized large box model polynomials given in Theorem 3.1.1, this implies that for any $n \in \mathbb{N}$, the probability that the polynomial

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

with $a_{n-1}, \ldots, a_1, a_0$ uniformly distributed on $[0, N] \cap \mathbb{Z}$ and jointly independent has all real roots is also positive, as $N \to \infty$. Other exciting results include Lemma 3.5.6 and Proposition 3.5.11, which state that the probability of a monic generalized bounded height polynomial of degree n with coefficients iid on $[0, 1]$ having all real roots is positive and monotonically decreasing in n.

Section 3.5 also considers when $0 \in [\alpha_i, \beta_i]$ for each $1 \leq i \leq n$. In this case, the probability that the non-monic generalized bounded height polynomial has all real roots is also positive.

3.2 Tools and preliminary results for generalized bounded height model

The purpose of this section is to gather the tools that will be needed to prove results related to the generalized bounded height model. One of our main goals is to prove the following key lemma.

Lemma 3.2.1. *Define the polynomial*

$$f_n(z) = a_n z^n + \ldots + a_1 z + a_0,$$

where the $0 \leq a_i \leq n$ are independent, continuously distributed real random variables that are absolutely continuous with respect to the Lebesgue measure. The probability that $f_n(z)$ has a repeated root is zero.

Showing that random polynomials with coefficients obeying the generalized bounded height model have repeated roots with probability zero will let us make assumptions later that will greatly simplify proofs related to this model in future sections. We must first prove an auxiliary result and introduce Vieta's formulas and related notation before we are able to prove the lemma. At the end of the section, we also include other preparatory material.

Lemma 3.2.2. *Let $f(z_1, \ldots, z_n)$ be a non-zero polynomial in n variables. If $X_1, \ldots, X_n$ are independent real random variables with probability distributions $\mu_1, \mu_2, \ldots, \mu_n$ that are absolutely continuous with respect to the Lebesgue measure, then the probability that $f(X_1, \ldots, X_n) = 0$ is zero.*

Proof. [1] The polynomial $f(z_1, \ldots, z_n)$ is a continuous (and hence measurable) function, so the set $\{(z_1, \ldots, z_n) \in \mathbb{R}^n \mid f(z_1, \ldots, z_n) = 0\}$ is measurable for every $n \in \mathbb{N}$. Let $\mathbb{1}_A$ denote the indicator function on a set A. We proceed by induction on n.

When $n = 1$, consider a non-constant degree k polynomial in a single variable z_1, namely $f(z_1)$. By the fundamental theorem of algebra, this polynomial has at most k zeros. Let M be the set of zeros of $f(z_1)$. Then $\lambda(M) = 0$ since M is finite, and $\mu_1(M) = 0$ since μ_1 is absolutely continuous with respect to the Lebesgue measure. Then

$$\mathbb{P}\left(f(X_1) = 0\right) = \mathbb{P}(X_1 \in M) = \mu_1(M) = 0$$

for the non-zero polynomial $f(z_1)$.

As our inductive hypothesis, suppose that the statement holds for non-zero polynomials in $n-1$ variables. In other words, letting

$$M = \{(z_1, \ldots, z_{n-1}) \in \mathbb{R}^n \mid f(z_1, \ldots, z_{n-1}) = 0\},$$

suppose that

$$\begin{aligned} \mathbb{P}\left(f(X_1, \ldots, X_{n-1}) = 0\right) &= \mathbb{P}((X_1, \ldots, X_{n-1}) \in M) \\ &:= \left(\mu_1 \times \mu_2 \times \cdots \times \mu_{n-1}\right)(M) \\ &= 0. \end{aligned}$$

Now consider a non-zero polynomial $f(z_1, \ldots, z_n)$ in the variables $z_1, \ldots, z_n$, and assume that the largest power of z_n appearing is $k \geq 1$. (If z_n does not appear in the non-zero polynomial, simply relabel the variables so that it does.) Then we have

$$f(z_1, \ldots, z_n) = f_k {z_n}^k + f_{k-1} {z_n}^{k-1} + \ldots + f_1 z_n + f_0,$$

where each f_i, $0 \leq i \leq m$ is a polynomial in the variables $z_1, \ldots, z_{n-1}$. Since we assumed that the power k of z_n is at least one, this implies that f_k must be non-zero.

[1] An outline for this proof in the case of the Lebesgue measure is given by Nathaniel Eldredge here: https://math.stackexchange.com/questions/1920302/the-lebesgue-measure-of-zero-set-of-a-polynomial-function-is-zero. We followed this outline, but provided more details and used different measures.

By our inductive hypothesis,

$$\mathbb{P}\left(f_k(X_1, \dots, X_{n-1}) = 0\right) = 0.$$

Letting M_k denote the zero set of f_k, this means that

$$\left(\mu_1 \times \mu_2 \times \cdots \times \mu_{n-1}\right)(M_k) = 0.$$

Let

$$M = \{(z_1, \dots, z_n) \in \mathbb{R}^n \mid f(z_1, \dots, z_n) = 0\}$$

By Tonelli's theorem,

$$\begin{aligned}
(\mu_1 \times \mu_2 \times \cdots \times \mu_n)(M) &= \int_{\mathbb{R}^n} \mathbb{1}_M(z_1, \dots, z_n)\, d\left(\mu_1 \times \mu_2 \times \cdots \times \mu_n\right)(z_1, z_2, \dots, z_n) \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} \mathbb{1}_M(z_1, \dots, z_{n-1}, z_n)\, d\left(\mu_1 \times \mu_2 \times \cdots \times \mu_{n-1}\right)(z_1, \dots, z_{n-1}) \right) d\mu_n(z_n) \\
&= \int_{\mathbb{R}} \left(\int_{M_k} \mathbb{1}_M(z_1, \dots, z_{n-1}, z_n)\, d\left(\mu_1 \times \mu_2 \times \cdots \times \mu_{n-1}\right)(z_1, \dots, z_{n-1}) \right) d\mu_n(z_n) \\
&+ \int_{\mathbb{R}^{n-1} \setminus M_k} \left(\int_{\mathbb{R}} \mathbb{1}_M(z_1, \dots, z_{n-1}, z_n)\, d\mu_n(z_n) \right) d\left(\mu_1 \times \mu_2 \times \cdots \times \mu_{n-1}\right)(z_1, \dots, z_{n-1}) \\
&= 0 + 0,
\end{aligned}$$

since the first integral in the summation is zero by the inductive hypothesis and the second integral in the summation is zero because a non-zero polynomial in n variables with $n-1$ variables fixed has only finitely many solutions, so this solution set has measure zero. Since $\mu_1, \dots, \mu_n$ are probability measures, they are each σ-finite. Additionally, $\mu_1 \times \cdots \times \mu_{n-1}$ is also σ-finite, as it is a product of σ-finite measures. This, along with the fact that the set M is measurable by the continuity of the polynomial, and that the indicator function is non-negative justifies the use of Tonelli's theorem. The statement holds by the principle of mathematical induction. □

Theorem 3.2.3 (Vieta's formulas, see Page 6 in [136])**.** *Suppose the polynomial*

$$f_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \quad a_n \neq 0$$

has coefficients in $\mathbb{R}$ *or* $\mathbb{C}$ *and roots given by* $\xi_1, \ldots, \xi_n$. *Then Vieta's formulas give the relationship between the roots and coefficients of* $f(z)$, *which is*

$$\sum_{1 \leq l_1 < l_2 < \ldots < l_k \leq n} \left(\prod_{j=1}^{k} \xi_{l_j} \right) = (-1)^k \frac{a_{n-k}}{a_n}$$

for $1 \leq k \leq n$.

We get the elementary symmetric polynomials from the left-hand side of Vieta's formulas.

Definition 3.2.4 (Elementary symmetric polynomials, see Page 6 in [136])**.** Let $n \in \mathbb{N}$. The n elementary symmetric polynomials in the variables $\xi_1, \ldots, \xi_n$ are defined to be

$$S_k(\xi_1, \ldots, \xi_n) = \sum_{1 \leq l_1 < l_2 < \ldots < l_k \leq n} \left(\prod_{j=1}^{k} \xi_{l_j} \right), \quad \text{for } 1 \leq k \leq n.$$

We also define

$$S_0(\xi_1, \ldots, \xi_n) = 1.$$

We call S_k the k^{th} elementary symmetric polynomial.

With these tools in hand, we may now prove Lemma 3.2.1.

Proof of Lemma 3.2.1. Let $\xi_1, \ldots, \xi_n$ denote the roots of $f_n(z)$. Since a_n is non-zero with probability one, we may assume that a_n is non-zero for the remainder of the proof. From Vieta's formulas in Theorem 3.2.3, the relationship between the roots and coefficients of $f_n(z)$ is given by

$$\sum_{1 \leq l_1 < l_2 < \ldots < l_k \leq n} \left(\prod_{j=1}^{k} \xi_{l_j} \right) = (-1)^k \frac{a_{n-k}}{a_n}. \tag{3.1}$$

Note that the left-hand sides of Vieta's equations are precisely the n elementary symmetric polynomials in the variables $\xi_1, \ldots, \xi_n$, while the right-hand sides contain only expressions in the coefficients of $f_n(z)$.

Consider the product

$$\prod_{i \neq j} (\xi_i - \xi_j),$$

which is also a symmetric polynomial and evaluates to zero if and only if $\xi_i = \xi_j$ for some $i \neq j$. The fundamental theorem of symmetric functions states that any symmetric polynomial q can be uniquely represented as a polynomial in the elementary symmetric polynomials in the variables of q; see Chapter 1, Section 2 in [113] for all of the details, with emphasis on (2.4) and the following remark. In our case, this means that $\prod_{i\neq j}(\xi_i - \xi_j)$ can be expressed as a polynomial in the elementary symmetric polynomials on $\xi_1, \ldots, \xi_n$, i.e.,

$$\prod_{i\neq j}(\xi_i - \xi_j) = Q(S_1(\xi_1, \ldots, \xi_n), \ldots, S_n(\xi_1, \ldots, \xi_n)),$$

where $S_i(\xi_1, \ldots, \xi_n)$ denotes the i-th elementary symmetric polynomial in the variables $\xi_1, \ldots, \xi_n$, and Q is a polynomial. Replacing the elementary symmetric polynomials in the polynomial Q by the right-hand sides in (3.1), we have that

$$\prod_{i\neq j}(\xi_i - \xi_j) = Q\left(-\frac{a_{n-1}}{a_n}, \ldots, (-1)^n \frac{a_0}{a_n}\right).$$

Since a_n is non-zero by assumption, we can multiply both sides of the above equation by the highest power of a_n occurring in the denominator of Q, to get that

$$a_n^{\deg(Q)} \prod_{i\neq j}(\xi_i - \xi_j) = \tilde{Q}\left(a_0, \ldots, a_{n-1}, a_n\right),$$

where $\tilde{Q}$ is the polynomial obtained after clearing denominators and making sign changes. By Lemma 3.2.2, the probability that $\tilde{Q}(a_0, \ldots, a_n) = 0$ is zero, so the probability that $\prod_{i\neq j}(\xi_i - \xi_j) = 0$ is also zero. Hence, the probability that $f_n(z)$ has a repeated root is zero. □

The following two results will be needed to bound the roots of $f_n(z)$.

Theorem 3.2.5 (Laguerre's inequality, see [86, 103])**.** *Given that the real polynomial $\sum_{k=0}^n a_k x^k$ with $a_n \neq 0$ has all real roots, the roots are located in the interval*

$$\left[-\frac{a_{n-1}}{na_n} - \frac{n-1}{na_n}\sqrt{a_{n-1}^2 - \frac{2n}{n-1}a_n a_{n-2}}, -\frac{a_{n-1}}{na_n} + \frac{n-1}{na_n}\sqrt{a_{n-1}^2 - \frac{2n}{n-1}a_n a_{n-2}}\right].$$

Theorem 3.2.6 (Landau's inequality, see [120])**.** *Let*

$$f_n(z) = a_n z^n + \cdots + a_0 = a_n \prod_{i=1}^n (z - \xi_i)$$

*with $a_n \neq 0$ be a polynomial with complex coefficients. Then its **Mahler measure** $M(f_n)$ is given by*

$$M(f_n) = |a_n| \prod_{i=1}^{n} \max\{1, |\xi_i|\}.$$

Landau's inequality is that

$$M(f_n) \leq \left(\sum_{i=0}^{n} |a_i|^2 \right)^{1/2}.$$

Using a polynomial's Mahler measure is a common method to bound the roots of a polynomial, see for instance [45].

3.3 Equivalence of generalized large box model and generalized bounded height model

In this section, we use Rouché's theorem to show that as $N \to \infty$, the probability that a monic degree n polynomial with coefficients uniformly distributed on $[\alpha_j N, \beta_j N] \cap \mathbb{Z}$ for $0 \leq j \leq n-1$ and jointly independent has all real roots converges to the probability that a non-monic polynomial of degree $n-1$ with independent coefficients each uniformly distributed on $[\alpha_j, \beta_j]$ for $0 \leq j \leq n-1$ has all real roots. This will allow us to work with the continuously distributed coefficients going forward, which simplifies our problem.

We will need to couple the coefficients of these polynomials to obtain the result.

Definition 3.3.1 (Definition 4.1.1 in [140])**.** Let μ and ν be probability measures on the same measurable space $(S, \mathcal{S})$. A **coupling** of μ and ν is a probability measure γ on the product space $(S \times S, \mathcal{S} \times \mathcal{S})$ such that the marginals of γ coincide with μ and ν, that is

$$\gamma(A \times S) = \mu(A) \quad \text{and} \quad \gamma(S \times A) = \nu(A), \quad \forall A \in S.$$

For two random variables X and Y taking values in $(S, \mathcal{S})$, a **coupling** of X and Y is a joint variable (X', Y') taking values in $(S \times S, \mathcal{S} \times S)$ whose law as a probability measure is a coupling of the laws of X and Y. Note that, under this definition, X and Y need not be defined on the same probability space (but X' and Y' do need to). We also say that (X', Y') is a **coupling** of μ and ν if the law of (X', Y') is a coupling of μ and ν.

Lemma 3.3.2. *Let $\alpha < \beta \in \mathbb{R}$. Let $N \in \mathbb{N}$ and let a be uniformly distributed on $[\alpha N, \beta N] \cap \mathbb{Z}$. Let b be uniformly distributed on $[\alpha, \beta]$. Then a coupling of a and b is given by the joint variable (a', b'), where $b' \sim b$, and for each $1 \leq i \leq \lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1$ and*

$$a' = \begin{cases} \lceil \alpha N \rceil + (i-1), & \text{if } b' \in \left[\alpha + \frac{(\beta-\alpha)(i-1)}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}, \alpha + \frac{(\beta-\alpha)i}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}\right) \\ \lfloor \beta N \rfloor + 1, & \text{if } b' = \beta \end{cases} \tag{3.2}$$

so that $a' \sim a$.

Proof. We begin by describing some details of the coupling. Notice first that the interval $[\alpha N, \beta N] \cap \mathbb{Z}$ contains $\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1$ integers. Dividing the interval $[\alpha, \beta)$ into $\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1$ pieces of length $\frac{\beta-\alpha}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}$, one obtains the partition

$$[\alpha, \beta) = \bigcup_{i=1}^{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1} \left[\alpha + \frac{(\beta-\alpha)(i-1)}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}, \alpha + \frac{(\beta-\alpha)i}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}\right). \tag{3.3}$$

Thus, (3.2) simply tells us that a' is the i-th integer in $[\alpha_j N, \beta_j N] \cap \mathbb{Z}$ whenever b' falls into the i-th interval of (3.3), and that $a' = \lfloor \beta N \rfloor + 1$ if $b' = \beta$. This means that for each $0 \leq j \leq n-1$, a' is simply defined in terms of a piecewise function of b', which we immediately see is measurable because each component is a constant. Therefore, a' and b' share a probability space.

It is given that $b' \sim b$. Let us now verify the marginal distribution of a'. Observe that for every integer k in $[\alpha N, \beta N] \cap \mathbb{Z}$, we have that $\mathbb{P}(a = k) = \frac{1}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}$ since the probability of b' falling into the i-th interval is $\frac{1}{\lfloor \beta N \rfloor - \lceil \alpha N \rceil + 1}$. In other words, a' is uniformly distributed on $[\alpha N, \beta N] \cap \mathbb{Z}$, as expected. Therefore, (a', b') is indeed a coupling of a and b. □

Remark 3.3.3. Suppose that $X_1, \ldots, X_n$ are jointly independent and $Y_1, \ldots, Y_n$ are jointly independent random variables. Suppose that (X_i', Y_i') is a coupling of X_i and Y_i for each $1 \leq i \leq n$, taking values in $(S_i \times S_i, \mathcal{S}_i \times \mathcal{S}_i)$. Then one can obtain a coupling $(X_1', Y_1', X_2', Y_2', \ldots, X_n', Y_n')$ of $(X_1, Y_1), \ldots, (X_n, Y_n)$ by embedding $(X_1', Y_1', X_2', Y_2', \ldots, X_n', Y_n')$ into the product space.

We can now prove Theorem 3.1.1.

Proof of Theorem 3.1.1. Since f and f/N have the same zeros, we will use Theorem 2.6.5 on f/N and g to compare the zeros of f and g. Suppose that the coefficients a_j and b_j, $0 \leq j \leq n-1$, of the random polynomials f and g have been coupled as in Lemma 3.3.2 and Remark 3.3.3. Refer to the coupled coefficients as a'_j and b'_j and the coupled polynomials as f' and g'. Then

$$|f'(z)/N - g'(z)| = |z^n/N + (a'_{n-1}/N - b'_{n-1})z^{n-1} + \cdots + (a'_0/N - b'_0)| \tag{3.4}$$

$$\leq |z^n/N| + \sum_{i=0}^{n-1} |a'_i/N - b'_i||z|^i. \tag{3.5}$$

We now obtain an upper bound for the term $|a'_j/N - b'_j|$, for each $0 \leq j \leq n-1$.

If $b'_j = \beta_j$, we have that

$$|a'_j/N - b'_j| = \left|\frac{\lfloor \beta_j N \rfloor + 1}{N} - \beta_j\right| \leq \left|\frac{\beta_j N + 1}{N} - \beta_j\right| \leq \frac{1}{N} = O(1/N).$$

If $b'_j \neq \beta_j$, we examine the distance of a'_j/N from both endpoints of the interval containing b'_j. A similar computation shows that for any $1 \leq l \leq \lfloor \beta_j N \rfloor - \lceil \alpha_j N \rceil + 1$,

$$\left|\frac{\lceil \alpha_j N \rceil + (l-1)}{N} - \left(\alpha_j + \frac{(\beta_j - \alpha_j)(l-1)}{\lfloor \beta_j N \rfloor - \lceil \alpha_j N \rceil + 1}\right)\right| = O(1/N)$$

and

$$\left|\frac{\lceil \alpha_j N \rceil + (l-1)}{N} - \left(\alpha_j + \frac{(\beta_j - \alpha_j)l}{\lfloor \beta_j N \rfloor - \lceil \alpha_j N \rceil + 1}\right)\right| = O(1/N).$$

Altogether, for each $0 \leq j \leq n-1$,

$$|a'_j/N - b'_j| \leq \frac{C_j}{N} \tag{3.6}$$

where the implicit constant $C_j > 0$ depends on α_j and β_j.

Combining (3.4) and (3.6), we have that

$$|f'(z)/N - g'(z)| \leq |z^n/N| + \left|\frac{1}{N}\right| \sum_{i=0}^{n-1} C_i |z|^i.$$

Next, fix the coefficients of the polynomial g' so that g' has a non-zero leading coefficient (which happens with probability one since $b_{n-1} \sim U(\alpha_j, \beta_j)$) and so that g' does not have any repeated roots (which happens with probability one by Lemma 3.2.1). By fixing the coefficients of the polynomial g', the coefficients of the polynomial f' are also fixed due to the assumed coupling.

Let $\xi_1', \ldots, \xi_{n-1}'$ denote the roots of g', ordered by increasing magnitude and then by increasing argument in $[0, 2\pi)$. Let $C_{\xi_j',\epsilon}$ denote the circle of radius ϵ centered at ξ_j', for $1 \leq j \leq n-1$. Since the fixed polynomial g' does not have any repeated roots, we can choose $\epsilon > 0$ to be sufficiently small such that the circles $C_{\xi_1',\epsilon}, \ldots, C_{\xi_{n-1}',\epsilon}$ are all disjoint.

We will show that for all N sufficiently large, $|f'(z)/N - g'(z)| < |g'(z)|$ for all $z \in C_{\xi_j',\epsilon}$, $1 \leq j \leq n-1$.

For every $1 \leq j \leq n-1$, $|g'(z)| = |b_{n-1}' z^{n-1} + \cdots + b_1' z + b_0'|$ is a continuous function on $C_{\xi_j',\epsilon}$. Since these circles were chosen to not contain any zeros of g', $|g'|$ obtains a minimum on each $C_{\xi_j',\epsilon}$ that is strictly positive.

For any $z \in C_{\xi_j',\epsilon}$, we may take N sufficiently large so that

$$\begin{aligned} |f'(z)/N - g'(z)| &\leq |z^n/N| + \left|\frac{1}{N}\right| \sum_{i=0}^{n-1} C_i |z|^i \\ &\leq \frac{\left(|\xi_j'| + \epsilon\right)^n}{N} + \frac{\sum_{i=0}^{n-1} C_i \left(|\xi_j'| + \epsilon\right)^i}{N} \\ &< \min_{z \in C_{\xi_j',\epsilon}} |g'(z)| \\ &\leq |g'(z)|, \end{aligned}$$

where the second line follows since $|z| \leq |\xi_j'| + \epsilon$ by Lemma 2.6.4, and the third line follows when N is sufficiently large since the minimum of $|g'(z)|$ for $z \in C_{\xi_j',\epsilon}$ is strictly positive because ϵ was chosen sufficiently small so that $C_{\xi_j',\epsilon}$ avoids the zeros of g'.

By Theorem 2.6.5, f'/N (and hence f') and g' have the same number of zeros in each $C_{\xi_j',\epsilon}$ for $1 \leq j \leq n-1$. This number must be one, because we assumed that the circles $C_{\xi_j',\epsilon}$ for $1 \leq j \leq n-1$ were disjoint.

Moreover, for N sufficiently large, g' has all real roots if and only if f' has all real roots. To see this, observe that if ξ_j' is a real root of g', then since ξ_j' is a single root by assumption, the circle $C_{\xi_j',\epsilon}$ contains exactly one root of f'. This root of f' must also be real, because otherwise $C_{\xi_j',\epsilon}$, whose center is on the real axis, would contain the root of f' and its complex conjugate. Since f' has one more root than g', it must also be real because complex roots occur in pairs.

On the other hand, suppose by contrapositive that not all roots of g' are real. If ξ'_j is a complex root of g' and $\bar{\xi}'_j$ is its conjugate, by assumption ϵ is sufficiently small so that $C_{\xi'_j,\epsilon} \cap C_{\bar{\xi}'_j,\epsilon} = \emptyset$. This means that the circles cannot intersect the real axis, so the corresponding roots of f' in each circle must also be complex.

Recall that the coefficients of g' were fixed so that g' has a non-zero leading coefficient and so that g' does not have any repeated roots. Since these events occur with probability one, the result holds for almost every realization of g'. Since $f \sim f'$ and since $g \sim g'$, this implies that as $N \to \infty$, the probability that f has all real roots converges to the probability that g has all real roots.

Next, we show the almost sure convergence of $n-1$ roots of f' to those of g'. Recall that we have shown that for every fixed realization of g' and coupled polynomial f', and for each $\epsilon > 0$ sufficiently small, the disjoint circles $C_{\xi'_j,\epsilon}$ contain exactly one root of g' and one root of f' when N is sufficiently large. For each $1 \le j \le n-1$, order the roots of f' by letting $\tilde{\xi}'^N_j$ be the root in the circle $C_{\xi'_j,\epsilon}$. Since f' has one more root than g', label the remaining root $\tilde{\xi}'^N_n$. This implies that for every fixed realization of g', coupled polynomial f', and any $\epsilon > 0$,

$$\left|\xi'_j - \tilde{\xi}'^N_j\right| < \epsilon$$

for each $1 \le j \le n-1$ and for all N sufficiently large. Now, stop fixing the polynomial g', but continue to suppose that g' does not have any repeated roots, and continue to couple the polynomial f' with g'. We wish to show that

$$\mathbb{P}\left(\bigcap_{1\le j\le n-1}\left\{\lim_{N\to\infty}\left|\xi'_j - \tilde{\xi}'^N_j\right| = 0\right\}\right) = 1.$$

Let $\epsilon = 1/k$ for $k \in \mathbb{N}$ and define the events

$$E_{j,k} = \left\{\limsup_{N\to\infty}\left|\xi'_j - \tilde{\xi}'^N_j\right| < 1/k.\right\}$$

From the discussion above, we know that $\mathbb{P}(E_{j,k}) = 1$ for each $1 \le j \le n-1$ and each $1 \le k < \infty$. Define

$$E_j = \bigcap_{k=1}^{\infty} E_{j,k}, \qquad \text{for } 1 \le j \le n.$$

Observe that $E_{j,1} \supseteq E_{j,2} \supseteq \ldots$ is a nested and decreasing sequence of events. Therefore, by continuity from above of the probability measure,

$$\begin{aligned}\mathbb{P}(E_j) &= \mathbb{P}\left(\bigcap_{k=1}^{\infty} E_{j,k}\right) \\ &= \lim_{k\to\infty} \mathbb{P}(E_{j,k}) \\ &= \inf_{k\geq 1} \mathbb{P}(E_{j,k}) \\ &= 1.\end{aligned}$$

Let

$$E = \bigcap_{1\leq j\leq n} E_j.$$

Then clearly

$$\mathbb{P}(E) = 1$$

as well. This shows that

$$\mathbb{P}\left(\bigcap_{1\leq j\leq n-1} \left\{\limsup_{N\to\infty} \left|\xi'_j - \tilde{\xi'}_j^N\right| = 0\right\}\right) = 1.$$

On the other hand, clearly

$$\mathbb{P}\left(\bigcap_{1\leq j\leq n-1} \left\{\liminf_{N\to\infty} \left|\xi'_j - \tilde{\xi'}_j^N\right| \geq 0\right\}\right) = 1.$$

Hence, the limit exists, and

$$\mathbb{P}\left(\bigcap_{1\leq j\leq n-1} \left\{\lim_{N\to\infty} \left|\xi'_j - \tilde{\xi'}_j^N\right| = 0\right\}\right) = 1,$$

so $\tilde{\xi'}_j^N \longrightarrow \xi'_j$ almost surely, for each $1 \leq j \leq n-1$.

Finally, we argue that $\tilde{\xi'}_n^N$ is the extremely dominant root of f' found in Section 2.6. In Section 2.6, we see that ξ'_n tends $\pm\infty$ as $N \to \infty$, depending on the sign of the coefficients of f'. By Theorem 3.2.6, we have that the product of all roots that have magnitude greater than 1 of g' must be bounded, since the coefficients of g' are all bounded. This means that the extra root of f', which we called $\tilde{\xi}'^N_n$, that is not contained in any of the circles we constructed around the roots of g' must be the extremely dominant root found in Section 2.6. $\square$

The following lemma proves that the root ξ_n is extremely dominant, as promised in Chapter 2.

Lemma 3.3.4. *For $0 \leq i \leq n-1$, let $\alpha_i < \beta_i \in \mathbb{R}$. Let $f_n(z) = z^n + a_{n-1}z + \cdots + a_1 z + a_0$, where for each $0 \leq i \leq n-1$, a_i is uniformly distributed on $[\alpha_i N, \beta_i N] \cap \mathbb{Z}$, with $a_0, \ldots, a_{n-1}$ jointly independent. Denoting the roots of $f_n(z)$ by ξ_n as defined in Theorem 2.1.5 and $\xi_1, \ldots, \xi_{n-1}$ for the remaining roots, we have that the modulus of ξ_n is not tight as $N \to \infty$, whereas the moduli of $\xi_1, \ldots, \xi_{n-1}$ are tight as $N \to \infty$.*

Proof. Let b_j be uniformly distributed on $[\alpha_j, \beta_j]$ for each $0 \leq j \leq n-1$ with $b_0, \ldots, b_{n-1}$ jointly independent. Define

$$g(z) = b_{n-1}z^{n-1} + \cdots + b_1 z + b_0.$$

Now couple the polynomials f and g as in Theorem 3.1.1, using the same notation for the coupled polynomials and roots therein. By Theorem 3.1.1, we know that $\tilde{\xi}'^{N}_j \longrightarrow \xi'_j$ almost surely, for each $1 \leq j \leq n-1$. By Theorem 3.2.6, we have that the product of all roots that have magnitude greater than 1 of g' must be bounded, since the coefficients of g' are all bounded. Hence, for each $1 \leq j \leq n-1$, and M sufficiently large,

$$\lim_{N \to \infty} \mathbb{P}\left(\left|\tilde{\xi}'^{N}_j\right| < M\right) = \mathbb{P}\left(\left|\xi'_j\right| < M\right) = 1,$$

implying the uniform tightness for $\left\{\left|\tilde{\xi}'^{N}_j\right| : N \in \mathbb{N}\right\}$ for each $1 \leq j \leq n-1$. Since $f \sim f'$, the roots $\xi_1, \ldots, \xi_{n-1}$ of f must be each be tight as well.

On the other hand, let ξ and ξ_n be defined as in the statement of Theorem 2.1.5. From Theorem 2.1.5, we know that for any fixed $\epsilon > 0$, $\xi_n \in B(\xi, \epsilon)$ with probability tending to 1 as $N \to \infty$. Moreover, for $a_{n-1} \neq 0$,

$$\begin{aligned}
|\xi| &= \left| \frac{-a_{n-1} - \operatorname{sgn}(a_{n-1})\sqrt{a_{n-1}^2 - 4a_{n-2}}}{2} \right| \\
&= |a_{n-1}| \left| \frac{1 + \sqrt{1 - \frac{4a_{n-2}}{a_{n-1}^2}}}{2} \right| \\
&\geq \frac{|a_{n-1}|}{2}.
\end{aligned}$$

If $0 \notin [\alpha_{n-1}N, \beta_{n-1}N]$, for every positive $M \in \mathbb{R}$ and for every $0 < \epsilon < 1$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(|\xi_n| > M\right) &\geq \lim_{N\to\infty} \mathbb{P}\left(|\xi| > M + \epsilon\right) \\ &\geq \lim_{N\to\infty} \mathbb{P}\left(\frac{|a_{n-1}|}{2} > M + \epsilon\right) \\ &= 1.\end{aligned}$$

If $0 \in [\alpha_{n-1}N, \beta_{n-1}N]$, for every positive $M \in \mathbb{R}$ and for every $0 < \epsilon < 1$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(|\xi_n| > M\right) &\geq \lim_{N\to\infty} \mathbb{P}\left(|\xi| > M + \epsilon\right) \\ &\geq \lim_{N\to\infty} \mathbb{P}\left(|\xi| > M + \epsilon \mid a_{n-1} \neq 0\right) \mathbb{P}\left(a_{n-1} \neq 0\right) \\ &\geq \lim_{N\to\infty} \mathbb{P}\left(\frac{|a_{n-1}|}{2} > M + \epsilon \mid a_{n-1} \neq 0\right)\left(\frac{\lfloor \beta_{n-2}N\rfloor - \lceil \alpha_{n-2}N\rceil}{\lfloor \beta_{n-2}N\rfloor - \lceil \alpha_{n-2}N\rceil + 1}\right) \\ &= 1.\end{aligned}$$

In either case, this shows that $\{|\xi_n| : N \in \mathbb{N}\}$ is not tight. □

3.4 Integral formulas for generalized bounded height model polynomials and the probability of all real roots

For $0 \leq j \leq n$, let $\alpha_j, \beta_j \in \mathbb{R}$, with $\alpha_j < \beta_j$. Consider the polynomial

$$f_n(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

with a_j uniformly distributed on $[\alpha_j, \beta_j]$ for $0 \leq j \leq n$ and $a_0, \ldots, a_n$ jointly independent. Denote the zeros of $f_n(z)$ by $\xi_1, \ldots, \xi_n$.

The main goal of this section is to find the probability that $f_n(z)$ has all real roots, which we will do by calculating the n-th dimensional Lebesgue measure of the zero set of the polynomial $f_n(z)$ and scaling by the appropriate constant. Results for finding the Lebesgue measures of the appropriate sets are given by Theorems 3.4.1 and 3.4.3. An integral formula for the probability that a generalized bounded height model polynomial has all real roots is given in Corollary 3.4.4. We also include a few hands on examples to show how to apply the theorems that are derived.

The notation and proof ideas come from [18], where for $N \geq 1$, the problem was tackled when $[\alpha_j, \beta_j] = [-N, N]$, $0 \leq j \leq n-1$. Some proof methods come from [2], which considered roots (rather than coefficients) in $[-N, N]$ instead.

We identify the vector $(a_0, \ldots, a_n) \in [\alpha_0, \beta_0] \times \cdots \times [\alpha_n, \beta_n] \subset \mathbb{R}^{n+1}$ with the polynomial $f_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$. Let $\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]$ denote the Cartesian product of the intervals for each coefficient, which is

$$\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right] := [\alpha_0, \beta_0] \times [\alpha_1, \beta_1] \times \cdots \times [\alpha_n, \beta_n].$$

When dividing all points in an interval $[\alpha, \beta]$ where $\alpha < \beta$ by $c < 0$, one obtains the interval $[\alpha/c, \beta/c]$. This resulting interval is now improperly ordered, since $\alpha/c > \beta/c$. Therefore, when $c < 0 \in \mathbb{R}$, we use the following convention.

$$\begin{aligned} \frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{c} &= \left[\frac{\beta_0}{c}, \frac{\alpha_0}{c}\right] \times \left[\frac{\beta_1}{c}, \frac{\alpha_1}{c}\right] \times \cdots \times \left[\frac{\beta_n}{c}, \frac{\alpha_n}{c}\right] \\ &= \frac{\left[\overrightarrow{\beta}, \overrightarrow{\alpha}\right]}{c}. \end{aligned}$$

When $c > 0$, an improper interval is not obtained, so we let

$$\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{c} = \left[\frac{\alpha_0}{c}, \frac{\beta_0}{c}\right] \times \left[\frac{\alpha_1}{c}, \frac{\beta_1}{c}\right] \times \cdots \times \left[\frac{\alpha_n}{c}, \frac{\beta_n}{c}\right].$$

Recalling that the j-th elementary symmetric polynomial in n variables is given by Definition 3.2.4, define the sets

$$\mathcal{H}_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \{(a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1} \mid a_j \in [\alpha_j, \beta_j], 0 \leq j \leq n, f_n(z) \text{ has all real roots}\},$$

$$\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \left\{(a_0, \ldots, a_{n-1}) \in \mathbb{R}^n \mid (a_0, \ldots, a_{n-1}, 1) \in \mathcal{H}_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)\right\}$$

and

$$H_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid \alpha_j \leq (-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n) \leq \beta_j,\ 0 \leq j \leq n-1\},$$

where we have tried to keep the notation consistent with [18].

The following theorem calculates the Lebesgue measure of $\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$. Using this result and varying the leading coefficient allows us to then calculate the Lebesgue measure of $\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$.

Theorem 3.4.1. *We have*

$$\lambda^n\left(\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)=\frac{1}{n!}\int_{H_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)}\prod_{1\le j<k\le n}|\xi_j-\xi_k|\;d\xi_1\dots d\xi_n.$$

Proof. The statement of this theorem is similar to Theorem 2.3 in [18], where for fixed $N\ge 1$, the authors state the result for $[\alpha_j,\beta_j]=[-N,N]$, $0\le j\le n-1$. Our case corresponds to $s=0$, where s is the number of complex pairs of roots of $f_n(z)$, but without placing any restrictions on α_j and β_j. Their proof relies on Theorem 2.1 in [2], so we proceed as in that proof. The main idea there is to use the change of variables between the coefficients and roots of the polynomial given by Vieta's formulas in Theorem 3.2.3. Explicitly, this change of variables is given by

$$a_j=(-1)^{n-j}S_{n-j}(\xi_1,\dots,\xi_n),\quad j=0,\dots,n-1, \tag{3.7}$$

where a_j is the j-th coefficient of the polynomial, $a_n=1$, and $\xi_1,\dots,\xi_n$ are the roots of the polynomial. The only difference now between our proof and the one given in [2] is our region of integration. Since $a_n=1$ and the remaining coefficients of our polynomials are in $[\alpha_j,\beta_j]$ for $0\le j\le n-1$, Vieta's formulas must satisfy

$$\alpha_j\le(-1)^{n-j}S_{n-j}(\xi_1,\dots,\xi_n)\le\beta_j,\quad 0\le j\le n-1.$$

The vectors $(\xi_1,\dots,\xi_n)$ that satisfy the above inequality become our region of integration, which we have called $H_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$. Then we have

$$\begin{aligned}\lambda_n\left(\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)&=\int_{\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)}1\;da_0\dots da_{n-1}\\&=\frac{1}{n!}\int_{H_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)}\prod_{1\le j<k\le n}|\xi_j-\xi_k|\;d\xi_1\dots d\xi_n\end{aligned}$$

where the determinant of the Jacobian of this transformation is

$$\prod_{1\le j<k\le n}(\xi_j-\xi_k),$$

which is computed in Remark 2.1 in [2] (using that $s = 0$), and the $\frac{1}{n!}$ is due to the n possible orderings of the rows of the Jacobian. □

The following is an example of a random monic polynomial that has zero probability of having all real roots.

Example 3.4.2. Consider $f_2(z) = z^2 + a_1 z + a_0$, with a_1 uniformly distributed on $[0, 1]$ and independent of a_0, which is uniformly distributed on $[1, 2]$. Then

$$\begin{aligned} H_2\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) &= H_2([0,1] \times [1,2] \times \{1\}) \\ &= \{(\xi_1, \xi_2) \in \mathbb{R}^2 \mid \alpha_j \leq (-1)^{2-j} S_{2-j}(\xi_1, \xi_2) \leq \beta_j,\ 0 \leq j \leq 1\}, \end{aligned}$$

which is the set satisfying the equations

$$0 \leq -\xi_1 - \xi_2 \leq 1$$

$$1 \leq \xi_1 \xi_2 \leq 2.$$

By inspection, we can see that no points (ξ_1, ξ_2) satisfy both of these equations, so

$$\lambda_2(\mathcal{H}_2^*([0,1] \times [1,2] \times \{1\})) = 0.$$

This makes sense because the discriminant $a_1^2 - 4a_0 \leq 1^2 - 4 \cdot 1 = -3$ of $f_2(z)$ is always negative, so this polynomial will never have all real roots.

Now define

$$\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \left\{(a_0, \ldots, a_n) \in \mathcal{H}_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) \mid a_n > 0\right\}$$

and

$$\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \left\{(a_0, \ldots, a_n) \in \mathcal{H}_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) \mid a_n < 0\right\}.$$

Theorem 3.4.3. *We have*

$$\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)\right) = \begin{cases} \beta_n^{n+1} \int_{\max\left\{0, \frac{\alpha_n}{\beta_n}\right\}}^{1} u^n \lambda^n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n u}\right)\right) du & \beta_n > 0 \\ 0 & \textit{otherwise} \end{cases}$$

and

$$\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)=\begin{cases}|\alpha_n|^{n+1}\int_{\max\left\{0,\frac{\beta_n}{\alpha_n}\right\}}^{1}u^n\lambda^n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\beta},\overrightarrow{\alpha}\right]}{\alpha_n u}\right)\right)du & \alpha_n<0\\ 0 & \textit{otherwise}\end{cases}.$$

In addition,

$$\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)=\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)+\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right). \tag{3.8}$$

Proof. First, notice that if $\beta_n \le 0$, then the interval $[\alpha_n, \beta_n]$ from $\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]$ is contained in $(-\infty, 0]$, so the leading coefficient of the polynomial $f_n(z)$ must be non-positive. In other words, $a_n < 0$, so $\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$ is the empty set, and has Lebesgue measure zero.

Similarly, if $\alpha_n \ge 0$, then $[\alpha_n, \beta_n] \subseteq [0, \infty)$, so $a_n > 0$. In this case, $\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$ is the empty set, and has Lebesgue measure zero.

So suppose $\beta_n > 0$ for the computation of $\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$ and that $\alpha_n < 0$ for the computation of $\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$.

Now we proceed similarly to the proof of Theorem 2.3 given in [18], where for $N \ge 1$, the result is stated for $[\alpha_j, \beta_j] = [-N, N]$, $0 \le j \le n-1$. There are a few main differences between our proof and the proof of Theorem 2.3. First, we must consider the possibility that our coefficients come from different intervals. We must also consider the cases for $\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$ and $\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$ separately since the intervals for our coefficients are not necessarily symmetric about the origin. We must modify the constants in the change of variables to α_n and β_n from the fixed height of the polynomials, and we must integrate over a slightly different region. The computation of the integral is then very similar, but we provide some details for completeness.

To compute

$$\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)=\int_{\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)}1\,da_0\dots da_n,$$

we make the following changes to the proof of Theorem 2.3 given in [18]:

- For the polynomial

$$f_n(z) = a_n z^n + \cdots + a_1 z + a_0,$$

use the change of variables

$$a_n = \beta_n q_n, \quad a_j = \beta_n q_n q_j, \quad 0 \le j \le n-1,$$

where β_n has replaced the fixed height of the polynomials from [18].

We can see that

$$\frac{\partial a_n}{\partial q_k} = \begin{cases} \beta_n & k = n \\ 0 & 0 \le k \le n-1 \end{cases},$$

and for $0 \le j \le n-1$,

$$\frac{\partial a_j}{\partial q_k} = \begin{cases} \beta_n q_n & k = j \\ \beta_n q_j & k = n \\ 0 & 0 \le k \ne j \le n-1 \end{cases}.$$

Therefore, the Jacobian of this change of variables is given by the $(n+1)\times(n+1)$ upper triangular matrix

$$\begin{pmatrix} \frac{\partial a_0}{\partial q_0} & \dots & \frac{\partial a_0}{\partial q_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_n}{\partial q_0} & \dots & \frac{\partial a_n}{\partial q_n} \end{pmatrix} = \begin{pmatrix} \beta_n q_n & 0 & \dots & \dots & 0 & \beta_n q_0 \\ 0 & \beta_n q_n & \ddots & & \vdots & \beta_n q_1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \beta_n q_n & 0 & \beta_n q_{n-2} \\ 0 & \dots & \dots & 0 & \beta_n q_n & \beta_n q_{n-1} \\ 0 & \dots & \dots & \dots & 0 & \beta_n \end{pmatrix}, \tag{3.9}$$

which has determinant $\beta_n^{n+1} q_n^n$. After applying this change of variables, the polynomial

$$f_n(z) = a_n z^n + \cdots + a_1 z + a_0$$

becomes

$$Q(z) = q_n \beta_n z^n + q_n q_{n-1} \beta_n z^{n-1} + \cdots + q_n \beta_n q_0.$$

- Since the leading coefficient a_n must be positive in $H_n^+\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)$, we have that

$$a_n \in \begin{cases} (0, \beta_n] & \text{if } \alpha_n \leq 0 \\ & \\ [\alpha_n, \beta_n] & \text{if } \alpha_n > 0 \end{cases}.$$

Therefore, when integrating over the region $\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)$, we must satisfy

$$\begin{cases} 0 < q_n = \dfrac{a_n}{\beta_n} \leq 1 & \text{if } \alpha_n \leq 0 \\ \dfrac{\alpha_n}{\beta_n} \leq q_n = \dfrac{a_n}{\beta_n} \leq 1 & \text{if } \alpha_n > 0. \end{cases}$$

However, since the boundary of the above set has Lebesgue measure zero, we may combine these cases into

$$\max\left\{0, \frac{\alpha_n}{\beta_n}\right\} < q_n \leq 1.$$

Furthermore,

$$\frac{\alpha_j}{\beta_n q_n} \leq q_j = \frac{a_j}{\beta_n q_n} \leq \frac{\beta_j}{\beta_n q_n}, \quad 0 \leq j \leq n-1$$

must also be satisfied when integrating over the region $\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)$.

- Note that

$$\begin{aligned} Q(z) &= q_n \beta_n z^n + q_n q_{n-1} \beta_n z^{n-1} + \cdots + q_n \beta_n q_0 \\ &= q_n \beta_n \left(z^n + q_{n-1} z^{n-1} + \cdots + q_1 z + q_0\right), \end{aligned}$$

so $Q(z)$ and the polynomial $z^n + q_{n-1}z^{n-1} + \cdots + q_1 z + q_0$ have the same roots. Therefore, we are interested in integrating over the region $(q_0, \ldots, q_{n-1}, q_n) \subseteq \mathbb{R}^{n+1}$ such that $q_n \in \left[\max\{0, \frac{\alpha_n}{\beta_n}\}, 1\right]$, $(q_0, \ldots, q_{n-1}) \in \left[\frac{\alpha_0}{\beta_n q_n}, \frac{\beta_0}{\beta_n q_n}\right] \times \cdots \times \left[\frac{\alpha_{n-1}}{\beta_n q_n}, \frac{\beta_{n-1}}{\beta_n q_n}\right]$, and such that the polynomial $z^n + q_{n-1}z^{n-1} + \cdots + q_1 z + q_0$ has all real roots. Recall that

$$\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n q_n}\right) = \left\{(q_0, \ldots, q_{n-1}) \in \mathbb{R}^n \mid (q_0, \ldots, q_{n-1}, 1) \in \mathcal{H}_n\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n q_n}\right)\right\}$$

and $(q_0, \ldots, q_{n-1}, 1) \in \mathcal{H}_n\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n q_n}\right)$ precisely when $z^n + q_{n-1}z^{n-1} + \cdots + q_1 z + q_0$ has all real roots and the coefficients $q_0, \ldots, q_n$ fall into the intervals described in the previous paragraph.

Therefore, we have

$$\begin{aligned}
\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right) &= \int_{\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)} 1 \; da_0 \dots da_n \\
&= \beta_n^{n+1}\int_{\max\left\{0,\frac{\alpha_n}{\beta_n}\right\}}^{1}\int_{\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]}{\beta_n q_n}\right)} q_n^n \; dq_0 \dots dq_{n-1} dq_n \\
&= \beta_n^{n+1}\int_{\max\left\{0,\frac{\alpha_n}{\beta_n}\right\}}^{1} q_n^n \lambda_n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]}{\beta_n q_n}\right)\right) dq_n.
\end{aligned}$$

To calculate $\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)$, we will use the change of variables

$$a_n = \alpha_n q_n, \quad a_j = \alpha_n q_n q_j, \quad 0 \le j \le n-1.$$

The Jacobian of this transformation looks similar to (3.9) but with α_n replacing all instances of β_n. Thus, the determinant of the Jacobian of this transformation is $\alpha_n^{n+1} q_n^n$.

Since the leading coefficient a_n must be negative in $H_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$, we have that

$$a_n \in \begin{cases} [\alpha_n, 0) & \text{if } \beta_n \ge 0 \\ [\alpha_n, \beta_n] & \text{if } \beta_n < 0 \end{cases}.$$

If $\beta_n \ge 0$, then $\frac{\beta_n}{\alpha_n} \le 0$ since $\alpha_n < 0$, and if $\beta_n < 0$, then $\frac{\beta_n}{\alpha_n} > 0$. Since the boundary has Lebesgue measure zero, the equations

$$\max\left\{0, \frac{\beta_n}{\alpha_n}\right\} < q_n = \frac{a_n}{\alpha_n} \le 1$$

and

$$\frac{\beta_j}{\alpha_n q_n} \le q_j \le \frac{\alpha_j}{\alpha_n q_n}, \quad 0 \le j \le n-1$$

must be satisfied when integrating over the region $\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$. Altogether, we are interested in integrating over the region $(q_0, \dots, q_{n-1}, q_n) \subseteq \mathbb{R}^{n+1}$ such that

$$q_n \in \left[\max\{0, \frac{\beta_n}{\alpha_n}\}, 1\right],$$

$$(q_0, \dots, q_{n-1}) \in \left[\frac{\beta_0}{\alpha_n q_n}, \frac{\alpha_0}{\alpha_n q_n}\right] \times \dots \times \left[\frac{\beta_{n-1}}{\alpha_n q_n}, \frac{\alpha_{n-1}}{\alpha_n q_n}\right],$$

and such that the polynomial $z^n + q_{n-1}z^{n-1} + \dots + q_1 z + q_0$ has all real roots.

Integrating like before, we get

$$\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)=|\alpha_n|^{n+1}\int_{\max\left\{0,\frac{\beta_n}{\alpha_n}\right\}}^1 q_n^n\lambda_n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\beta},\overrightarrow{\alpha}\right]}{\alpha_n q_n}\right)\right)dq_n,$$

where $|\alpha_n|$ comes from taking the absolute value of the determinant of the Jacobian. Now define the set

$$\mathcal{H}_n^0\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)=\left\{(a_0,\ldots,a_n)\in\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\ \mid a_n=0\right\}.$$

Then

$$0\le\lambda^{n+1}\left(\mathcal{H}_n^0\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)\le\lambda^{n+1}\left(\{[\alpha_0,\beta_0]\times\cdots\times[\alpha_{n-1},\beta_{n-1}]\times\{0\}\}\right)=0.$$

Finally, since

$$\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)=\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\cup\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)+\mathcal{H}_n^0\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$$

is a decomposition into disjoint sets,

$$\begin{aligned}\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)&=\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)+\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)+\lambda^{n+1}\left(\mathcal{H}_n^0\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)\\&=\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)+\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)+0.\end{aligned}$$

□

Note that the difficult part of the integration comes from finding the region $H_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)$.

An immediate corollary to Theorem 3.4.3 gives us the probability that a random polynomial whose coefficients obey the generalized bounded height model has all real roots.

Corollary 3.4.4. *For $0\le j\le n$, let $\alpha_j,\beta_j\in\mathbb{R}$, with $\alpha_j<\beta_j$. Then the probability that the polynomial*

$$f_n(z)=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0$$

with a_j uniformly distributed on $[\alpha_j,\beta_j]$ for $0\le j\le n$ and $a_0,\ldots,a_n$ jointly independent is given by

$$\frac{\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)}{\lambda^{n+1}\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)},$$

which may be evaluated using Theorem 3.4.3.

The following two examples show that there exist random non-monic polynomials with zero probability of having all real roots. This is in contrast to Theorem 3.5.18, which shows that if zero is in all of the intervals, we always have positive probability of having all real roots. Observe that the coefficients staying away from zero is crucial for these examples to work.

Example 3.4.5. Consider the polynomial $f_2(z) = a_2z^2 + a_1z + a_0$, where $a_0, a_1, a_2 \in [2,3]$ are independently and uniformly distributed. Combining Theorems 3.4.1 and 3.4.3, we see that since $\beta_2 = 3 > 0$,

$$\begin{aligned}\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right) &= \lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]\right)\right)\\ &= \beta_n^{n+1}\int_{\max\left\{0,\frac{\alpha_n}{\beta_n}\right\}}^1 u^n\lambda^n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]}{\beta_n u}\right)\right)du\\ &= \beta_n^{n+1}\int_{\max\left\{0,\frac{\alpha_n}{\beta_n}\right\}}^1 u^n\left(\frac{1}{n!}\int_{H_n\left(\frac{\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]}{\beta_n u}\right)}\prod_{1\le j<k\le n}|\xi_j-\xi_k|\,d\xi_1\ldots d\xi_n\right)du.\end{aligned}$$

Then

$$\begin{aligned}H_2\left(\frac{\left[\overrightarrow{\alpha},\overrightarrow{\beta}\right]}{\beta_n u}\right) &= H_2\left(\frac{[2,3]\times[2,3]\times[2,3]}{3u}\right)\\ &= \left\{(\xi_1,\xi_2)\in\mathbb{R}^2 \mid \frac{\alpha_j}{3u}\le(-1)^{2-j}S_{2-j}(\xi_1,\xi_2)\le\frac{\beta_j}{3u},\ 0\le j\le 1\right\},\end{aligned}$$

which is the set satisfying the equations

$$\frac{2}{3u}\le-\xi_1-\xi_2\le\frac{1}{u}$$
$$\frac{2}{3u}\le\xi_1\xi_2\le\frac{1}{u}.$$

Since $\max\{0,\alpha_2/\beta_2\} = 2/3$, we see that $2/3 \le u \le 1$. By inspection, we can see that no points (ξ_1,ξ_2) satisfy the above equations when $u \in [2/3,1]$. This implies that $\lambda_2(\mathcal{H}_2^*\left(\frac{[2,3]\times[2,3]\times[2,3]}{3u}\right) = 0$ for $u \in [2/3,1]$, so

$$\lambda^3\left(\mathcal{H}_n\left([2,3]\times[2,3]\times[2,3]\right)\right) = 0$$

as well. This makes sense because the discriminant of this polynomial is $a_1^2-4a_0a_2 \le 3^2-4\cdot2\cdot2 = -7$ is always negative, so the probability that $f_2(z)$ has all real roots is zero.

Example 3.4.6. Consider the polynomial $f_2(z) = a_2z^2+a_1z+a_0$, where $a_0 \in [1,2], a_1 \in [0,1], a_2 \in [1,2]$ are independently and uniformly distributed. Proceeding as in the previous example, we now consider

$$H_2\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n u}\right) = H_2\left(\frac{[1,2]\times[0,1]\times[1,2]}{2u}\right)$$
$$= \left\{(\xi_1,\xi_2) \in \mathbb{R}^2 \mid \frac{\alpha_j}{2u} \leq (-1)^{2-j} S_{2-j}(\xi_1,\xi_2) \leq \frac{\beta_j}{2u},\ 0 \leq j \leq 1\right\},$$

which is the set satisfying the equations

$$0 \leq -\xi_1 - \xi_2 \leq \frac{1}{2u}$$
$$\frac{1}{2u} \leq \xi_1\xi_2 \leq \frac{1}{u}.$$

Since $\max\{0, \alpha_2/\beta_2\} = 1/2$, we see that $1/2 \leq u \leq 1$. By inspection, we can see that no points (ξ_1,ξ_2) satisfy the above equations when $u \in [1/2,1]$. This implies that $\lambda_2(\mathcal{H}_2^*\left(\frac{[1,2]\times[0,1]\times[1,2]}{2u}\right) = 0$ for $u \in [1/2,1]$, so

$$\lambda^3\left(\mathcal{H}_n\left([1,2]\times[0,1]\times[1,2]\right)\right) = 0$$

as well. This makes sense because the discriminant of this polynomial is $a_1^2-4a_0a_2 \leq 1^2-4\cdot1\cdot1 = -3$ is always negative, so the probability that $f_2(z)$ has all real roots is zero.

3.5 Corollaries to Theorems 3.4.1 and 3.4.3 for specific choices of coefficients

We now focus on some results where the coefficients are chosen to satisfy certain properties. When all coefficients are chosen independently from the same interval $[\alpha,\beta]$, we can simply write α,β instead of $\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]$ for $[\alpha,\beta]^{n+1}$. Main results of this section include showing that $\lambda^n(\mathcal{H}_n^*(0,1))$ is positive for every $n \in \mathbb{N}$ and monotonically decreasing in n, obtaining lower bounds for $\lambda^n(\mathcal{H}_n^*(0,1))$, and finding a computationally simpler integral formula for $\lambda^n(\mathcal{H}_n^*(0,1)$. We investigate a conjectured relationship between $\lambda^n(\mathcal{H}_n^*(0,1))$ and the Selberg integral. Finally, we show that $\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right) > 0$ when zero is contained in $[\alpha_j,\beta_j]$ for each $0 \leq j \leq n$.

3.5.1 Coefficients restricted to the unit interval

Let $\alpha_0, \alpha_1, \ldots, \alpha_n = \alpha$ and $\beta_0, \beta_1, \ldots, \beta_n = \beta$ for the remainder of this section. Of particular interest is the case when $\alpha = 0$ and $\beta = 1$, which we will consider now. The main result of this section is the following.

Theorem 3.5.1. *The probability that the polynomial*

$$f_n(z) = a_n z^n + \ldots + a_1 z + a_0$$

with iid coefficients that are uniformly distributed on $[0,1]$ *has all real roots is positive.*

Additional results include corollaries establishing positive lower bounds for this probability. These lower bounds are given in terms of the Selberg integral, which we introduce now.

Definition 3.5.2. For $\mathrm{Re}(s) > 0$, define the **gamma function** by

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt.$$

The following theorem is used for obtaining a lower bound for $\lambda^{n+1}\left(\mathcal{H}_n(0,1)\right).$

Theorem 3.5.3 (Selberg's Integral Formula, (1.1) in [64])**.** *The Selberg integral formula shows that the n-dimensional integral*

$$\begin{aligned}\mathcal{S}_n(\alpha,\beta,\gamma) &:= \int_0^1 \cdots \int_0^1 \prod_{j=1}^n \xi_j^{\alpha-1}(1-\xi_j)^{\beta-1} \prod_{1\le j<k\le n} |\xi_j - \xi_k|^{2\gamma}\, d\xi_1 \ldots d\xi_n \\ &= \prod_{j=0}^{n-1} \frac{\Gamma(\alpha + j\gamma)\Gamma(\beta + j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(\alpha+\beta+(n+j-1)\gamma)\Gamma(1+\gamma)}\end{aligned}$$

as long as $\mathrm{Re}(\alpha) > 0, \mathrm{Re}(\beta) > 0, \mathrm{Re}(\gamma) > -\min\{1/n, \mathrm{Re}(\alpha)/(n-1), \mathrm{Re}(\beta)/(n-1)\}$.

Conjecture 3.5.4. *There exists an integer constant $C_n > 0$ depending on n such that*

$$\lambda^n(\mathcal{H}_n^*(0,1)) = \frac{\mathcal{S}_n(1,1,1/2)}{C_n},$$

where $\mathcal{S}_n$ denotes the n-dimensional Selberg integral and

$$\mathcal{S}_n(1,1,1/2) = \int_0^1 \cdots \int_0^1 \prod_{1\le i<j\le n} |\xi_i - \xi_j| \quad d\xi_1 \ldots d\xi_n.$$

When Akiyama and Pethő [2] considered bounded roots, rather than bounded coefficients, they were able to give exact answers to their questions in terms of Selberg integrals. Bertók, Hajdu, and Pethő [18] commented on this, stating that their own answers looked similar to the formulas obtained by the previous authors, but that they were unable to prove a relationship to the Selberg integral. They provided exact answers only for degrees 2 and 3, giving numerical approximations otherwise.

With Mathematica, we were able to verify the Conjecture 3.5.4 for the cases $n = 2$, $n = 3$, and $n = 4$ by integrating the function given in Theorem 3.4.1. In those cases, Mathematica computes that

$$\mathcal{S}_2(1, 1, 1/2) = 1/3, \quad \lambda^2(\mathcal{H}_2^*(0, 1)) = 1/12, \quad C_2 = 2^2$$

$$\mathcal{S}_3(1, 1, 1/2) = 1/30, \quad \lambda^3(\mathcal{H}_3^*(0, 1)) = 1/2880, \quad C_3 = 2^5 \cdot 3$$

$$\mathcal{S}_4(1, 1, 1/2) = 1/1050, \quad \lambda^4(\mathcal{H}_4^*(0, 1)) = 1/19353600, \quad C_4 = 2^{11} \cdot 3^2.$$

Mathematica was unable to evaluate $\lambda^5(\mathcal{H}_5^*(0, 1))$ from the formula given in Theorem 3.4.1. After finding a new way to express the integral, which is given later in Theorem 3.5.14, Mathematica computes that

$$\mathcal{S}_5(1, 1, 1/2) = 1/132300, \quad \lambda^5(\mathcal{H}_5^*(0, 1)) = 1/4649508864000, \quad C_5 = 2^{14} \cdot 3 \cdot 5 \cdot 11 \cdot 13.$$

For the case $n = 2$, we can prove the result geometrically, which we show in the following proposition.

Proposition 3.5.5. *We have that*

$$\lambda^2(\mathcal{H}_2^*(0, 1)) = \frac{\mathcal{S}_2(1, 1, 1/2)}{4},$$

where

$$\mathcal{S}_2(1, 1, 1/2) = \int_0^1 \int_0^1 |\xi_1 - \xi_2| \quad d\xi_1 d\xi_2.$$

Proof. When $n = 2$, we have that

$$\begin{aligned}
\mathcal{S}_2(1,1,1/2) &= \int_0^1 \int_0^1 |\xi_1 - \xi_2| \; d\xi_2 d\xi_1 \\
&= 2\int_0^1 \int_0^{-\xi_1+1} |\xi_1 - \xi_2| \; d\xi_2 d\xi_1 \\
&= 1/3
\end{aligned}$$

because the line $x_2 = -x_1 + 1$ divides $[0,1]^2$ into two right triangles, which are reflections of each other. By the symmetry of $|x_1 - x_2|$, the integrals over both of the triangles are equal.

On the other hand, applying Theorem 3.4.1,

$$\begin{aligned}
\lambda^2(\mathcal{H}_2^*(0,1)) &= \frac{1}{2}\int_{H_2(0,1)} |\xi_1 - \xi_2| \; d\xi_2 d\xi_1 \\
&= \frac{1}{2}\int_{-1}^0 \int_{-\xi_1-1}^0 |\xi_1 - \xi_2| \; d\xi_2 d\xi_1 \\
&= \frac{1}{2}\int_0^1 \int_0^{-\xi_1+1} |\xi_1 - \xi_2| \; d\xi_2 d\xi_1 \\
&= \frac{\mathcal{S}_2(1,1,1/2)}{2^2},
\end{aligned}$$

where the third equality holds from reflecting the region of integration of the second integral into the first quadrant. □

Lemma 3.5.6. *We have that*

$$\lambda^n(\mathcal{H}_n^*(0,1)) > 0$$

for all $n \in \mathbb{N}$.

Proof. From Theorem 3.4.1,

$$\lambda^n(\mathcal{H}_n^*(0,1)) = \frac{1}{n!}\int_{H_n(0,1)} \prod_{1\le j<k\le n} |\xi_j - \xi_k| \; d\xi_1 \dots d\xi_n$$

where

$$H_n(0,1) = \{(\xi_1,\dots,\xi_n) \in \mathbb{R}^n \mid 0 \le (-1)^{n-j} S_{n-j}(\xi_1,\dots,\xi_n) \le 1, 0 \le j \le n-1\}.$$

Since we are integrating a non-negative function, we will show that $\lambda^n(H_n(0,1)) > 0$ by showing that n-dimensional box $[-1/n, 0]^n \subseteq H_n(0,1)$. Suppose $(\xi_1, \ldots, \xi_n) \in [-1/n, 0]^n$. Then for $0 \leq j \leq n-1$,

$$S_{n-j}(\xi_1, \ldots, \xi_n) = \sum_{1 \leq l_1 < l_2 < \cdots < l_{n-j} \leq n} \xi_{l_1} \cdots \xi_{l_{n-j}}$$

is the sum of $\binom{n}{n-j}$ summands. For each summand, we have that

$$\begin{cases} 0 \leq \xi_{l_1} \cdots \xi_{l_{n-j}} \leq (-1)^{n-j} \left(\frac{1}{n}\right)^{n-j}, & \text{if } n-j \text{ even} \\ (-1)^{n-j} \left(\frac{1}{n}\right)^{n-j} \leq \xi_{l_1} \cdots \xi_{l_{n-j}} \leq 0, & \text{if } n-j \text{ odd} \end{cases}.$$

Hence,

$$\begin{aligned} 0 &\leq (-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n) \\ &\leq \binom{n}{n-j} \left(\frac{1}{n}\right)^{n-j} \\ &= \left(\frac{n!}{(n-j)!j!}\right) \left(\frac{1}{n}\right)^{n-j} \\ &= \frac{n(n-1)\cdots(n-j+1)}{n^{n-j} j!} \\ &\leq 1. \end{aligned}$$

This shows that $[-1/n, 0]^n \subseteq H_n(0,1)$, and this box clearly has positive Lebesgue measure. Since the integrand $\prod_{1 \leq j < k \leq n} |\xi_j - \xi_k|$ is positive unless $\xi_j = \xi_k, j \neq k$, which happens only on a set of measure zero, we see that $\lambda^n(\mathcal{H}_n^*(0,1))$ must be positive. □

Corollary 3.5.7. *If $\beta \geq 1$,*

$$\lambda^n(\mathcal{H}_n^*(0, \beta)) > 0.$$

Proof. If $\beta \geq 1$, then $\mathcal{H}_n^*(0,1) \subseteq \mathcal{H}_n^*(0,\beta)$ since any vector $(a_0, \ldots, a_{n-1}) \in [0,1]^n$ such that the polynomial $z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ has all real roots also satisfies $(a_0, \ldots, a_{n-1}) \in [0, \beta]^n$. Then the result follows immediately from Lemma 3.5.6. □

We are now able to prove the main result for this section.

Proof of Theorem 3.5.1. By Theorem 3.4.3,

$$\lambda^{n+1}(\mathcal{H}_n^+(0,1)) = \int_0^1 u^n \lambda^n \left(\mathcal{H}_n^*(0,1/u)\right)\; du.$$

Note that $1/u \in (1,\infty)$ as u ranges from 0 to 1. We have seen from Lemma 3.5.6 and Corollary 3.5.7 that $\lambda^n\left(\mathcal{H}_n^*(0,\beta)\right) > 0$ for $\beta \geq 1$, so $\lambda^{n+1}(\mathcal{H}_n^+(0,1) > 0$ as well. Since the volume of the $(n+1)$-dimensional box $[0,1]^{n+1}$ is 1, the value of $\lambda^{n+1}(\mathcal{H}_n^+(0,1)$ is precisely the probability that the polynomial with iid coefficients uniformly distributed on $[0,1]$ has all real roots. □

The following two results are both corollaries to Theorem 3.4.3.

Corollary 3.5.8. *We have that*

$$\lambda^{n+1}(\mathcal{H}_n^+(0,1)) \geq \int_0^1 u^n \lambda^n \left(\mathcal{H}_n^*(0,1)\right)\; du \geq \frac{\lambda^n(\mathcal{H}_n^*(0,1)}{n+1}.$$

Proof. By Theorem 3.4.3,

$$\lambda^{n+1}(\mathcal{H}_n^+(0,1)) = \int_0^1 u^n \lambda^n \left(\mathcal{H}_n^*\left(0,\frac{1}{u}\right)\right) du.$$

The first lower bound follows from observing that $\mathcal{H}_n^*(0,1) \subseteq \mathcal{H}_n^*(0,\beta)$ for $\beta \geq 1$, and clearly $1/u > 1$ for $u \in (0,1]$. The second lower bound follows since $\int_0^1 u^n\; du = 1/(n+1)$. □

Using the values $\lambda^2(\mathcal{H}_2^*(0,1)) = 1/6, \lambda^3(\mathcal{H}_3^*(0,1)) = 1/2880$, $\lambda^4(\mathcal{H}_4^*(0,1)) = 1/19353600$ and $\lambda^5(\mathcal{H}_5^*(0,1)) = 1/4649508864000$ computed by Mathematica, Corollary 3.5.8 gives us the following lower bounds:

$$\lambda^3(\mathcal{H}_2^+(0,1) \geq 1/18 \approx 0.0555556$$

$$\lambda^4(\mathcal{H}_3^+(0,1) \geq 1/11520 \approx 0.0000868056$$

$$\lambda^5(\mathcal{H}_4^+(0,1) \geq 1/77414400 \approx 0.0000000129175$$

$$\lambda^6(\mathcal{H}_5^+(0,1) \geq 1/464950886400 \approx 0.000000000000358461.$$

Corollary 3.5.9. *We have that*

$$\lambda^n\left(\mathcal{H}_n^*(0,1)\right) \geq \int_0^{1/n} \cdots \int_0^{1/n} \prod_{1\leq j<k\leq n} |\xi_j - \xi_k|\; d\xi_1 \cdots d\xi_n.$$

$$= \frac{\mathcal{S}_n(1,1,1/2)}{n^{(n^2+n)/2}}.$$

Proof. The first line follows from the fact that $[-1/n, 0]^n \subseteq H_n(0,1)$, which was seen in the proof of Lemma 3.5.6, and symmetry. For the second line, consider the change of variables

$$\xi_1 = \frac{u_1}{n}, \ldots, \xi_n = \frac{u_1}{n}.$$

The determinant of the Jacobian of this transformation is $\frac{1}{n^n}$. Then we have

$$\begin{aligned}
\prod_{1\le j<k\le n} |\xi_j - \xi_k| &= \prod_{1\le j<k\le n} \frac{1}{n} |u_j - u_k| \\
&= \frac{1}{n^{\binom{n}{2}}} \prod_{1\le j<k\le n} |u_j - u_k| \\
&= \frac{1}{n^{\left(\frac{n^2-n}{2}\right)}} \prod_{1\le j<k\le n} |u_j - u_k|.
\end{aligned}$$

Altogether, we have

$$\begin{aligned}
&\int_0^{1/n} \cdots \int_0^{1/n} \prod_{1\le j<k\le n} |\xi_j - \xi_k| \; d\xi_1 \cdots d\xi_n \\
=&\left(\frac{1}{n^n}\right)\left(\frac{1}{n^{\left(\frac{n^2-n}{2}\right)}}\right) \int_0^1 \cdots \int_0^1 \prod_{1\le j<k\le n} |u_j - u_k| \; du_1 \cdots du_n \\
=&\frac{\mathcal{S}_n(1,1,1/2)}{n^{(n^2+n)/2}}
\end{aligned}$$

□

Corollary 3.5.10. *We have that*

$$\lambda^{n+1}(\mathcal{H}_n^+(0,1)) \ge \frac{\mathcal{S}_n(1,1,1/2)}{n^{(n^2+2)/2}(n+1)}.$$

Proof. This is simply combining the lower bounds obtained in Corollaries 3.5.8 and 3.5.9. □

3.5.2 Monotonicity of $\lambda^n(\mathcal{H}_n^*(0,1))$

We say that a sequence of real numbers $\{a_n\}$ is **monotonically decreasing** if $a_i \ge a_j$ whenever $i < j$. The goal of this subsection is to prove the following.

Proposition 3.5.11. *The sequence $\lambda^n(\mathcal{H}_n^*(0,1))$ is monotonically decreasing in n.*

We will need the following two lemmas that let us see the region $H_n(0,1)$ in a new way to accomplish this. Lemma 3.5.12 is not a new result, but we were unable to locate a published reference.[2]

Lemma 3.5.12. *We may write the elementary symmetric polynomials in the following recursive manner. Let $1 \leq k \leq n$. Then*

$$S_k(\xi_1, \ldots, \xi_n) = \xi_n S_{k-1}(\xi_1, \ldots, \xi_{n-1}) + \xi_{n-1} S_{k-1}(\xi_1, \ldots, \xi_{n-2}) + \cdots + \xi_k S_{k-1}(\xi_1, \ldots, \xi_{k-1}),$$

where for the case $k = 1$, we just let $S_{k-1}(\xi_1, \ldots, \xi_{k-1}) = S_0(\xi_1) = 1$ by convention.

Proof. Recall that $S_0(\xi_m, \ldots, \xi_l) = 1$ for any counting numbers $0 \leq m \leq l$, including the case $S_0(\xi_1) = 1$. Now let $k = 1$. Then we have that

$$\begin{aligned} S_1(\xi_1, \ldots, \xi_n) &= \xi_1 + \xi_2 + \ldots + \xi_n \\ &= \xi_n S_0(\xi_1, \ldots, \xi_{n-1}) + \xi_{n-1} S_0(\xi_1, \ldots, \xi_{n-2}) \\ &\quad + \cdots + \xi_2 S_0(\xi_1) + \xi_1 S_0(\xi_1), \end{aligned}$$

so the result holds.

In general, recall that for $1 \leq k \leq l$,

$$S_k(\xi_1, \ldots, \xi_l) = \sum_{1 \leq l_1 < l_2 < \cdots < l_k \leq l} \xi_{l_1} \xi_{l_2} \cdots \xi_{l_k}.$$

The idea behind the proof is to start with the expression for $S_k(\xi_1, \ldots, \xi_l)$, and then to factor out ξ_n from all the terms that have an ξ_n. Then we write the terms that were multiplied by ξ_n as elementary symmetric polynomials of degree $k - 1$. Then, considering only the terms that did not contain an ξ_n, repeat the process for ξ_{n-1}, and so forth, until ξ_k has been factored out. The last

[2] The result is documented on the ProofWiki: https://proofwiki.org/wiki/Recursion_Property_of_Elementary_Symmetric_Function

term will always be $\xi_k\xi_{k-1}\cdots\xi_1 = \xi_k S_{k-1}(\xi_1,\ldots,\xi_{k-1})$ because this is the last possible combination of k variables for $S_k(\xi_1,\ldots,\xi_n)$.

Now let $2 \le k \le n$. Performing the process above, we see that

$$\begin{aligned}
S_k(\xi_1,\ldots,\xi_n) &= \sum_{1\le l_1<l_2<\cdots<l_k\le n} \xi_{l_1}\xi_{l_2}\cdots\xi_{l_k} \\
&= \xi_n\left(\sum_{1\le l_1<l_2<\cdots<l_{k-1}\le n-1} \xi_{l_1}\xi_{l_2}\cdots\xi_{l_{k-1}}\right) \\
&\quad + \xi_{n-1}\left(\sum_{1\le l_1<l_2<\cdots<l_{k-1}\le n-2} \xi_{l_1}\xi_{l_2}\cdots\xi_{l_{k-1}}\right) \\
&\quad + \cdots + \xi_k\left(\sum_{1\le l_1<l_2<\cdots<l_{k-1}\le k-1} \xi_{l_1}x_{l_2}\cdots\xi_{l_{k-1}}\right) \\
&= \xi_n S_k(\xi_1,\ldots,\xi_{n-1}) + \xi_{n-1}S_k(\xi_1,\ldots,\xi_{n-2}) + \cdots + \xi_k S_k(\xi_1,\ldots\xi_{k-1}).
\end{aligned}$$

□

Lemma 3.5.13. *For $n \in \mathbb{N}$ and $\xi_1,\ldots,\xi_n \in \mathbb{R}$, consider the set of equations*

$$\begin{cases}
-1-\xi_n-\cdots-\xi_2 \le \xi_1 \le 0 \\
-1-\xi_n-\cdots-\xi_3 \le \xi_2 \le 0 \\
\qquad\vdots \\
-1-\xi_n \le \xi_{n-1} \le 0 \\
-1 \le \xi_n \le 0.
\end{cases} \tag{3.10}$$

Then the region

$$H_n(0,1) = \{(\xi_1,\ldots,\xi_n) \in \mathbb{R}^n \mid 0 \le (-1)^{n-j}S_{n-j}(\xi_1,\ldots,\xi_n) \le 1,\ 0 \le j \le n-1\}$$

and the region defined by

$$\tilde{H}_n(0,1) = \{(\xi_1,\ldots,\xi_n) \in \mathbb{R}^n \mid \textit{equations (3.10) are satisfied}\}$$

are the same.

Proof. Writing each $S_{n-j}(\xi_1, \ldots, \xi_n)$ from $H_n(0,1)$ in summation notation for $0 \le j \le n-1$ and multiplying by -1, we want to show that the set of equations

$$\begin{cases} -1 \le \sum_{1 \le l_1 \le n} \xi_{l_1} \le 0 \\ -1 \le -\sum_{1 \le l_1 < l_2 \le n} \xi_{l_1}\xi_{l_2} \le 0 \\ \qquad \vdots \\ -1 \le (-1)^n \sum_{1 \le l_1 < l_2 < \cdots < l_{n-1} \le n} \xi_{l_1} \cdots \xi_{l_{n-1}} \le 0 \\ -1 \le (-1)^{n+1} \xi_1 \xi_2 \cdots \xi_n \le 0 \end{cases} \tag{3.11}$$

and the set of equations (3.10) are equivalent.

First, we will show that if a point $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ satisfies either set of equations, then we must also have that $(\xi_1, \ldots, \xi_n) \in [-1, 0]^n$.

Considering equations (3.11), define

$$a_{n-k} = (-1)^k \sum_{1 \le l_1 < l_2 < \ldots < l_k \le n} \left(\prod_{j=1}^{k} \xi_{l_j} \right)$$

for $1 \le k \le n$. Then equations (3.11) imply that each a_j is contained in $[0,1]$ for $0 \le j \le n-1$, with some restrictions to ensure that $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$. Examining Vieta's formulas given in Theorem 3.2.3, we see that the $\xi_1, \ldots, \xi_n$ may be interpreted as being the roots of the polynomial $f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$. For any positive $x \in \mathbb{R}$, $f_n(x) \ge x^n > 0$, so x cannot be a root of $f_n(z)$. Therefore, all roots of $f_n(z)$ must be in $(-\infty, 0]$. Now, consider the left endpoint of the interval from Theorem 3.2.5, given by

$$-\frac{a_{n-1}}{na_n} - \frac{n-1}{na_n} \sqrt{a_{n-1}^2 - \frac{2n}{n-1} a_n a_{n-2}},$$

where $a_n = 1$ since $f_n(z)$ is monic. Since we have assumed that $f_n(z)$ has all real roots, the discriminant $a_{n-1}^2 - \frac{2n}{n-1} a_n a_{n-2}$ is non-negative. Recalling that $a_{n-1}, a_{n-2} \in [0,1]$, we have that

$$0 \le a_{n-1}^2 - \frac{2n}{n-1} a_n a_{n-2} \le 1^2 - 0 = 1.$$

Therefore,

$$-\frac{a_{n-1}}{na_n} - \frac{n-1}{na_n}\sqrt{a_{n-1}^2 - \frac{2n}{n-1}a_n a_{n-2}} \geq -\frac{a_{n-1}}{n} - \frac{n-1}{n}$$
$$\geq -\frac{1}{n} - \frac{n-1}{n}$$
$$= -1.$$

Therefore, the roots of $f_n(z)$ are in $[-1, 0]$.

For equations (3.10), we can see that $\xi_1, \ldots, \xi_n \in [-1, 0]$ by working backward. In the last equation from the set (3.10), it is given that $-1 \leq \xi_n \leq 0$. Examining the previous equation in (3.10), we have that

$$-1 - \xi_n \leq \xi_{n-1} \leq 0.$$

Since $\xi_n \in [-1, 0]$, this implies that

$$-1 \leq \xi_{n-1} \leq 0.$$

Notice that for $1 \leq k \leq n$,

$$-1 - \xi_n - \cdots - \xi_{k+1} \leq \xi_k \leq 0$$

is given, so if it has been shown that $\xi_n, \xi_{n-1}, \ldots, \xi_{k+1} \in [-1, 0]$, then $\xi_k \in [-1, 0]$ also. Therefore, by induction, equations (3.10) imply that $\xi_1, \ldots, \xi_n \in [-1, 0]$

Next, we will show that equations (3.11) imply equations (3.10). Suppose that the point $(\xi_1, \xi_2, \ldots, \xi_n)$ satisfies equations (3.11), which implies that each $\xi_1, \ldots, \xi_n \in [-1, 0]$. Then, combining that $\xi_1 \in [-1, 0]$ with the rearranged the first inequality in (3.11), we have that

$$-1 - \xi_2 - \cdots - \xi_n \leq \xi_1 \leq 0,$$

so the first inequality in (3.10) holds. Suppose now as our inductive hypothesis that

$$-1 - \xi_n - \cdots - \xi_{k+1} \leq \xi_k \leq 0$$

holds for some $k \in \mathbb{N}$. Adding ξ_{k+1}, we have that

$$-1 - \xi_n - \cdots - \xi_{k+2} \leq \xi_k + \xi_{k+1} \leq \xi_{k+1} \leq 0$$

since $\xi_k, \xi_{k+1} \in [-1, 0]$. Hence, equations (3.10) hold by the principle of mathematical induction.

Now, we will show that equations (3.10) imply equations (3.11). Suppose that the point $(\xi_1, \xi_2, \ldots, \xi_n)$ satisfies equations (3.10), which implies that $\xi_1, \ldots, \xi_n \in [-1, 0]$. From rearranging the first inequality in (3.10) and bounding above, we see that

$$-1 \leq \xi_1 + \cdots + \xi_n \leq \xi_2 + \cdots + \xi_n \leq 0$$

since $\xi_2, \ldots, \xi_n \in [-1, 0]$, so the first equation in (3.11) holds.

Lemma 3.5.12 tells us that for $1 \leq k \leq n$,

$$\begin{aligned} S_k(\xi_1, \ldots, \xi_n) = \xi_n S_{k-1}(\xi_1, \ldots, \xi_{n-1}) + \xi_{n-1} S_{k-1}(\xi_1, \ldots, \xi_{n-2}) \\ + \cdots + \xi_k S_{k-1}(\xi_1, \ldots, \xi_{k-1}). \end{aligned}$$

As our inductive hypothesis, suppose that when equations (3.10) hold, it has been shown that

$$0 \leq -S_1(\xi_1, \ldots, \xi_n) \leq 1, 0 \leq S_2(\xi_1, \ldots, \xi_n) \leq 1, \ldots, 0 \leq (-1)^{k-1} S_{k-1}(\xi_1, \ldots, \xi_n) \leq 1,$$

i.e., that the first $k-1$ equations of equation (3.11) hold, where $k \geq 2$.

If k is odd, observe that $S_k(\xi_1, \ldots, \xi_n) \leq 0$ since $\xi_1, \ldots, \xi_n \in [-1, 0]$. Moreover, for any $1 \leq l \leq n$, we have that

$$0 \leq S_{k-1}(\xi_1, \ldots, \xi_l) \leq S_{k-1}(\xi_1, \ldots, \xi_n) \leq 1,$$

since $\xi_1, \ldots, \xi_n \in [-1, 0]$. From this, we see that

$$\begin{aligned} S_k(\xi_1, \ldots, \xi_n) &= \xi_n S_{k-1}(\xi_1, \ldots, \xi_{n-1}) + \xi_{n-1} S_{k-1}(\xi_1, \ldots, \xi_{n-2}) + \cdots + \xi_k S_{k-1}(\xi_1, \ldots, \xi_{k-1}) \\ &\geq \xi_n + \cdots + \xi_k \\ &\geq \xi_n + \cdots + \xi_1 \\ &\geq -1. \end{aligned}$$

Combining, we have shown that $0 \leq (-1)^k S_k(\xi_1, \ldots, \xi_n) \leq 1$.

Similarly, if k is even, observe that $0 \le S_k(\xi_1, \ldots, \xi_n)$ since $\xi_1, \ldots, \xi_n \in [-1, 0]$. Moreover, for any $1 \le l \le n$, we have that

$$0 \ge S_{k-1}(\xi_1, \ldots, \xi_l) \ge S_{k-1}(\xi_1, \ldots, \xi_n) \ge -1$$

and so

$$\begin{aligned} S_k(\xi_1, \ldots, \xi_n) &= \xi_n S_{k-1}(\xi_1, \ldots, \xi_{n-1}) + \xi_{n-1} S_{k-1}(\xi_1, \ldots, \xi_{n-2}) + \cdots + \xi_k S_{k-1}(\xi_1, \ldots, \xi_{k-1}) \\ &\le -\xi_n - \cdots - \xi_k \\ &\le -\xi_n - \cdots - \xi_1 \\ &\le 1. \end{aligned}$$

Combining, we have shown that $0 \le (-1)^k S_k(\xi_1, \ldots, \xi_n) \le 1$. In either case, we can see that equations (3.11) hold. □

We are now able to prove that the sequence $\lambda^n(\mathcal{H}_n^*(0,1))$ is monotonically decreasing in n.

Proof of Proposition 3.5.11. We want to show that

$$\lambda^{n+1}\left(\mathcal{H}_{n+1}^*(0,1)\right) \le \lambda^n\left(\mathcal{H}_n^*(0,1)\right)$$

for all $n \in \mathbb{N}$.

From Theorem 3.4.1, we have that

$$\lambda^n(\mathcal{H}_n^*(0,1)) = \frac{1}{n!} \int_{H_n(0,1)} \prod_{1 \le j < k \le n} |\xi_j - \xi_k| \, d\xi_1 \ldots d\xi_n,$$

so simplifying the factorials, we wish to show that

$$\int_{H_{n+1}(0,1)} \prod_{1 \le j < k \le n+1} |\xi_j - \xi_k| \, d\xi_1 \ldots d\xi_{n+1} \le (n+1) \int_{H_n(0,1)} \prod_{1 \le j < k \le n} |\xi_j - \xi_k| \, d\xi_1 \ldots d\xi_n.$$

By Lemma 3.5.13, for every $n \in \mathbb{N}$, the region $H_{n+1}(0,1)$ is given by the equations

$$-1 \le \xi_{n+1} \le 0$$

$$-1 - \xi_{n+1} \le \xi_n \le 0$$

$$-1-\xi_{n+1}-\xi_n \leq \xi_{n-1} \leq 0$$

$$\vdots$$

$$-1-\xi_{n+1}-\cdots-\xi_2 \leq \xi_1 \leq 0.$$

From the proof of Lemma 3.5.13, we see that $-1 \leq \xi_j \leq 0$ for each $1 \leq j \leq n+1$. Therefore, $|\xi_1 - \xi_j| \leq 1$ for each $2 \leq j \leq n+1$. Hence, we may bound

$$\prod_{1\leq j<k\leq n+1} |\xi_j - \xi_k| = \prod_{2\leq j\leq n+1} |\xi_1 - \xi_j| \prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k| \leq \prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k|$$

over the region $H_{n+1}(0,1)$. Since $\prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k|$ does not include ξ_1, it is now straightforward to integrate with respect to ξ_1. This gives

$$\begin{aligned}
&\int_{H_{n+1}(0,1)} \prod_{1\leq j<k\leq n+1} |\xi_j - \xi_k| \; d\xi_1 \dots d\xi_{n+1} \\
&= \int_{-1}^{0} \cdots \int_{-1-\xi_{n+1}-\cdots-\xi_2}^{0} \prod_{1\leq j<k\leq n+1} |\xi_j - \xi_k| \; d\xi_1 \dots d\xi_{n+1} \\
&\leq \int_{-1}^{0} \cdots \int_{-1-\xi_{n+1}-\cdots-\xi_2}^{0} \prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k| \; d\xi_1 \dots d\xi_{n+1} \\
&\leq \int_{-1}^{0} \cdots \int_{-1-\xi_{n+1}-\cdots-\xi_3}^{0} (n+1) \prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k| \; d\xi_2 \dots d\xi_{n+1},
\end{aligned}$$

where in the last line we used the fact that

$$\int_{-1-\xi_{n+1}-\cdots-\xi_2}^{0} \prod_{2\leq j<k\leq n} |\xi_j - \xi_k| \; dx_1 \leq (n+1) \prod_{2\leq j<k\leq n} |\xi_j - \xi_k|$$

since

$$-1-\xi_{n+1}-\cdots-\xi_2 \geq -(n+1).$$

Then

$$\begin{aligned}
&\int_{-1}^{0} \cdots \int_{-1-\xi_{n+1}-\cdots-\xi_3}^{0} (n+1) \prod_{2\leq j<k\leq n+1} |\xi_j - \xi_k| \; d\xi_2 \dots d\xi_{n+1} \\
&= \int_{-1}^{0} \cdots \int_{-1-\xi_n-\cdots-\xi_2}^{0} (n+1) \prod_{1\leq j<k\leq n} |\xi_j - \xi_k| \; d\xi_1 \dots d\xi_n
\end{aligned}$$

by relabeling the variables. $\square$

3.5.3 An $(n-1)$-dimensional integral formula for $\lambda^n(\mathcal{H}_n^*(0,1))$

The following theorem allows us to rewrite the integral formula for $\lambda^n(\mathcal{H}_n^*(0,1))$ given by Theorem 3.4.1. The resulting integral is symmetric and only of dimension $n-1$. It was using this result that Mathematica was able to obtain the precise value for $\lambda^5(\mathcal{H}_5^*(0,1))$, which was presented following Conjecture 3.5.4.

Theorem 3.5.14. *For $n \geq 3$, we have*

$$\lambda^n(\mathcal{H}_n^*(0,1)) = \frac{2}{(n+1)!n} \int_0^\infty \cdots \int_0^\infty \frac{\prod_{1\leq j<k\leq n-1} |\xi_j - \xi_k| \prod_{j=1}^{n-1} |1-\xi_j|}{(\xi_1 + \cdots + \xi_{n-1} + 1)^{\frac{(n+1)n}{2}}} \, d\xi_1 \cdots d\xi_{n-1}.$$

Proof. From Theorem 3.4.1 and Lemma 3.5.13, we know that

$$\begin{aligned}\lambda^n\left(\mathcal{H}_n^*(0,1)\right) &= \frac{1}{n!} \int_{H_n([0,1])} \prod_{1\leq j<k\leq n} |\xi_j - \xi_k| \, d\xi_1 \ldots d\xi_n \\ &= \frac{1}{n!} \int_{-1}^0 \int_{-1-\xi_n}^0 \cdots \int_{-1-\xi_n-\cdots-\xi_2}^0 \prod_{1\leq i<j\leq n} |\xi_i - \xi_j| \, d\xi_1 \cdots d_{\xi_n}.\end{aligned}$$

First, perform the change of variables $w_1 = -\xi_1, \ldots, w_n = -\xi_n$ to transform the region of integration into the first orthant. This gives

$$\lambda^n\left(\mathcal{H}_n^*(0,1)\right) = \frac{1}{n!} \int_0^1 \int_0^{-w_n+1} \cdots \int_0^{-w_n-\cdots-w_2+1} \prod_{1\leq i<j\leq n} |w_i - w_j| \, dw_1 \cdots dw_n.$$

Next, perform the change of variables $z = w_1, zt_2 = w_2, \ldots, zt_n = w_n$. The determinant of the Jacobian of this transformation is z^{n-1}. This trick, similar to the one used in Selberg's original proof of the Selberg integral given in Mehta's book [119] gives

$$\begin{aligned}\prod_{1\leq i<j\leq n} |w_i - w_j| \, dw_1 \ldots dw_n &= |z|^{n-1} \prod_{2\leq j\leq n} |z - zt_i| \prod_{2\leq i<j\leq n} |zt_i - zt_j| \, dzdt_2 \ldots dt_n \\ &= |z|^{n(n-1)/2} \cdot |z|^{n-1} \prod_{2\leq j\leq n} |1-t_j| \prod_{2\leq i<j\leq n} |t_i - t_j| \, dzdt_2 \ldots dt_n \\ &= |z|^{(n^2+n-2)/2} \prod_{2\leq j\leq n} |1-t_j| \prod_{2\leq i<j\leq n} |t_i - t_j| \, dzdt_2 \ldots dt_n.\end{aligned}$$

Rewriting the bounds on the integrands in terms of the new variables, this gives the set of equations

$$0 \leq z \leq -zt_2 - zt_3 - \ldots - zt_n + 1$$

$$0 \leq zt_2 \leq -zt_3 - \ldots - zt_n + 1$$

$$\vdots$$

$$0 \leq zt_{n-1} \leq -zt_n + 1$$

$$0 \leq zt_n \leq 1.$$

Rewriting the first equation gives

$$0 \leq z \leq \frac{1}{t_2 + \cdots + t_n + 1}$$

and rewriting all other equations while taking the first into consideration gives

$$0 \leq t_j < \infty, \quad 2 \leq j \leq n.$$

Thus, we have that

$$\begin{aligned}
&\lambda^n \left(\mathcal{H}_n^* \left(0, 1\right)\right) \\
=&\frac{1}{n!} \int_0^\infty \cdots \int_0^\infty \int_0^{\frac{1}{t_1 + \cdots + t_{n-1} + 1}} z^{(n^2+n-2)/2} \prod_{2 \leq j \leq n} |1 - t_j| \prod_{2 \leq i < j \leq n} |t_i - t_j| \; dz dt_2 \cdots dt_n \\
=&\frac{2}{n!(n+1)n} \int_0^\infty \cdots \int_0^\infty \frac{\prod_{2 \leq j \leq n} |1 - t_j| \prod_{2 \leq i < j \leq n} |t_i - t_j|}{(t_2 + \cdots + t_n + 1)^{\frac{(n+1)n}{2}}} \; dt_2 \cdots dt_n.
\end{aligned}$$

A final relabeling of the variables gives the result. □

The integral formula obtained in Theorem 3.5.14 seems closely related to the Selberg integral $S(1, 1, 1/2)$. In [2], which considered the volume of the set of random polynomials whose roots were bounded within a ball of a certain height centered at the origin, the authors were able to compute this volume of their sets in terms of Selberg integrals. In [18], which considered random polynomials whose coefficients had bounded height instead, the authors conjectured a relation between Selberg integrals and the volumes of their sets, but were unable to prove this relation.

3.5.4 Coefficients restricted to $[0,\beta]$

We are interested in which results from the previous three subsections hold when all coefficients are restricted to $[0,\beta]$ rather than $[0,1]$. When $\beta \in (0,1)$, using the technique presented in the proof of Lemma 3.5.6, one can show that n-dimensional box $[-\beta/n, 0]^n \subseteq H_n(0,\beta)$. The proof of this relies on the fact that $\beta^k \leq \beta$ for all $k \geq 1$, and the statement does not hold in general when $\beta > 1$. Thus, results from the previous three subsections which rely on this fact will work for $\beta < 1$, and are not necessarily true for $\beta > 1$. Hence, when $\beta < 1$, one can also show

$$\lambda^n(\mathcal{H}_n^*(0,\beta)) > 0$$

by showing that the n-dimensional box $[-\beta/n, 0]^n \subseteq H_n(0,\beta)$; the proof is otherwise analogous to the proof of Lemma 3.5.6.

Moreover, by the proof of Theorem 3.5.1 and applying Theorem 3.4.3, we see that for any $\beta > 0$, the probability that the polynomial $f_n = a_n z^n + \cdots + a_1 z + a_0$ with iid coefficients that are uniformly distributed on $[0,\beta]$ has all real roots is positive. To see this, observe that

$$\lambda^{n+1}(\mathcal{H}_n^+(0,\beta)) = \beta^n \int_0^1 u^n \lambda^n(\mathcal{H}_n^*(0,1/u))\, du,$$

which is positive by the proof of Theorem 3.5.1. Dividing by β^{n+1} gives the desired probability. An analogous lower bound to Corollary 3.5.8 may be obtained for any $\beta > 0$, however an analogous lower bound given in Corollary 3.5.9, and hence also Corollary 3.5.10 only applies when $\beta < 1$; the problem is again whether or not $[-\beta/n, 0]^n \subseteq H_n(0,\beta)$.

We also have the following.

Lemma 3.5.15. *Let $\beta > 0$. For $n \in \mathbb{N}$ and $\xi_1, \ldots, \xi_n \in \mathbb{R}$, consider the set of equations*

$$\begin{cases} -\beta - \xi_n - \cdots - \xi_2 \leq \xi_1 \leq 0 \\ -\beta - \xi_n - \cdots - \xi_3 \leq \xi_2 \leq 0 \\ \qquad \vdots \\ -\beta - \xi_n \leq \xi_{n-1} \leq 0 \\ -\beta \leq \xi_n \leq 0 \end{cases}. \tag{3.12}$$

When $\beta \in (0,1]$, the region

$$H_n(0,\beta) = \{(\xi_1,\dots,\xi_n) \in \mathbb{R}^n \mid 0 \le (-1)^{n-j} S_{n-j}(\xi_1,\dots,\xi_n) \le \beta,\ 0 \le j \le n-1\}$$

and the region defined by

$$\tilde{H}_n(0,\beta) = \{(\xi_1,\dots,\xi_n) \in \mathbb{R}^n \mid \textit{equations } (3.12) \textit{ are satisfied}\}$$

are the same.

When $\beta > 1$, $H_n(0,\beta) \subseteq \tilde{H}_n(0,\beta)$.

Proof. The proof is very similar to the proof of Lemma 3.5.13. Writing each $S_{n-j}(\xi_1,\dots,\xi_n)$ from $H_n(0,\beta)$ in summation notation and multiplying by -1, we may consider the set of equations

$$\begin{cases} -\beta \le \sum_{1\le l_1 \le n} \xi_{l_1} \le 0 \\ -\beta \le -\sum_{1\le l_1<l_2\le n} \xi_{l_1}\xi_{l_2} \le 0 \\ \quad\vdots \\ -\beta \le (-1)^n \sum_{1\le l_1<l_2<\cdots<l_{n-1}\le n} \xi_{l_1}\cdots\xi_{l_{n-1}} \le 0 \\ -\beta \le (-1)^{n+1}\xi_1\xi_2\cdots\xi_n \le 0 \end{cases} \tag{3.13}$$

One can show that for any $\beta > 0$, if the point $(\xi_1,\dots,\xi_n) \in \mathbb{R}^n$ satisfies either equations (3.12) or equations (3.13), then $(\xi_1,\dots,\xi_n) \in [-\beta,0]^n$; the proof is the same as when $\beta = 1$. From there, it is straightforward to show if a point $(\xi_1,\dots,\xi_n)$ satisfies equations (3.13), equations (3.12) also hold; again, the proof is the same. Now suppose that $(\xi_1,\dots,\xi_n)$ satisfies equations (3.12). Examining the proof of Lemma 3.5.13 and using the analagous inductive hypothesis, we see that for k odd,

$$\begin{aligned} S_k(\xi_1,\dots,\xi_n) &= \xi_n S_{k-1}(\xi_1,\dots,\xi_{n-1}) + \xi_{n-1}S_{k-1}(\xi_1,\dots,\xi_{n-2}) + \cdots + \xi_k S_{k-1}(\xi_1,\dots,\xi_{k-1}) \\ &\ge \xi_n\beta + \cdots + \xi_k\beta, \end{aligned}$$

and since $\xi_1,\dots,\xi_n \in [-\beta,0]$, this only shows that $S_k(\xi_1,\dots,\xi_n) \ge -\beta$ when $\beta \le 1$. The proof for when k is even follows similarly. □

With Lemma 3.5.15, we can show that $\lambda^n(\mathcal{H}_n^*(0,\beta))$ is monotonically decreasing in n when $\beta \leq 1$ with a very small adaptation of the proof of Proposition 3.5.11. When $\beta > 1$, we present the following counterexample.

Example 3.5.16. Evaluating the integral given in Theorem 3.4.1 in Mathematica for $\alpha = 0$, $\beta = 20$ and then approximating, we have that

$$\lambda^2(\mathcal{H}_2^*(0,20)) \approx 280.743$$

and

$$\lambda^3(\mathcal{H}_3^*(0,20)) \approx 732.056.$$

However, we make the following conjecture.

Conjecture 3.5.17. *For any $\beta > 1$ and $n \geq 1$,*

$$\frac{\lambda^{n+1}\left(\mathcal{H}_{n+1}^*(0,\beta)\right)}{\beta^{n+1}} \leq \frac{\lambda^n\left(\mathcal{H}_n^*(0,\beta)\right)}{\beta^n}.$$

In other words, the probability that the polynomial

$$f_n(z) = a_n z^n + \cdots + a_1 z + a_0$$

with iid coefficients that are uniformly distributed on $[0,\beta]$ has all real roots is monotonically decreasing in n.

If Conjecture 3.5.17 is true, applying it together with Theorem 3.4.3, one would be able to show that

$$\begin{aligned}
\lambda^{n+1}\left(\mathcal{H}_n^+([0,1])\right) &= \int_0^1 u^n \lambda^n\left(\mathcal{H}_n^*\left(\left[0,\frac{1}{u}\right]\right)\right)\, du \\
&\leq \int_0^1 u^{n-1}\lambda^{n-1}\left(\mathcal{H}_{n-1}^*\left(\left[0,\frac{1}{u}\right]\right)\right)\, du \\
&= \lambda^n\left(\mathcal{H}_{n-1}^+([0,1])\right).
\end{aligned}$$

This would mean that the probability that a degree n non-monic generalized bounded height model polynomial with iid coefficients on $[0,1]$ has all real roots is monotonically decreasing in n.

By the equivalence between these polynomials and the generalized large box model polynomials given in Theorem 3.1.1, this implies that as $N \to \infty$, the probability that the polynomial

$$f_n(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

with coefficients $a_0, \ldots, a_{n-1}$ uniformly distributed on $[0, N] \cap \mathbb{Z}$ and jointly independent has all real roots is also monotonically decreasing in n.

3.5.5 Positivity of $\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right)$ when every coefficient is allowed to be zero

The main result of this section is the following theorem.

Theorem 3.5.18. *For $0 \le j \le n$, let $\alpha_j, \beta_j \in \mathbb{R}$, with $\alpha_j < \beta_j$. Suppose further that $0 \in [\alpha_j, \beta_j]$ for each $0 \le j \le n$. Then $\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right) > 0$.*

We begin with two necessary lemmas.

Lemma 3.5.19. *Define the recursive sequence where $k_1 = 1$, and $k_l = k_{l-1} + l$, $2 \le l \le n$. Let $c \in \mathbb{R}$ and suppose that $c \neq 0$. Let*

$$x_1 = \frac{c}{n^{k_1}}, x_2 = \frac{c}{n^{k_2}}, \ldots, x_n = \frac{c}{n^{k_n}}.$$

Then for $n \ge 2$, the equations

$$\begin{aligned}
|x_1| &> \sum_{2 \le l_1 \le n} |x_{l_1}| \\
|x_1 x_2| &> \sum_{\substack{1 \le l_1 < l_2 \le n \\ (l_1, l_2) \neq (1,2)}} |x_i x_j| \\
&\vdots \\
|x_1 x_2 \cdots x_{n-1}| &> \sum_{\substack{1 \le l_1 < \cdots < l_{n-1} \le n \\ (l_1, \ldots, l_{n-1}) \neq (1, 2, \ldots, n-1)}} |x_{l_1} x_{l_2} \cdots x_{l_{n-1}}|
\end{aligned}$$

are satisfied.

Proof. First, observe that dividing both sides of the equation

$$|x_1 x_2 \cdots x_k| > \sum_{\substack{1 \le l_1 < \cdots < l_k \le n \\ (l_1, \ldots, l_k) \ne (1,2,\ldots,k)}} |x_{l_1} x_{l_2} \cdots x_{l_k}|$$

by $|c|^k$ preserves the inequality, so it suffices to show that the equations above hold with the choices

$$x_1 = \frac{1}{n^{k_1}}, x_2 = \frac{1}{n^{k_2}}, \ldots, x_n = \frac{1}{n^{k_n}}.$$

Let $x_1, \ldots, x_n$ be these values for the remainder of the proof. Next, notice that $|x_1| > |x_2| > \cdots > |x_n|$ since k_l is increasing for $1 \le l \le n$, so it suffices to show that

$$|x_1| > \binom{n}{1} |x_2|$$

$$|x_1 x_2| > \binom{n}{2} |x_1 x_3|$$

$$\vdots$$

$$|x_1 x_2 \cdots x_{n-1}| > \binom{n}{n-1} |x_1 x_2 \cdots x_{n-2} x_n|.$$

In general, this is

$$|x_1 x_2 \cdots x_l| > \binom{n}{l} |x_1 x_2 \cdots x_{l-2} x_{l+1}|.$$

Cancelling the $|x_1 \cdots x_{l-1}|$ from both sides, it suffices to show that

$$|x_l| > \binom{n}{l} |x_{l+1}|.$$

Finally, for $1 \le l \le n$, we have

$$\frac{|x_l|}{\binom{n}{l}} = \frac{l!}{n^{k_l} n(n-1)\cdots(n-l+1)} > \frac{1}{n^{k_l} n^{l+1}} = \frac{1}{n^{k_l+(l+1)}} = |x_{l+1}|$$

where the inequality follows since $l! \ge 1$ and $n(n-1)\cdots(n-l+1) \le n^l < n^{l+1}$. □

Lemma 3.5.20. *The coefficients described in the statement of Lemma 3.5.19 are* $k_i = \frac{i(i+1)}{2}$ *for* $1 \le i \le n$.

Proof. We claim that $k_l = \sum_{j=1}^{l} j$. The statement clearly holds for $l = 1$. Suppose that the statement is true for the first i integers. Then

$$k_{i+1} = k_i + (i+1) = \left(\sum_{j=1}^{i} j\right) + (i+1) = \sum_{j=1}^{i+1} j,$$

so the claim holds by the principle of mathematical induction. Since the sum of the first i integers is given by

$$\sum_{j=0}^{i} j = \frac{i(i+1)}{2},$$

the proof of the lemma is complete. □

We can now prove that $\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right) > 0$ when $0 \in [\alpha_j, \beta_j]$ for each $0 \leq j \leq n$.

Proof of Theorem 3.5.18. Let c be the minimum of the set $\{|\alpha_0|, |\beta_0|, \ldots, |\alpha_n|, |\beta_n|\} \setminus \{0\}$. Then clearly $c > 0$. If $c \geq 1$, we can restrict ourselves to a subset of $[\alpha_0, \beta_0] \times \cdots \times [\alpha_n, \beta_n]$ so that $c < 1$, for instance by taking

$$\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right] \cap [-1/2, 1/2] = [\alpha_0, \beta_0] \cap [-1/2, 1/2] \times \cdots \times [\alpha_n, \beta_n] \cap [-1/2, 1/2].$$

Then showing

$$\lambda^{n+1}\left(\mathcal{H}_n\left(\overrightarrow{[\alpha,\beta]} \cap [-1/2, 1/2]\right)\right) > 0$$

implies that

$$\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right) > 0.$$

This will simplify the proof, so suppose that $c < 1$.

We begin by showing that $\lambda^n\left(\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right) > 0$. To do this, we will argue that the set $H_n\left(\left[\overrightarrow{\alpha,\beta}\right]\right)$ has positive measure. Recall that

$$H_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right) = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid \alpha_j \leq (-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n) \leq \beta_j,\ 0 \leq j \leq n-1\}.$$

For every $0 \leq j \leq n-1$, notice that either $[0, c] \subseteq [\alpha_j, \beta_j]$ or $[-c, 0] \subseteq [\alpha_j, \beta_j]$ (or both) by the definition of c and since all intervals were assumed to contain 0. For each $0 \leq j \leq n-1$, define the set

$$[c_{\alpha_j}, c_{\beta_j}] = \begin{cases} [0,c] & \text{if } [0,c] \subseteq [\alpha_j, \beta_j] \\ [-c,0] & \text{otherwise} \end{cases}.$$

Notice that

$$[c_{\alpha_0}, c_{\beta_0}] \times \cdots \times [c_{\alpha_{n-1}}, c_{\beta_{n-1}}] \subseteq [\alpha_0, \beta_0] \times \cdots \times [\alpha_{n-1}, \beta_{n-1}].$$

We will begin by showing that

$$H_n\left(\overrightarrow{[c_\alpha, c_\beta]}\right) = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid c_{\alpha_j} \le (-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n) \le c_{\beta_j},\ 0 \le j \le n-1\}$$

has positive measure. Consider the point $(\xi_1, \ldots, \xi_n)$ where

$$\xi_1 = -\frac{c}{n},\ \xi_2 = -\frac{c}{n^3}, \ldots,\ \xi_n = -\frac{c}{n^{\left(\frac{n(n+1)}{2}\right)}}. \tag{3.14}$$

Let $n \ge 2$. Then since $\frac{c}{n} = |\xi_1| > |\xi_2| > \cdots > |\xi_n|$, we have that

$$|S_k(\xi_1, \ldots, \xi_n)| = \left| \sum_{1 \le l_1 < l_2 < \cdots < l_k \le n} \xi_{l_1} \cdots \xi_{l_k} \right| \tag{3.15}$$

$$\le \sum_{1 \le l_1 < l_2 < \cdots < l_k \le n} |\xi_{l_1} \cdots \xi_{l_k}| \tag{3.16}$$

$$< \binom{n}{k} |\xi_1|^k \tag{3.17}$$

$$= \binom{n}{k} \left(\frac{c}{n}\right)^k \tag{3.18}$$

$$= \frac{n(n-1)\cdots(n-k+1)c^k}{k!n^k} \tag{3.19}$$

$$\le \left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right) c^k \tag{3.20}$$

$$< c \tag{3.21}$$

since $c < 1$ so $c^k < c$. Therefore,

$$-c < (-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n) < c, \quad 0 \le j \le n-1, \tag{3.22}$$

and it remains to show that each $(-1)^{n-j} S_{n-j}(\xi_1, \ldots, \xi_n)$ is actually contained in either $(-c, 0)$ or $(0, c)$, depending on $[c_{\alpha_j}, c_{\beta_j}]$ for $0 \le j \le n-1$. For this purpose, consider instead the set of equations

$$c_{\alpha_{n-1}} < -\sum_{1\le l_1\le n} y_{l_1} < c_{\beta_{n-1}} \tag{3.23}$$

$$c_{\alpha_{n-2}} < \sum_{1\le l_1<l_2\le n} y_{l_1}y_{l_2} < c_{\beta_{n-2}} \tag{3.24}$$

$$\vdots$$

$$c_{\alpha_0} < (-1)^n y_1 y_2 \cdots y_n < c_{\beta_0}. \tag{3.25}$$

We show that changing only the signs of $\xi_1, \dots, \xi_n$ given in (3.14) will let us find a point $(y_1, \dots, y_n)$ such that the above equations are satisfied.

If $c_{\alpha_{n-1}} = 0$ and $c_{\beta_{n-1}} = c$, let $y_1 = \xi_1 = -\frac{c}{n}$, and $y_2, \dots, y_n$ be chosen in any way such that $|y_2| = |\xi_2|, \dots, |y_n| = |\xi_n|$. Then $-y_1 = |y_1|$, and

$$0 < -y_1 - \sum_{j=2}^{n} |y_j| \le -y_1 - y_2 - \cdots - y_n \le -y_1 + \sum_{j=2}^{n} |y_j| < c,$$

where the first inequality follows from Lemma 3.5.19 and the last inequality follows from the bound on the absolute value of the terms of $S_1(\xi_1, \dots, \xi_n)$ given by (3.16) and (3.21).

If $c_{\alpha_{n-1}} = -c$ and $c_{\beta_{n-1}} = 0$, we proceed similarly by letting $y_1 = -\xi_1 = \frac{c}{n}$, and letting $y_2, \dots, y_n$ be chosen in any way such that $|y_2| = |\xi_2|, \dots, |y_n| = |\xi_n|$. Then now $-y_1 = -|y_1|$, and

$$-c < -y_1 - \sum_{j=2}^{n} |y_j| \le -y_1 - y_2 - \cdots - y_n \le -y_1 + \sum_{j=2}^{n} |y_j| < 0,$$

where this time the first inequality follows from negating (3.16) and (3.21) and the last inequality follows from Lemma 3.5.19.

Suppose by induction that the first k variables $y_1, \dots, y_k$ have been fixed such that the first k equations of $c_{\alpha_{n-k}} < (-1)^{n-k} S_k(y_1, \dots, y_n) < c_{\beta_{n-k}}$ have been satisfied by letting $y_1 = \pm\xi_1, \dots, y_k = \pm\xi_k$ in the appropriate manner and letting $y_{k+1}, \dots, y_n$ be chosen in any way such that $|y_{k+1}| = |\xi_{k+1}|, \dots, |y_n| = |\xi_n|$.

We want to show

$$c_{\alpha_{n-k-1}} < (-1)^{n-k-1} S_{k+1}(y_1, \dots, y_n) < c_{\beta_{n-k-1}}.$$

We consider four possibilities:

If $c_{\alpha_{n-k-1}} = c_{\alpha_{n-k}} = 0$, then let $y_{k+1} = -|\xi_{k+1}|$. Then

$$0 < (-1)^{n-k} y_1 \cdots y_k$$

implies that

$$0 < (-1)^{n-k-1} y_1 \cdots y_k y_{k+1}$$

also since the extra factor of -1 and the negative coefficient of y_{k+1} cancel.

If $c_{\alpha_{n-k-1}} = c_{\alpha_{n-k}} = -c$, then let $y_{k+1} = -|\xi_{k+1}|$. Then

$$(-1)^{n-k} y_1 \cdots y_k < 0$$

implies that

$$(-1)^{n-k-1} y_1 \cdots y_k y_{k+1} < 0$$

also since the extra factor of -1 and the negative coefficient of y_{k+1} cancel.

If $c_{\alpha_{n-k-1}} = 0$ and $c_{\alpha_{n-k}} = -c$, then let $y_{k+1} = |\xi_{k+1}|$. Then

$$(-1)^{n-k} y_1 \cdots y_k < 0$$

implies that

$$0 < (-1)^{n-k-1} y_1 \cdots y_k y_{k+1}$$

since the single factor of -1 negates the negative sign.

If $c_{\alpha_{n-k-1}} = 0$ and $c_{\alpha_{n-k}} = -c$, then let $y_{k+1} = |\xi_{k+1}|$. Then

$$0 < (-1)^{n-k} y_1 \cdots y_k$$

implies that

$$(-1)^{n-k-1} y_1 \cdots y_k y_{k+1} < 0$$

since the single factor of -1 negates the negative sign. By the principle of mathematical induction, we have shown that either $(-1)^{n-j} S_{n-j}(y_1, \ldots, y_n)$ is negative when $[c_{\alpha_j}, c_{\beta_j}] = [-c, 0]$ or

$(-1)^{n-j}S_{n-j}(y_1,\ldots,y_n)$ is positive when $[c_{\alpha_j}, c_{\beta_j}] = [0,c]$. To see that $(-1)^{n-j}S_{n-j}(y_1,\ldots,y_n)$ is actually contained in $[c_{\alpha_j}, c_{\beta_j}]$, observe that, as in equations (3.15) through (3.21),

$$|S_{n-j}(y_1,\ldots,y_n)| < \binom{n}{n-j}|y_1|^{n-j} < c.$$

Thus, the point $(y_1,\ldots,y_n)$ satisfies (3.23), (3.24), and (3.25); this means that the intersection of the open sets defined by equations (3.23), (3.24), and (3.25) is a non-empty subset of $\mathbb{R}^n$; call this set H_C. In particular, since the intersection of open sets is also open, there exists a small ball around $(y_1,\ldots,y_n)$ contained in H_C. Examining the definitions of both sets, we see that $H_C \subseteq H_n\left(\overrightarrow{c_\alpha, c_\beta}\right)$, so this small ball around $(y_1,\ldots,y_n)$ is also contained in $H_n\left(\overrightarrow{c_\alpha, c_\beta}\right)$, showing that $H_n\left(\overrightarrow{c_\alpha, c_\beta}\right)$ and hence $H_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)$ has positive measure. By Theorem 3.4.1, we have that

$$\lambda^n\left(\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)\right) = \frac{1}{n!}\int_{H_n\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)} \prod_{1\le j<k\le n} |\xi_j - \xi_k|\; d\xi_1 \ldots d\xi_n,$$

so since $\prod_{1\le j<k\le n} |\xi_j - \xi_k| > 0$ except for on a set of measure zero, we see that

$$\lambda^n\left(\mathcal{H}_n^*\left(\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]\right)\right) > 0$$

as well.

Finally, we wish to show that $\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right) > 0$. Let us break this into the calculations of $\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right)$ and $\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right)$. Recall that we assumed that $0 \in [\alpha_n, \beta_n]$, so either $\alpha_n < 0$, $\beta_n > 0$, or both.

For $\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right)$, if $\beta_n > 0$, we have from Theorem 3.4.3 that

$$\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right) = \beta_n^{n+1}\int_{\max\left\{0, \frac{\alpha_n}{\beta_n}\right\}}^1 u^n \lambda^n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n u}\right)\right) du.$$

In the calculation of the integral above, since $\beta_n > 0$, $u > 0$ (since we do not need to consider the endpoints of the bounds of integration), and the interval $[\alpha_j, \beta_j]$ was assumed to contain 0 for each $0 \le j \le n$, we see that the sets $\left[\frac{\alpha_j}{\beta_n u}, \frac{\beta_j}{\beta_n u}\right]$ also contain 0 for each $0 \le j \le n$. Hence, $\lambda^n\left(\mathcal{H}_n^*\left(\frac{\left[\overrightarrow{\alpha}, \overrightarrow{\beta}\right]}{\beta_n u}\right)\right) > 0$ for all $u \in (0,1)$, so $\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right) > 0$ as well.

If $\alpha_n < 0$, we can see that $\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha, \beta}\right]\right)\right) > 0$ by an analogous argument.

Since at least one of $\lambda^{n+1}\left(\mathcal{H}_n^+\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right)$ or $\lambda^{n+1}\left(\mathcal{H}_n^-\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right)$ must be positive, by (3.8)

$$\lambda^{n+1}\left(\mathcal{H}_n\left(\left[\overrightarrow{\alpha,\beta}\right]\right)\right) > 0$$

as desired. □

Chapter 4

Probability of all real eigenvalues for generalized bounded height model random matrices

4.1 Introduction and main results

We now turn our attention to random matrices whose entries are given by the generalized bounded height model. In particular, let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for $1 \leq i, j \leq n$ and consider the n-by-n random matrices whose entries a_{ij} are independently and uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$. We wish to find the probability that all eigenvalues of a random matrix from this ensemble are real. In the case of the random matrix ensemble whose elements are real iid standard Gaussians, this was accomplished by Edelman in [54].

In Chapter 5, Theorem 5.1.1 shows that as $N \to \infty$, the probability that a random matrix with entries independently and uniformly distributed on $[\alpha_{ij}N, \beta_{ij}N] \cap \mathbb{Z}$ for each $1 \leq i, j \leq n$ has all real eigenvalues converges to the probability that a random matrix with entries independently and uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$ for each $1 \leq i, j \leq n$ has all real eigenvalues. Therefore, the results of this chapter will allow us to calculate the probability that, as $N \to \infty$, a generalized large box model random matrix has all real eigenvalues.

For a 2-by-2 random matrix from the bounded height model ensemble, we have the following main result; the proof is the content of Section 4.6.

Theorem 4.1.1. *Let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for $1 \leq i, j \leq 2$ and consider the 2-by-2 random matrix A whose entries a_{ij} are independently and uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$. For any real λ_1, λ_2,* *define*

$$\begin{aligned}A_{\lambda_1,\lambda_2} = \{(r,q) \mid\ & q \in (0,1),\\ & (\lambda_1-\lambda_2)q^2+\lambda_2+rq\sqrt{1-q^2} \in (\alpha_{11},\beta_{11}),\\ & (\lambda_1-\lambda_2)q\sqrt{1-q^2}+q^2r \in (-\beta_{12},-\alpha_{12})\cup(\alpha_{12},\beta_{12}),\\ & (\lambda_1-\lambda_2)q\sqrt{1-q^2}+r-q^2r \in (-\beta_{21},-\alpha_{21})\cup(\alpha_{21},\beta_{21}),\\ & -(\lambda_1-\lambda_2)q^2+\lambda_1-rq\sqrt{1-q^2} \in (\alpha_{22},\beta_{22})\}.\end{aligned}$$

The probability that all eigenvalues of A are real is given by

$$\left(\prod_{1\le i\le j\le 2}\frac{1}{(\beta_{ij}-\alpha_{ij})}\right)\int_{-\infty}^{\infty}\int_{\lambda_2}^{\infty}\int_{A_{\lambda_1,\lambda_2}} |\lambda_1-\lambda_2|\ drdqd\lambda_1 d\lambda_2.$$

To obtain an analogous result for an n-by-n random matrix from the bounded height model ensemble, we consider such matrices A with all real eigenvalues, and let $A = QRQ^T$ be decomposed into its real Schur decomposition (discussed in detail in Section 4.3). Proposition 4.5.3 shows that when all eigenvalues of A are real and distinct, this decomposition can be made unique by specifying that Q is an n-by-n orthogonal matrix where the first non-zero entry in each column is positive and

$$R = \begin{pmatrix}\lambda_1 & r_{12} & \cdots & r_{1n}\\ & \ddots & \ddots & \vdots\\ & & \ddots & r_{n-1,n}\\ & & & \lambda_n\end{pmatrix},$$

where $\lambda_1,\ldots,\lambda_n$ are the real eigenvalues of A listed in decreasing order, $r_{12},\ldots,r_{n-1,n}\in\mathbb{R}$, and the entries below the main diagonal are all zero.

Let

$$\rho = \{r_{12}, r_{13}, \ldots, r_{1n}, r_{23}, \ldots r_{2n}, \ldots, r_{n-1,n}\}.$$

The algorithm given by Raffenetti and Ruedenberg in [135] (and discussed in detail in Section 4.7) allows us to see that Q is specified by $\frac{1}{2}n(n-1)$ independent angular parameters

$$\Theta = \{\theta_{12}, \theta_{13}, \ldots, \theta_{1n}, \theta_{23}, \ldots \theta_{2n}, \ldots, \theta_{n-1,n}\},$$

where each angle in Θ is contained in a subset of $[0, 2\pi)$. Furthermore, letting $(d\tilde{R})$ denote the wedge product over the independent strictly upper triangular elements of R and letting $(Q^T dQ)$ denote the wedge product over the independent elements of the antisymmetric matrix $Q^T dQ$ (discussed in detail in Section 4.2), we arrive at the following main result, whose proof is the content of Section 4.7.

Theorem 4.1.2. *Let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for $1 \leq i, j \leq n$ and consider the n-by-n random matrix A whose entries a_{ij} are independently and uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$. For any real $\lambda_1, \lambda_2, \ldots, \lambda_n$, define*

$$\begin{aligned} A_{\lambda_1,\ldots,\lambda_n} = \{(\rho, \Theta) \mid\ & Q_{*j} > 0 \text{ for } 1 \leq j \leq n, \\ & Q^T Q = I_n, \\ & \alpha_{ij} < (QRQ^T)_{ij} < \beta_{ij} \text{ for } 1 \leq i, j \leq n\}, \end{aligned}$$

where Q is an n-by-n orthogonal matrix parameterized by the Raffenetti and Ruedenberg algorithm (discussed in Section 4.7) in the variables

$$\Theta = \{\theta_{12}, \theta_{13}, \ldots, \theta_{1n}, \theta_{23}, \ldots \theta_{2n}, \ldots, \theta_{n-1,n}\},$$

where each angle is contained in a subset of $[0, 2\pi)$, Q_{*j} is the first non-zero entry in the j-th column of Q, R is the upper triangular matrix with all real entries

$$R = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ & \ddots & \ddots & \vdots \\ & & \ddots & r_{n-1,n} \\ & & & \lambda_n \end{pmatrix},$$

and

$$\rho = \{r_{12}, r_{13}, \ldots, r_{1n}, r_{23}, \ldots r_{2n}, \ldots, r_{n-1,n}\}.$$

The probability that all eigenvalues of A are real is given by

$$\left(\prod_{1 \leq i \leq j \leq n} \frac{1}{(\beta_{ij} - \alpha_{ij})}\right) \int_{-\infty}^{\infty} \int_{\lambda_{n-1}}^{\infty} \cdots \int_{\lambda_2}^{\infty} \int_{A_{\lambda_1,\ldots,\lambda_n}} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j| \ (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n.$$

We begin by introducing the notation used by Edelman before delving into his method to prove the theorems.

4.2 The wedge product

In this section, we define the wedge product, used by Edelman to compute the Jacobian which arises when performing a change of variables on a matrix. Complete details regarding the construction of the wedge product in general can be found in algebra or differential geometry texts, such as [47, 104]. Theorem 4.2.1 below is a compilation of the requisite terminology and results needed to define the wedge product for differential forms on $\mathbb{R}^n$. All statements therein are taken directly from those provided by Shifrin in [147].

Theorem 4.2.1 (The wedge product, see Chapters 7 and 8 in [147])**.** *A function* $\varphi : (\mathbb{R}^n)^m \to \mathbb{R}$ *is called **multilinear** if for each* $1 \leq i \leq m$*, we have*

$$\varphi(v_1, \ldots, v_{i-1}, cv_i + c'v_i', v_{i+1}, \ldots, v_m)$$
$$= c\varphi(v_1, \ldots, v_i, \ldots, v_m) + c'\varphi(v_1, \ldots, v_i', \ldots, v_m)$$

for all $v_i, v_i' \in \mathbb{R}^n$ *and* $c, c' \in \mathbb{R}$*. In addition,* φ *is called **alternating** if* φ *changes sign whenever any pair of vectors* $v_1, \ldots, v_m \in \mathbb{R}^n$ *are exchanged, i.e., for any* $1 \leq i < j \leq m$,

$$\varphi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m) = -\varphi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m).$$

In particular, this implies that $\varphi(v_1, v_2, \ldots, v_m) = 0$ *whenever* $v_i = v_{i+1}$ *for some* $1 \leq i \leq m-1$.

Define the n *linear maps* $dx_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, n$*, by assigning to each vector* $v = (v_1, v_2, \ldots, v_n)^T \in \mathbb{R}^n$ *the* i*-th component of* v*, i.e.,* $dx_i(v) = v_i$*. The set of linear maps from* $\mathbb{R}^n$ *to* $\mathbb{R}$ *is an* n*-dimensional vector space, denoted* $(\mathbb{R}^n)^*$*, with basis* $\{dx_1, \ldots, dx_n\}$*. Let* $\{e_1, \ldots, e_n\}$ *be the standard basis for* $\mathbb{R}^n$*. If* $I = (i_1, \ldots, i_k)$ *is an ordered* k*-tuple, define* $dx_I : (\mathbb{R}^n)^k \to \mathbb{R}$ *by*

$$dx_I(v_1, \ldots, v_k) = \det \begin{pmatrix} dx_{i_1}(v_1) & \cdots & dx_{i_1}(v_k) \\ \vdots & \ddots & \vdots \\ dx_{i_k}(v_1) & \cdots & dx_{i_k}(v_k) \end{pmatrix}.$$

As is the case with the determinant, dx_I defines an alternating, multilinear function of k vectors in $\mathbb{R}^n$. If we write

$$v_i = \begin{pmatrix} v_{i,1} \\ v_{i,2} \\ \vdots \\ v_{i,n} \end{pmatrix}, \quad i = 1, \ldots, k,$$

then

$$dx_I(v_1, \ldots, v_k) = \det \begin{pmatrix} v_{1,i_1} & \cdots & v_{k,i_1} \\ \vdots & \ddots & \vdots \\ v_{1,i_k} & \cdots & v_{k,i_k} \end{pmatrix}.$$

*When $i_1 < i_2 < \cdots < i_k$, this the determinant of the $k \times k$ matrix obtained by taking rows $i_1, \ldots, i_k$ of the matrix which has column vectors $v_1, v_2, \ldots, v_k$. When $i_1 < i_2 < \cdots < i_k$, we say that the ordered k-tuple $I = (i_1, \ldots, i_k)$ is **strictly increasing**. If I is a k-tuple with no repeated index, we denote by $I^<$ the associated strictly increasing k-tuple. The set of dx_I with I strictly increasing forms a basis for the vector space of alternating multilinear functions from $(\mathbb{R}^n)^k$ to $\mathbb{R}$, denoted $\Lambda^k(\mathbb{R}^n)^*$. If I and J are ordered k- and l-tuples respectively, we define $dx_I \wedge dx_J = dx_{(I,J)}$, where by (I, J) we mean the ordered $(k+l)$-tuple obtained by concatenating I and J. If $\omega = \sum a_I dx_I$ and $\eta = \sum b_J dx_J$, then we extend by linearity and set $\omega \wedge \eta = \sum (a_I b_J) dx_I \wedge dx_J = \sum (a_I b_J) dx_{(I,J)}$. This is called the **wedge product** of ω and η.*

*A **differential** 0**-form** on $\mathbb{R}^n$ is a smooth function, i.e., a function whose partial derivatives of all orders exist. A **differential** n**-form** on $\mathbb{R}^n$ is an expression of the form*

$$\omega = f(x) dx_1 \wedge \cdots \wedge dx_n$$

*for some smooth function f. A **differential** k**-form** on $\mathbb{R}^n$ is an expression*

$$\omega = \sum_{\text{strictly increasing } k\text{-tuples } I} f_I(x) dx_I = \sum_{i_1 < \cdots < i_k} f_I dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

for some smooth functions f_I. The set of differential k-forms on $\mathbb{R}^n$ is a vector space and denoted by $\mathcal{A}^k(\mathbb{R}^n)$. Let $U \subseteq \mathbb{R}^n$ be an open set. Let $\omega \in \mathcal{A}^k(U), \eta \in \mathcal{A}^l(U)$, and $\phi \in \mathcal{A}^m(U)$, with $k, l, m \in \mathbb{N} \cup \{0\}$. Let c be a smooth function. The following algebraic properties hold.

(1) When $k = l = m$, $\omega + \eta = \eta + \omega$ and $(\omega + \eta) + \phi = \omega + (\eta + \phi)$.

(2) Antisymmetry: $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$.

(3) Associativty: $(\omega \wedge \eta) \wedge \phi = \omega \wedge (\eta \wedge \phi)$.

(4) Bilinearity: $c(\omega \wedge \eta) = (c\omega) \wedge \eta$, and when $k = l$, $(\omega + \eta) \wedge \phi = (\omega \wedge \phi) + (\eta \wedge \phi)$.

Given an n-form $\omega = f(x) dx_1 \wedge \cdots \wedge dx_n$ on a region $\Omega \subseteq \mathbb{R}^n$, we define $\int_\Omega \omega = \int_\Omega f dV$, where dV denotes the volume element on $\mathbb{R}^n$. In other words,

$$\int_\Omega f(x_1, \ldots, x_n) dx_1 \wedge \cdots \wedge dx_n = \int_\Omega f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n.$$

For our purposes, we need the following result, which shows the connection between the wedge product and the Jacobian for the change of variables that arises when performing matrix decompositions. The following theorem is the combination of three definitions by Forrester [65], though proofs of results stated therein can be found in Section 4.1 of [85] or Chapter 2 of [121].

Theorem 4.2.2 (See Definitions 1.2.1, 1.2.2, 1.2.3 in [65])**.** *Denote the wedge product by*

$$du_1 \wedge \cdots \wedge du_n = \bigwedge_{j=1}^{n} du_j. \tag{4.1}$$

Consider the integral of an n-form $f(u_1, \ldots, u_n) du_1 \wedge \cdots \wedge du_n$ over a region $\Omega \subseteq \mathbb{R}^n$,

$$\int_\Omega f(u_1, \ldots, u_n)\ du_1 \wedge \cdots \wedge du_n.$$

When changing variables from $\{u_1, \ldots, u_n\}$ to $\{v_1, \ldots, v_n\}$, the fundamental formula

$$du_i = \sum_{l=1}^{n} \frac{\partial u_i}{\partial v_l} dv_l$$

applies. Substituting this in (4.1) *and simplifying by using the associativity, bilinearity, and anti-symmetry of the wedge product, it follows that*

$$\bigwedge_{j=1}^{n} du_j = \det\left(\frac{\partial u_i}{\partial v_j}\right)_{i,j=1,\dots,n} \bigwedge_{j=1}^{n} dv_j. \tag{4.2}$$

The determinant in (4.2) *is precisely the Jacobian for the change of variables.*

*For any n-by-n matrix X, the **matrix of differentials** is defined as*

$$dX = \begin{pmatrix} dx_{11} & dx_{12} & \cdots & dx_{1n} \\ dx_{21} & dx_{22} & \cdots & dx_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ dx_{n1} & dx_{n2} & \cdots & dx_{nn} \end{pmatrix}.$$

For an arbitrary n-by-n matrix X, the symbol (dX) denotes the wedge product of the n^2 independent elements of dX. If X is an n-by-n symmetric matrix, the symbol (dX) denotes the wedge product of the $\frac{1}{2}n(n+1)$ distinct elements of dX. If X is an n-by-n antisymmetric matrix (meaning that $X = -X^T$), then (dX) denotes the wedge product of the $\frac{1}{2}n(n-1)$ distinct elements of dX. If X is upper triangular, then (dX) denotes the wedge over all upper triangular elements.

4.3 Real Schur decomposition

The real Schur decomposition is a matrix factorization that expresses a matrix as an orthogonal conjugation of a block upper triangular matrix; when all eigenvalues of the original matrix are real, the block upper triangular matrix is simply upper triangular, and the diagonal entries of the upper triangular matrix are conveniently the eigenvalues of the original matrix.

Definition 4.3.1 (Real Schur Decomposition, see equations 15.200-15.202 in [65] or Theorem 7.4.1 in [74])**.** Let A be an n-by-n matrix with real entries and exactly k real eigenvalues (where n and k have the same parity). Then the **real Schur decomposition** of A is given by

$$A = QRQ^T,$$

where Q is an n-by-n orthogonal matrix and

$$R = \begin{pmatrix} \lambda_1 & \cdots & R_{1,k} & R_{1,k+1} & \cdots & R_{1,m} \\ & \ddots & \vdots & \vdots & \cdots & \vdots \\ & & \lambda_k & R_{k,k+1} & \cdots & R_{k,m} \\ & & & Z_{k+1} & \cdots & R_{k+1,m} \\ & & & & \ddots & \vdots \\ & & & & & Z_m \end{pmatrix}$$

where all elements not explicitly shown are zero, $m = (n+k)/2$, and R_{ij} is of size $p \times q$ with

$$p \times q = \begin{cases} 1 \times 1 & \text{if } \ i \leq k, \ j \leq k, \\ 1 \times 2 & \text{if } \ i \leq k, \ j > k, \\ 2 \times 1 & \text{if } \ i > k, \ j \leq k, \\ 2 \times 2 & \text{if } \ i > k, \ j > k. \end{cases}$$

Furthermore, $\lambda_1, \ldots, \lambda_k$ are the real eigenvalues of A, while the 2×2 matrices $Z_{k+1}, \ldots, Z_m$ have the structure

$$Z_j = \begin{pmatrix} x_j & b_j \\ -c_j & x_j \end{pmatrix}, \quad b_j, c_j > 0,$$

such that the complex eigenvalues of A are $x_j \pm i y_j$, where $y_j = \sqrt{b_j c_j}$.

4.4 Edelman's method

In Edelman's paper [54], Edelman considers an n-by-n random matrix A whose entries are given by iid standard Gaussian random variables. He obtains the probability that A has exactly k real eigenvalues. His method involves first writing the matrix A in terms of its real Schur decomposition, and then computing the Jacobian of the associated change of variables. This result is summarized by the following proposition, which does not assume a specific distribution for the entries of A.

Proposition 4.4.1 (Proposition 15.10.1 from [65] or Theorem 5.1 from [54])**.** *Let A be an n-by-n matrix with real entries written in real Schur form $A = QRQ^T$, as in Definition 4.3.1. Let k denote the number of real eigenvalues of A. Interpreting this decomposition as a change of variables, the Jacobian of the change of variables is given by*

$$(dA) = 2^{(n-k)/2} \prod_{j<p} |\lambda(R_{pp}) - \lambda(R_{jj})| \; (d\tilde{R})(Q^T dQ)$$
$$\times \prod_{l=k+1}^{(n+k)/2} |b_l - c_l| \; d\lambda_1 \ldots d\lambda_k (dZ),$$

where $\lambda(R_{ll}) = \lambda_l$ are the real eigenvalues of A for $l \leq k$, while $\lambda(R_{ll}) = x_l \pm iy_l$ are the complex eigenvalues of A for $l > k$, $\tilde{R}$ denotes the strictly upper triangular part of R,

$$(Q^T dQ) = \bigwedge_{i>j} q_i^T dq_j,$$

and

$$(dZ) = \prod_{j=k+1}^{(n+k)/2} dx_j db_j dc_j,$$

where all variables are defined as in Definition 4.3.1.

Let $\mathbb{P}(A)\; dA$ denote the joint probability density function for the random matrix A on the space of n-by-n real matrices with respect to the Lebesgue measure on $\mathbb{R}^{n^2}$; when all entries are independent, $\mathbb{P}(A)$ is simply the product of the pdf of each entry of A. By integrating $\mathbb{P}(A)\; dA$ over all variables except for the eigenvalues and normalizing, Edelman obtains the joint pdf of the eigenvalues of A, given that k eigenvalues are real. Continuing to integrate the eigenvalues over the real line, Edelman obtains the probability that exactly k eigenvalues of A are real, leading to the following main theorem.

Theorem 4.4.2 (Theorem 6.1 in [54])**.** *Let A be an n-by-n random matrix whose elements are independent random variables with standard normal distributions. Let $\mathcal{A}_k$ denote that set of matrices A with exactly k real eigenvalues. Let $p_{n,k}$ denote the probability that $A \in \mathcal{A}_k$. The ordered real eigenvalues of A are denoted $\lambda_i, i = 1, \ldots, k$, while the $l = (n-k)/2$ ordered complex eigenvalue* *pairs are denoted* $x_i \pm y_i\sqrt{-1}, i = k+1, \dots, m$. *Let*

$$c_{n,k} = \frac{2^{2l-n(n+1)/4}}{\prod_{i=1}^{n} \Gamma(i/2)}.$$

The joint distribution of the real and complex eigenvalues given that A *has* k *real eigenvalues is*

$$p_{n,k}^{-1} c_{n,k} \left(\prod_{1 \le j < i \le n} |\lambda(R_{ii}) - \lambda(R_{jj})| \right) e^{\left(\sum_{i=k+1}^{m} (y_i^2 - x_i^2) - \sum_{i=1}^{k} \lambda_i^2/2 \right)} \prod_{k+1 \le i \le m} \left(y_i \operatorname{erfc}(y_i \sqrt{2}) \right),$$

where $\lambda(R_{ll})$ *are the real eigenvalues of* A *for* $l \le k$, *while* $\lambda(R_{ll}) = x_l \pm i y_l$ *are the complex eigenvalues of* A *for* $l > k$, *and where* erfc *is the complementary error function* $\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2}\, dt$. *The probability that* $A \in \mathcal{A}_k$ *is*

$$p_{n,k} = \frac{c_{n,k}}{k!l!} \int_{\substack{x_i \in \mathbb{R} \\ y_i \in \mathbb{R}^+ \\ \lambda_i \in \mathbb{R}}} \left(\prod_{1 \le j < i \le n} |\lambda(R_{ii}) - \lambda(R_{jj})| \right) e^{\left(\sum_{i=k+1}^{m} (y_i^2 - x_i^2) - \sum_{i=1}^{k} \lambda_i^2/2 \right)}$$
$$\times \prod_{k+1 \le i \le m} \left(y_i \operatorname{erfc}(y_i \sqrt{2}) \right) d\lambda_1 \cdots d\lambda_k dx_1 \cdots dx_l dy_1 \cdots dy_l.$$

We now restrict ourselves to finding the probability that a random matrix whose entries are iid standard Gaussians has all real eigenvalues and take a close look at the steps involved. Let $\mathcal{A}$ be the set of n-by-n matrices with real entries which have exactly n real and distinct eigenvalues. When all entries of A are iid standard Gaussians, we have that

$$\mathbb{P}(A) = (2\pi)^{-n^2/2} e^{\frac{-\sum_{1 \le i,j \le n} a_{ij}^2}{2}}.$$

Edelman states that the probability that a random matrix with iid Gaussian entries has repeated eigenvalues is zero. Therefore, the probability that A has all real eigenvalues is given by $\int_{\mathcal{A}} \mathbb{P}(A)\, dA$, which equals

$$\int_{\mathcal{A}} (2\pi)^{-n^2/2} e^{\frac{-\sum_{1 \le i,j \le n} a_{ij}^2}{2}}\, dA. \tag{4.3}$$

Edelman also states that the real Schur decomposition of an n-by-n real matrix A with iid Gaussian entries can be made unique with probability one if some restrictions are imposed on R and Q, namely that the elements on the diagonal of R are ordered $\lambda_1 > \cdots > \lambda_k, x_{k+1} > \cdots > x_m$ and that the first row of Q is chosen to be positive. For $A \in \mathcal{A}$, letting $A = QRQ^T$ be decomposed into a real Schur decomposition with these constraints satisfied, this change of variables gives

$$\begin{aligned}
(2\pi)^{-n^2/2}e^{\frac{-\sum_{1\le i,j\le n} a_{ij}^2}{2}} &= (2\pi)^{-n^2/2}e^{-\operatorname{tr}(AA^T)/2} \\
&= (2\pi)^{-n^2/2}e^{-\operatorname{tr}((QRQ^T)(QR^TQ^T))/2} \\
&= (2\pi)^{-n^2/2}e^{-\operatorname{tr}((QRR^TQ))/2} \\
&= (2\pi)^{-n^2/2}e^{\frac{-\sum_{1\le i,j\le n} r_{ij}^2}{2}} \\
&= (2\pi)^{-n^2/2}e^{-\sum_{1\le i<j\le n} r_{ij}^2/2}e^{-\sum_{j=1}^{n}\lambda_j^2/2},
\end{aligned}$$

where the r_{ij} are the strictly upper triangular entries of R and the λ_j are the eigenvalues of A. Proposition 4.4.1 shows that the Jacobian of this change of variables is given by

$$\prod_{1\le i<j\le n} |\lambda_i - \lambda_j|(d\tilde{R})(Q^T dQ)d\lambda_1 \dots d\lambda_n.$$

Therefore, (4.3) equals

$$\int_{\tilde{\mathcal{A}}} (2\pi)^{-n^2/2}e^{-\sum_{i<j} r_{ij}^2/2}e^{-\sum_{j=1}^{n}\lambda_j^2/2} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|(d\tilde{R})(Q^T dQ)d\lambda_1 \dots d\lambda_n, \tag{4.4}$$

where $\tilde{\mathcal{A}}$ is the image of the region $\mathcal{A}$ under the map which changes variables according to the real Schur decomposition. By the above restrictions on the variables in $\tilde{R}, Q$ and $\lambda_1, \dots, \lambda_n$ which ensure a generically unique real Schur decomposition, Edelman shows that the correct region of integration in this case is over the orthogonal matrices with positive first row, real $\lambda_1 > \cdots > \lambda_n$ and $r_{ij} \in \mathbb{R}$ for $1 \le i < j \le n$. Integrating over these constraints gives the exact probability that A has all real eigenvalues, which is $\frac{1}{2^{n(n-1)/4}}$.

The integrand in (4.4) has an extremely nice form; the eigenvalues $\lambda_1, \dots, \lambda_n$ completely separate from the other variables of integration and integrating r_{ij} over the real line is the same as integration over a standard Gaussian random variable. This allows Edelman to easily integrate over $(d\tilde{R})(Q^T dQ)$ to obtain the joint density of the eigenvalues. Unfortunately, this separation of the eigenvalues is only possible because of the structure of the joint probability density for A when the entries are all independent Gaussians. Indeed, it was the property that for any $a, b \in \mathbb{R}$, $e^{a+b} = e^a e^b$ that allowed the separation of the eigenvalues from the other variables to occur. For uniformly distributed entries, it is not as clear how to proceed. It does not seem likely that we will be able to simply separate out the eigenvalues as in the Gaussian case. Instead, we will carefully consider the region of integration.

4.5 Uniqueness of the real Schur decomposition

We now specify restrictions which make the real Schur decomposition completely unique when a matrix has all real eigenvalues. Edelman states (without proof) in [54] that if the n-by-n matrix A has entries that are iid standard Gaussians, the real Schur decomposition of A can be made unique with probability one by ordering the diagonal entries of the upper triangular matrix R and by assuming that the first row of the orthogonal matrix Q is positive.

We note now that that such a decomposition is not completely unique. For instance, consider

$$\begin{aligned}\begin{pmatrix}1&2&3\\0&4&5\\0&0&6\end{pmatrix} &= \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\0&4&5\\0&0&6\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\\ &= \begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&-2&3\\0&4&-5\\0&0&6\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}.\end{aligned}$$

The problem here is that the first row of the orthogonal matrix is not strictly positive, and cannot be made strictly positive (try!). We address the issue of uniqueness by showing that for matrices with real and distinct eigenvalues, the real Schur decomposition can be made completely unique if one requires that the first non-zero entry in each column of the orthogonal matrix Q be positive.

Let $n \in \mathbb{N}$ and let $M_n(\mathbb{R})$ denote the set of all n-by-n matrices over the field $\mathbb{R}$ and let $M_n(\mathbb{C})$ denote the set of all n-by-n matrices over the field $\mathbb{C}$. We begin with the following lemma, whose purpose is to help prove the uniqueness of our real Schur decomposition.

Lemma 4.5.1. *Suppose that $R, \tilde{R} \in M_n(\mathbb{R})$ are upper triangular with identical diagonal entries*

$r_{11} > r_{22} > \cdots > r_{nn}$. *Let* $S \in M_n(\mathbb{R})$ *be orthogonal. If* $SR = \tilde{R}S$*, then* S *has the form*

$$S = \begin{pmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & \pm 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm 1 \end{pmatrix}.$$

Proof. Let $A_{k,}$ and $A_{,k}$ denote the k-th row and k-th column of a matrix A, respectively. Suppose that $SR = \tilde{R}S$. Multiplying the n-th row of the left-hand matrices by the first column of the right-hand matrices on both sides of the preceding equation, we have that

$$S_{n,}R_{,1} = s_{n1}r_{11} = r_{nn}s_{n1} = \tilde{R}_{n,}S_{,1}.$$

Since $r_{11} > r_{nn}$, this implies that $s_{n1} = 0$. Now multiplying the n-th row of the left-hand matrices by the second column of the right-hand matrices, and using the fact that $s_{n1} = 0$, we have that

$$S_{n,}R_{,2} = s_{n2}r_{22} = r_{nn}s_{n2} = \tilde{R}_{n,}S_{,2},$$

implying that $s_{n2} = 0$ as well. One can proceed in this manner to see that the first $n-1$ entries of the n-th row of S must be zero. Finally, multiplying the n-th row of the left-hand matrices by the n-th column of the right-hand matrices and assuming that $s_{n1} = s_{n2} = \cdots = s_{n-1,n} = 0$, we have that

$$S_{n,}R_{,n} = s_{nn}r_{nn} = r_{nn}s_{nn} = \tilde{R}_{n,}S_{,n}.$$

Since S is orthogonal, we see that $s_{nn} = \pm 1$.

Let our inductive hypothesis be that for some $1 \leq k \leq n-1$, the k rows $S_{n,}, S_{n-1,}, \ldots, S_{n-k+1,}$ of S have been shown to have ± 1 in the diagonal entry, and are zero otherwise. Consider the $(n-k)$-th row of S. By orthogonality requirements, and from taking the dot product of the $(n-k)$-th row of S with every row below it, we see that all entries after the $(n-k)$-th position must be zero. For the first entry in $S_{n-k,}$, we have

$$S_{n-k,}R_{,1} = s_{n-k,1}r_{11} = r_{n-k,n-k}s_{n-k,1} = \tilde{R}_{n-k,}S_{,1},$$

since all entries below $s_{n-k,1}$ in the first column of S have been shown to be zero. Since $r_{11} > r_{nn}$, this implies that $s_{n-k,1} = 0$. Continue in this manner to show that the first $n-k-1$ entries $(n-k)$-th row of S are zero. Then the remaining entry must be ± 1 by orthogonality requirements. This proves that S has the specified form. $\square$

The following theorem is necessary to prove the existence of our real Schur decomposition.

Theorem 4.5.2 (Schur triangularization, Theorem 2.3.1 in [79]). *Let $A \in M_n(\mathbb{R})$ have eigenvalues $\lambda_1, \ldots, \lambda_n$ in any prescribed order and let $x \in \mathbb{C}^n$ be a unit vector such that $Ax = \lambda_1 x$. If A has only real eigenvalues, then x may be chosen to be real and there is a real orthogonal $Q = [x\ q_2\ \ldots\ q_n] \in M_n(\mathbb{R})$ such that $Q^T AQ = R$ is upper triangular with diagonal entries $r_{ii} = \lambda_i, i = 1, \ldots, n$.*

We are now ready to provide a completely unique real Schur decomposition for all matrices with real and distinct eigenvalues.

Proposition 4.5.3. *Suppose that $A \in M_n(\mathbb{R})$ has all real and distinct eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. Then A admits exactly one real Schur decomposition $A = QRQ^T$, such that the first non-zero entry in every column of Q is positive and such that the diagonal entries $r_{ii} = \lambda_i$ for $1 \leq i \leq n$.*

Proof. Let us first show that such a decomposition exists. Theorem 4.5.2 tells us that when the eigenvalues of A are all real, there exists a real orthogonal matrix Q and an upper triangular matrix R with diagonal entries $r_{ii} = \lambda_i$ for each $1 \leq i \leq n$ such that $A = QRQ^T$. Since Q is orthogonal, every column of Q has at least one non-zero entry. Denote the first non-zero entry in the i-th column of Q by q_{*i} for each $1 \leq i \leq n$. Let D be the diagonal matrix where

$$\begin{cases} d_{ii} = 1, & \text{if } q_{*i} > 0 \\ d_{ii} = -1, & \text{if } q_{*i} < 0 \end{cases}.$$

Letting $\tilde{Q} = QD$ and $\tilde{R} = DRD$, we have

$$\begin{aligned} A &= QRQ^T \\ &= QDDRDDQ^T \\ &= \tilde{Q}\tilde{R}\tilde{Q}^T. \end{aligned}$$

Observe that the orthogonal matrix $\tilde{Q}$ corresponds to negating the i-th column of Q whenever $d_{ii} = -1$, and leaving all other entries of Q unchanged. Therefore,

$$\tilde{q}_{*i} = q_{*i} d_{ii} = \begin{cases} q_{*i}, & \text{if } q_{*i} > 0 \\ -q_{*i}, & \text{if } q_{*i} < 0 \end{cases},$$

so $\tilde{q}_{*i} > 0$ for each $1 \leq i \leq n$. The matrix $\tilde{R}$ corresponds to negating the i-th row and then also negating i-th column of R whenever $d_{ii} = -1$; the diagonal entry is therefore negated twice, leaving the eigenvalues on the diagonal of R unchanged. This proves existence.

For uniqueness, suppose now that $A = QRQ^T$ and $A = \tilde{Q}\tilde{R}\tilde{Q}^T$ are two different real Schur decompositions of A, such that the upper triangular matrices R and $\tilde{R}$ have equal diagonals given by $r_{ii} = \tilde{r}_{ii} = \lambda_i$ for $1 \leq i \leq n$, and such that the first non-zero entry in each column of Q and $\tilde{Q}$ is positive. Then

$$QRQ^T = \tilde{Q}\tilde{R}\tilde{Q}^T$$

implies that

$$\tilde{Q}^T QR = \tilde{R}\tilde{Q}^T Q.$$

Let $M = \tilde{Q}^T Q$. Since M is an orthogonal matrix, Lemma 4.5.1 says that M must be a diagonal matrix where each entry is ± 1. From the equality

$$Q = \tilde{Q}M,$$

we see that if $m_{ii} = 1$, the i-th column of Q equals the i-th column of $\tilde{Q}$, and if $m_{ii} = -1$, the i-th column of Q is the negative of the i-th column of $\tilde{Q}$. In particular, this implies that Q and $\tilde{Q}$ have zero entries in exactly the same positions. Moreover, we see that the coordinates of the first non-zero entry in a given column of $\tilde{Q}$ are the same as the coordinates of the first non-zero entry in the corresponding column of Q. For the i-th column, these entries are called $q_{*,i}$ and $\tilde{q}_{*,i}$, and by hypothesis, both are positive. Since $q_{*,i} = \tilde{Q}_{*,}M_{,i} = \tilde{q}_{*,i}m_{ii}$, this implies that $m_{ii} = 1$ for every $1 \leq i \leq n$, so M is the identity matrix and uniqueness holds. □

The proof of the uniqueness of the real Schur decomposition provided by Proposition 4.5.3 relies on the assumption that the eigenvalues of the matrix A are distinct. The purpose of the following definition and lemma is to show that this assumption holds with probability one for a random matrix whose entries are absolutely continuous with respect to the Lebesgue measure.

Definition 4.5.4. Define the **power sum symmetric polynomials** in the n variables $x_1, \ldots, x_n$ as

$$p_j(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^j$$

for $j \geq 0$.

Lemma 4.5.5. *Let A be an n-by-n random matrix whose entries are independent and absolutely continuous with respect to the Lebesgue measure. The probability that A has repeated eigenvalues is zero.*

Proof. Let the eigenvalues of A be given by $\lambda_1, \ldots, \lambda_n$, which are complex-valued random variables. Then the product

$$\prod_{i \neq j} (\lambda_i - \lambda_j)^2 \tag{4.5}$$

is a symmetric polynomial in terms of the eigenvalues of A, which equals zero if and only if $\lambda_i = \lambda_j$ for some $i \neq j$. By the discussion in Chapter 1, Section 1.2 in [136], we see that the symmetric polynomial given by (4.5) can be expressed as a polynomial with rational coefficients in the power sum symmetric polynomials $p_1(\lambda_1, \ldots, \lambda_n), \ldots, p_n(\lambda_1, \ldots, \lambda_n)$ via Newton's identities (see page 8 of [136]). Since

$$\text{tr}(A^k) = \lambda_1^k + \cdots + \lambda_n^k$$

is precisely the k-th power sum symmetric polynomial in the variables $\lambda_1, \ldots, \lambda_n$, we may write

$$\prod_{i \neq j} (\lambda_i - \lambda_j)^2 = Q(\operatorname{tr}(A), \operatorname{tr}(A^2), \ldots, \operatorname{tr}(A^n))$$

for some polynomial Q, which is not identically zero. To see that Q is not identically zero, note that the equality must hold for all matrices A and consider a diagonal matrix where all entries are unique.

From the trace identity

$$\operatorname{tr}(A^k) = \sum_{1 \leq i_1, \ldots, i_k \leq n} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1},$$

we see that the polynomial Q can be written as a polynomial in the entries of A. Thus,

$$\prod_{i \neq j} (\lambda_i - \lambda_j)^2 = \tilde{Q}(a_{11}, a_{12}, \ldots, a_{nn}),$$

where $\tilde{Q}$ is a polynomial that is also not identically zero, because

$$\tilde{Q}(a_{11}, a_{12}, \ldots, a_{nn}) = Q(\operatorname{tr}(A), \operatorname{tr}(A^2), \ldots, \operatorname{tr}(A^n)).$$

Since the entries of A are independent, by Lemma 3.2.2, the probability that $\tilde{Q}(a_{11}, a_{12}, \ldots, a_{nn}) = 0$ is zero, so the probability that $\prod_{i \neq j} (\lambda_i - \lambda_j)^2 = 0$ is also zero. Hence, the probability that A has a repeated eigenvalue is zero. $\square$

4.6 Probability of all real eigenvalues for 2-by-2 generalized bounded height model ensemble random matrices

We now apply Edelman's method to 2-by-2 matrices from the generalized bounded height model ensemble, giving a proof of Theorem 4.1.1. The small matrix size allows us to explicitly work out the algebra.

Proof of Theorem 4.1.1. Let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for each $1 \leq i, j \leq n$. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

with jointly independent entries $a_{ij} \sim U(\alpha_{ij}, \beta_{ij})$, for $1 \leq i, j \leq n$. Let $\mathcal{A}$ be the set of 2-by-2 matrices with entries a_{ij} in $[\alpha_{ij}, \beta_{ij}]$ for each $1 \leq i, j \leq 2$ which have exactly two real and distinct eigenvalues. By Lemma 4.5.5, the probability that A has repeated eigenvalues is zero. Therefore, the probability that A has all real eigenvalues is given by the probability that $A \in \mathcal{A}$.

Let $\mathbb{P}(A)\ dA$ denote the joint probability density function for the random matrix A on the space of 2-by-2 real matrices with respect to the Lebesgue measure on $\mathbb{R}^4$. When all entries of A are independent, $\mathbb{P}(A)$ is simply the product of the pdf of each entry of A. In our case,

$$\mathbb{P}(A)\ dA = \left(\prod_{1 \leq i,j \leq 2} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \left(\prod_{1 \leq i,j \leq 2} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) \right)\ dA$$

and the probability that A has all real and distinct eigenvalues is given by

$$\int_{A \in \mathcal{A}} \left(\prod_{1 \leq i,j \leq 2} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \left(\prod_{1 \leq i,j \leq 2} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) \right)\ dA. \tag{4.6}$$

For $A \in \mathcal{A}$, letting $A = QRQ^T$ be decomposed into its unique real Schur decomposition given by Proposition 4.5.3, this change of variables gives

$$\prod_{1 \leq i,j \leq 2} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) = \prod_{1 \leq i,j \leq 2} \mathbb{1}_{\{\alpha_{ij} \leq (QRQ^T)_{ij} \leq \beta_{ij}\}}((QRQ^T)_{ij}).$$

This decomposition should be interpreted as a change of variables from the entries $a_{11}, a_{12}, a_{21}, a_{22}$ of A to the four new variables λ_1, λ_2, r, and q of R and Q, as defined in Definition 4.3.1. Here, λ_1 and λ_2 are the real eigenvalues of A, $r \in \mathbb{R}$ is the strictly upper triangular entry of R, and $q \in \mathbb{R}$ is the single parameter which specifies the 2-by-2 orthogonal matrix Q. (We will see momentarily that only one parameter is needed to specify a 2-by-2 orthogonal matrix with a given determinant. Alternatively, one may immediately read this fact off of the chart provided on page 217 of [56].) This decomposition is made unique by specifying that the first non-zero entry in every column of Q is positive and that the eigenvalues of A are listed in decreasing order on the diagonal of R, as according to Proposition 4.5.3.

Proposition 4.4.1 shows that the Jacobian of this change of variables is given by

$$|\lambda_1 - \lambda_2|\ dr dq d\lambda_1 d\lambda_2.$$

Therefore (4.6), and hence also the probability that A has exactly two real eigenvalues, equals

$$\int_{\tilde{\mathcal{A}}} \left(\prod_{1\le i,j\le 2} \frac{1}{(\beta_{ij}-\alpha_{ij})} \right) \left(\prod_{1\le i,j\le 2} \mathbb{1}_{\{\alpha_{ij}\le (QRQ^T)_{ij}\le \beta_{ij}\}}((QRQ^T)_{ij}) \right) |\lambda_1-\lambda_2| \; drdqd\lambda_1 d\lambda_2, \quad (4.7)$$

where $\tilde{\mathcal{A}}$ is the image of the region $\mathcal{A}$ under the map which changes variables according to the real Schur decomposition. In order to find the region $\tilde{\mathcal{A}}$, we must find the restrictions on the variables q, r, λ_1, and λ_2 which ensure a unique Schur decomposition and which count each matrix in $\mathcal{A}$ exactly once.

We now investigate the restrictions required by Proposition 4.5.3 which make the Schur factorization unique for the 2-by-2 case. Any 2-by-2 orthogonal matrix

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

must satisfy the equations $QQ^T = I$ and $Q^T Q = I$, i.e.,

$$\begin{cases} q_{11}^2 + q_{12}^2 = 1 \\ q_{21}^2 + q_{22}^2 = 1 \\ q_{11}q_{21} + q_{12}q_{22} = 0 \end{cases} . \quad (4.8)$$

A quick analysis shows that there are two possible types of matrices which satisfy equations (4.8). Solving the equations and relabeling, we obtain

$$Q = \begin{pmatrix} q & \sqrt{1-q^2} \\ -\sqrt{1-q^2} & q \end{pmatrix}$$

or

$$Q = \begin{pmatrix} q & \sqrt{1-q^2} \\ \sqrt{1-q^2} & -q \end{pmatrix}.$$

Restricting $0 \le q \le 1$ guarantees that the first row of Q is non-negative. If $q = 0$, we include only the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ due to the positivity requirements on the first non-zero entry of every column of Q. Similarly, if $q = 1$, we include only the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Thus, whenever $A \in \mathcal{A}$ has two real and distinct eigenvalues $\lambda_1 > \lambda_2$, Proposition 4.5.3 says the Schur decomposition of A is uniquely specified by

$$R = \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix}$$

and exactly one of the matrices Q described above.

Now, assume that $A \in \mathcal{A}$ has real eigenvalues $\lambda_1 > \lambda_2$ and let $A = QRQ^T$ be decomposed into its unique real Schur decomposition. If

$$Q = \begin{pmatrix} q & \sqrt{1-q^2} \\ -\sqrt{1-q^2} & q \end{pmatrix}$$

for some $0 < q < 1$, equating coefficients, we obtain the relation

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} q & \sqrt{1-q^2} \\ -\sqrt{1-q^2} & q \end{pmatrix} \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} q & -\sqrt{1-q^2} \\ \sqrt{1-q^2} & q \end{pmatrix}$$

which implies that

$$\begin{cases} a_{11} = (\lambda_1 - \lambda_2)q^2 + \lambda_2 + rq\sqrt{1-q^2} \\ a_{12} = -(\lambda_1 - \lambda_2)q\sqrt{1-q^2} + q^2 r \\ a_{21} = -(\lambda_1 - \lambda_2)q\sqrt{1-q^2} - r + q^2 r \\ a_{22} = -(\lambda_1 - \lambda_2)q^2 + \lambda_1 - rq\sqrt{1-q^2} \end{cases} . \tag{4.9}$$

Similarly, if

$$Q = \begin{pmatrix} q & \sqrt{1-q^2} \\ \sqrt{1-q^2} & -q \end{pmatrix}$$

for some $0 < q < 1$, equating coefficients, we obtain the relation

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} q & \sqrt{1-q^2} \\ \sqrt{1-q^2} & -q \end{pmatrix} \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} q & \sqrt{1-q^2} \\ \sqrt{1-q^2} & -q \end{pmatrix}$$

which implies that

$$\begin{cases} a_{11} = (\lambda_1 - \lambda_2)q^2 + \lambda_2 + rq\sqrt{1-q^2} \\ a_{12} = (\lambda_1 - \lambda_2)q\sqrt{1-q^2} - q^2 r \\ a_{21} = (\lambda_1 - \lambda_2)q\sqrt{1-q^2} + r - q^2 r \\ a_{22} = -(\lambda_1 - \lambda_2)q^2 + \lambda_1 - rq\sqrt{1-q^2} \end{cases} . \tag{4.10}$$

Notice the symmetry between the cases: the diagonal entries are equal, and the off-diagonal entries are negatives of each other. This allows us to combine the conditions.

If $q = 0$, we obtain the relation

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which implies that

$$a_{11} = \lambda_2, a_{12} = 0, a_{21} = r, a_{22} = \lambda_1.$$

Since $\mathbb{P}(a_{12} = 0) = 0$, such a decomposition will not contribute to an integral over $A \in \mathcal{A}$, so we exclude the possibility that $q = 0$.

When $q = 1$, we obtain the relation

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which implies that

$$a_{11} = \lambda_1, a_{12} = r, a_{21} = 0, a_{22} = \lambda_2.$$

Similarly, since $\mathbb{P}(a_{21} = 0) = 0$, we exclude the possibility that $q = 1$.

Therefore, the above discussion tells us that when $A \in \mathcal{A}$ has two real eigenvalues $\lambda_1 > \lambda_2$ (and $a_{12} \neq 0, a_{21} \neq 0$, with happens with probability one), there exists a unique point (r, q), with $r \in \mathbb{R}$ and $q \in (0, 1)$, such that the entries $a_{11}, a_{12}, a_{21}, a_{22}$ and the variables λ_1, λ_2, r and q satisfy exactly one set of the equations (4.9) or (4.10). Define the subset of $\mathbb{R}^4$

$$\begin{aligned}\mathcal{B} = \{(\lambda_1, \lambda_2, r, q) \mid \lambda_1 > \lambda_2,\\ &q \in (0,1),\\ &(\lambda_1-\lambda_2)q^2+\lambda_2+rq\sqrt{1-q^2} \in (\alpha_{11},\beta_{11}),\\ &(\lambda_1-\lambda_2)q\sqrt{1-q^2}+q^2r \in (-\beta_{12},-\alpha_{12})\cup(\alpha_{12},\beta_{12}),\\ &(\lambda_1-\lambda_2)q\sqrt{1-q^2}+r-q^2r \in (-\beta_{21},-\alpha_{21})\cup(\alpha_{21},\beta_{21}),\\ &-(\lambda_1-\lambda_2)q^2+\lambda_1-rq\sqrt{1-q^2} \in (\alpha_{22},\beta_{22})\}.\end{aligned}$$

Consider the map which sends a matrix $A \in \mathcal{A}$ (with $a_{12} \neq 0$ and $a_{21} \neq 0$) to its unique Schur decomposition, where the eigenvalues on the diagonal of R are listed in decreasing order and the first non-zero entry in each column of Q is positive. Proposition 4.5.3 shows that this map is injective. To see that this map is surjective onto $\mathcal{B}$, suppose that $(\lambda_1, \lambda_2, r, q) \in \mathcal{B}$. Then letting

$$\begin{aligned}a_{11} &= (\lambda_1-\lambda_2)q^2+\lambda_2+rq\sqrt{1-q^2}\\ a_{12} &= (\lambda_1-\lambda_2)q\sqrt{1-q^2}-q^2r\\ a_{21} &= (\lambda_1-\lambda_2)q\sqrt{1-q^2}+r-q^2r\\ a_{22} &= -(\lambda_1-\lambda_2)q^2+\lambda_1-rq\sqrt{1-q^2}\end{aligned}$$

gives us a matrix $A \in \mathcal{A}$ which maps to that point. Thus, $\mathcal{B}$ is the image of $\mathcal{A}$ under this map, and therefore (4.7) equals

$$\int_{\mathcal{B}} \left(\prod_{1\le i,j\le 2} \frac{1}{(\beta_{ij}-\alpha_{ij})}\right)\left(\prod_{1\le i,j\le 2} \mathbb{1}_{\{\alpha_{ij}\le (QRQ^T)_{ij}\le\beta_{ij}\}}((QRQ^T)_{ij})\right)|\lambda_1-\lambda_2|\; dr dq d\lambda_1 d\lambda_2.$$

This integral can be written as

$$\left(\prod_{1\le i,j\le 2} \frac{1}{(\beta_{ij}-\alpha_{ij})}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \times$$
$$\mathbb{1}_{\{0<q<1\}}(q)\mathbb{1}_{\{\lambda_1>\lambda_2\}}(\lambda_1,\lambda_2)\left(\prod_{1\le i,j\le 2} \mathbb{1}_{\{\alpha_{ij}\le (QRQ^T)_{ij}\le\beta_{ij}\}}((QRQ^T)_{ij})\right)|\lambda_1-\lambda_2|\; dr dq d\lambda_1 d\lambda_2. \tag{4.11}$$

Recalling now that

$$\begin{aligned} A_{\lambda_1,\lambda_2} = \{(r,q) \mid q \in (0,1), \\ &(\lambda_1-\lambda_2)q^2+\lambda_2+rq\sqrt{1-q^2} \in (\alpha_{11},\beta_{11}), \\ &(\lambda_1-\lambda_2)q\sqrt{1-q^2}+q^2r \in (-\beta_{12},-\alpha_{12}) \cup (\alpha_{12},\beta_{12}), \\ &(\lambda_1-\lambda_2)q\sqrt{1-q^2}+r-q^2r \in (-\beta_{21},-\alpha_{21}) \cup (\alpha_{21},\beta_{21}), \\ &-(\lambda_1-\lambda_2)q^2+\lambda_1-rq\sqrt{1-q^2} \in (\alpha_{22},\beta_{22})\}, \end{aligned}$$

we may rewrite (4.11) succinctly as

$$\left(\prod_{1\le i\le j\le 2} \frac{1}{(\beta_{ij}-\alpha_{ij})}\right) \int_{-\infty}^{\infty}\int_{\lambda_2}^{\infty}\int_{A_{\lambda_1,\lambda_2}} |\lambda_1-\lambda_2| \, drdqd\lambda_1 d\lambda_2, \tag{4.12}$$

proving Theorem 4.1.1. □

Let the quantity in (4.12) be denoted by ρ_2. If $\rho_2 \neq 0$, the joint density of the eigenvalues for the 2-by-2 generalized bounded height model random matrices, given that both eigenvalues are real, is given by

$$\rho(\lambda_1,\lambda_2) = \frac{1}{2\rho_2}\left(\prod_{1\le i\le j\le 2} \frac{1}{(\beta_{ij}-\alpha_{ij})}\right) \int_{A_{\lambda_1,\lambda_2}} |\lambda_1-\lambda_2| \, drdq,$$

where ρ_2 is the normalization constant and the factor of $1/2$ removes the ordering of the eigenvalues.

The probability given by (4.12) is very hard to calculate, even in so-called nice cases like when $\alpha_{11} = \alpha_{12} = \alpha_{21} = \alpha_{22} = 0$ and $\beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = 1$, which will be explored in Chapter 5.

4.7 Probability of all real eigenvalues for n-by-n generalized bounded height model random matrices

Again letting $\alpha_{ij} < \beta_{ij}$ for all $1 \le i,j \le n$, we now apply Edelman's method to the n-by-n matrices from the generalized bounded height model ensemble, giving a proof of Theorem 4.1.2.

Proof of Theorem 4.1.2. Let $\mathcal{A}$ be the set of n-by-n matrices with entries a_{ij} in $[\alpha_{ij}, \beta_{ij}]$ for each $1 \leq i, j \leq n$ which have exactly n real and distinct eigenvalues. By Lemma 4.5.5, the probability that A has repeated eigenvalues is zero. Therefore, the probability that A has all real eigenvalues is given by the probability that $A \in \mathcal{A}$.

Let $\mathbb{P}(A)\ dA$ denote joint probability density function for the random matrix A on the space of n-by-n real matrices with respect to the Lebesgue measure on $\mathbb{R}^{n^2}$. Since the entries of the matrix A are independent, we have

$$\mathbb{P}(A)\ dA = \left(\prod_{1 \leq i,j \leq n} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \left(\prod_{1 \leq i,j \leq n} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) \right)\ dA$$

and the probability that A has all real and distinct eigenvalues is given by

$$\int_{A \in \mathcal{A}} \left(\prod_{1 \leq i,j \leq n} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \left(\prod_{1 \leq i,j \leq n} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) \right)\ dA. \tag{4.13}$$

For $A \in \mathcal{A}$, letting QRQ^T be decomposed into its unique real Schur decomposition given by Proposition 4.5.3, this change of variables gives

$$\left(\prod_{1 \leq i,j \leq n} \mathbb{1}_{\{\alpha_{ij} \leq a_{ij} \leq \beta_{ij}\}}(a_{ij}) \right) = \mathbb{1}_{\{\alpha_{ij} \leq (QRQ^T)_{ij} \leq \beta_{ij}\}}((QRQ^T)_{ij}).$$

This decomposition, as defined in Definition 4.3.1, should be interpreted as a change of variables from the n^2 entries $a_{11}, \ldots, a_{nn}$ of A to the n^2 new variables given by the n real eigenvalues $\lambda_1, \ldots, \lambda_n$ of A, the $n(n-1)/2$ real strictly upper triangular entries $r_{12}, \ldots, r_{n-1,n}$ of R, and the $n(n-1)/2$ angular parameters $\theta_{12}, \ldots, \theta_{n-1,n}$, each contained in $[0, 2\pi)$, of the orthogonal matrix Q. (We will see momentarily that this provides a correct count and characterization of the new variables. Alternatively, one may immediately read this fact off of the chart provided on page 217 of [56].) This decomposition is made unique by specifying that the first non-zero entry in every column of Q is positive and that the eigenvalues of A are listed in decreasing order on the diagonal of R, as according to Proposition 4.5.3.

Proposition 4.4.1 shows that the Jacobian of this change of variables is given by

$$\prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|\ (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n,$$

where we recall that $(d\tilde{R})$ is the wedge product over the strictly upper triangular elements of R and $(Q^T dQ)$ denotes the wedge product over the $\frac{1}{2}n(n-1)$ independent elements of the antisymmetric matrix $Q^T dQ$, and $\lambda_1, \ldots, \lambda_n$ are the real eigenvalues of A. Therefore (4.13), and hence also the probability that A has all real eigenvalues, equals

$$\int_{\tilde{\mathcal{A}}} \left(\prod_{1\le i,j\le n} \frac{1}{(\beta_{ij}-\alpha_{ij})} \right) \times \left(\prod_{1\le i,j\le n} \mathbb{1}_{\{\alpha_{ij}\le (QRQ^T)_{ij} \le \beta_{ij}\}}((QRQ^T)_{ij}) \right) \prod_{1\le i<j\le n} |\lambda_i - \lambda_j| \; (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n, \tag{4.14}$$

where $\tilde{\mathcal{A}}$ is the image of the region $\mathcal{A}$ under the map which changes variables according to the real Schur decomposition. In order to find the region $\tilde{\mathcal{A}}$, we must find restrictions on the variables in $\tilde{R}$, Q, and $\lambda_1, \ldots, \lambda_n$, which ensure a unique real Schur decomposition and which count each matrix in $\mathcal{A}$ exactly once.

We now investigate the restrictions required by Proposition 4.5.3 which make the real Schur factorization unique for the the n-by-n case. We begin by considering any n-by-n orthogonal matrix Q given by

$$Q = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix}.$$

For each $1 \le j \le n$, let q_{*j} denote the first non-zero entry in the j-th column of Q. Due to positivity and orthogonality requirements, the entries of Q must satisfy the equations

$$\begin{cases} q_{*j} > 0 & 1 \le j \le n \\ \sum_{j=1}^{n} q_{ij}^2 = 1 & 1 \le i \le n \\ \sum_{1\le k\le n} q_{ik}q_{jk} = 0 & 1 \le i \ne j \le n. \end{cases}$$

In [135], Raffenetti and Ruedenberg give a recurrence relation which constructs a general n-by-n orthogonal matrix, which we will refer to as the **Raffenetti and Ruedenberg algorithm**. For $1 \le i \ne j \le n$, they define the matrix A_{ij} as the matrix with $\cos(\theta_{ij})$ as the diagonal entries in the

ith and jth columns, with $\sin(\theta_{ij})$ as the ij entry and with $-\sin(\theta_{ij})$ as the ji entry. Here, each θ_{ij} is an angular parameter contained in $[0, 2\pi)$. All other diagonal entries of the matrix are 1, and all other off-diagonal entries of the matrix are 0. They then give the sequence

$$\begin{aligned}
Q &= Q^{(n)} \\
Q^{(i)} &= A^{(i)} t^{(i)}, \quad i = 2, 3, \ldots, n \\
t^{(i)} &= \left(\begin{array}{c|c} Q^{(i-1)} & 0 \\ \hline 0 & 1 \end{array}\right), \quad i = 2, 3, \ldots, n \\
A^{(i)} &= A_{i-1,i} \cdots A_{2,i} A_{1,i}, \quad i = 2, 3, \ldots, n \\
Q^{(1)} &= \pm 1
\end{aligned}$$

where letting $Q^{(1)} = 1$ gives n-by-n orthogonal matrix with determinant 1 and letting $Q^{(1)} = -1$ gives n-by-n orthogonal matrix with determinant -1. We read off that this orthogonal matrix is given by the $\frac{1}{2}n(n-1)$ independent angular parameters, each contained in $[0, 2\pi)$,

$$\Theta = \{\theta_{12}, \theta_{13}, \ldots, \theta_{1n}, \theta_{23}, \ldots \theta_{2n}, \ldots, \theta_{n-1,n}\}.$$

This construction gives only an orthogonal matrix with the specified size and determinant, and does not make specific entries positive; we will address this by specifying the appropriate region of integration.

Let Q be parameterized as according to the algorithm, and let

$$R = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ & \ddots & \ddots & \vdots \\ & & \ddots & r_{n-1,n} \\ & & & \lambda_n \end{pmatrix}.$$

Let

$$\Theta = \{\theta_{12}, \theta_{13}, \ldots, \theta_{1n}, \theta_{23}, \ldots \theta_{2n}, \ldots, \theta_{n-1,n}\},$$

where each angular parameter in Θ is contained in $[0, 2\pi)$, and let

$$\rho = \{r_{12}, r_{13}, \ldots, r_{1n}, r_{23}, \ldots r_{2n}, \ldots, r_{n-1,n}\}.$$

Therefore, for any $A \in \mathcal{A}$ with real and distinct eigenvalues $\lambda_1 > \cdots > \lambda_n$, the uniqueness of the real Schur decomposition given in Proposition 4.5.3 tells us that there exists exactly one point $(r_{12}, r_{13}, \ldots, r_{1n}, r_{23}, \ldots r_{2n}, \ldots, r_{n-1,n}, \theta_{12}, \theta_{13}, \ldots, \theta_{1n}, \theta_{23}, \ldots \theta_{2n}, \ldots, \theta_{n-1,n}) \in (\rho, \Theta)$ satisfying the equations $Q_{*j} > 0$ for $1 \leq j \leq n$, $Q^T Q = I_n$, and $a_{ij} = (QRQ^T)_{ij}$ for $1 \leq i, j \leq n$. Define the subset of $\mathbb{R}^{n^2}$

$$\begin{aligned} \mathcal{B} = \{(\lambda_1, \ldots, \lambda_n, \rho, \Theta) \mid \lambda_1 > \cdots > \lambda_n & \\ & Q_{*j} > 0 \text{ for } 1 \leq j \leq n, \\ & Q^T Q = I_n, \\ & \alpha_{ij} < (QRQ^T)_{ij} < \beta_{ij} \text{ for } 1 \leq i, j \leq n\}. \end{aligned} \tag{4.15}$$

Consider the map which sends a matrix $A \in \mathcal{A}$ to its unique Schur decomposition, where the eigenvalues on the diagonal of R are listed in decreasing order and the first non-zero entry in each column of Q is positive. Proposition 4.5.3 shows that this map is injective. To see that this map is surjective onto $\mathcal{B}$, suppose that $(\lambda_1, \ldots, \lambda_n, \rho, \Theta) \in \mathcal{B}$. Letting $a_{ij} = (QRQ^T)_{ij}$ for each $1 \leq i, j \leq n$ gives us a matrix $A \in \mathcal{A}$ which maps to that point. Thus, $\mathcal{B}$ is the image of $\mathcal{A}$ under this map, and therefore (4.14) equals

$$\int_{\mathcal{B}} \left(\prod_{1 \leq i,j \leq n} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \times$$
$$\left(\prod_{1 \leq i,j \leq n} \mathbb{1}_{\{\alpha_{ij} \leq (QRQ^T)_{ij} \leq \beta_{ij}\}}((QRQ^T)_{ij}) \right) \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j| \; (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n.$$

This integral can be written as

$$\begin{aligned} &\left(\prod_{1 \leq i,j \leq n} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{1 \leq j \leq n} \mathbb{1}_{\{Q_{*j} > 0\}}(\theta_{12}, \ldots, \theta_{n-1,n}) \mathbb{1}_{\{Q^T Q = I_n\}}(\theta_{12}, \ldots, \theta_{n-1,n}) \\ &\times \mathbb{1}_{\{\lambda_1 > \cdots > \lambda_n\}}(\lambda_1, \ldots, \lambda_n) \left(\prod_{1 \leq i,j \leq 2} \mathbb{1}_{\{\alpha_{ij} \leq (QRQ^T)_{ij} \leq \beta_{ij}\}}((QRQ^T)_{ij}) \right) \\ &\times \prod_{1 \leq i,j \leq n} |\lambda_i - \lambda_j| \; (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n. \end{aligned} \tag{4.16}$$

For any real $\lambda_1, \ldots, \lambda_n$, recalling now that

$$\begin{aligned} A_{\lambda_1,\ldots,\lambda_n} = \{(\rho, \Theta) \mid\ & Q_{*j} > 0 \text{ for } 1 \le j \le n, \\ & Q^T Q = I_n, \\ & \alpha_{ij} < (QRQ^T)_{ij} < \beta_{ij} \text{ for } 1 \le i, j \le n\}, \end{aligned}$$

we may rewrite (4.16) as

$$\left(\prod_{1 \le i \le j \le n} \frac{1}{(\beta_{ij} - \alpha_{ij})} \right) \int_{-\infty}^{\infty} \int_{\lambda_{n-1}}^{\infty} \cdots \int_{\lambda_2}^{\infty} \int_{A_{\lambda_1,\ldots,\lambda_n}} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j| \ (d\tilde{R})(Q^T dQ) d\lambda_1 \ldots d\lambda_n,$$

proving Theorem 4.1.2. □

Chapter 5

Eigenvalues and fluctuations of perturbations of large box model random matrices

5.1 Introduction and main results

This chapter begins with a section which presents an example showing that, under mild assumptions on the iid matrix entries, the expected characteristic polynomial of a matrix A has all real roots. Does this mean that we should expect A to have all real eigenvalues with high probability? Unfortunately not. In fact, Figure 5.1 provides an illustration of the eigenvalues of integer matrices. We can see that most of the eigenvalues appear to not be real.

After introducing some preliminary results and tools in Section 5.3, the rest of the chapter investigates the eigenvalues of generalized large box model random matrices.

In Section 5.4, we let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for each $1 \leq i, j \leq n$ and consider the random matrices whose entries are uniformly distributed on $[\alpha_{ij}N, \beta_{ij}N] \cap \mathbb{Z}$ and jointly independent, as $N \to \infty$. We will see that when scaled appropriately, the empirical spectral measure of these matrices converges weakly in distribution to the empirical spectral measure of the random matrices whose entries are uniformly distributed on the continuous interval $[\alpha_{ij}, \beta_{ij}]$ for $1 \leq i, j \leq n$ and jointly independent. This gives a kind of equivalence between these types of matrices, allowing us to work with the continuously distributed entries when necessary. This result is contained in the following main theorem.

Theorem 5.1.1. *For each $1 \leq i, j \leq n$, let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$. Let A be an n-by-n random matrix with entries that are uniformly distributed on $[\alpha_{ij}N, \beta_{ij}N] \cap \mathbb{Z}$ for $1 \leq i, j \leq n$ and jointly independent.*

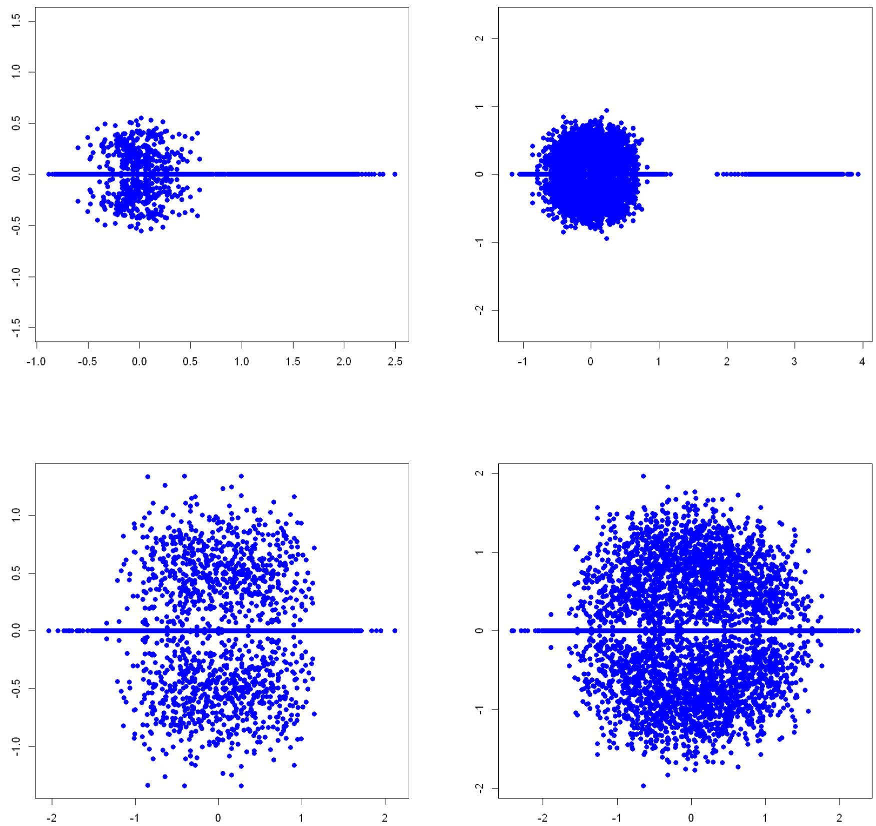

Figure 5.1: The top left figure shows the eigenvalues of 1000 3-by-3 random matrices with entries uniformly distributed on $[0, 10^{15}] \cap \mathbb{Z}$ and scaled by 10^{15}. The top right figure shows the eigenvalues of 1000 6-by-6 random matrices with entries uniformly distributed on $[0, 10^{15}] \cap \mathbb{Z}$, scaled by 10^{15}. The bottom left figure shows the eigenvalues of 1000 3-by-3 random matrices with entries uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$ and scaled by 10^{15}. The bottom right figure shows the eigenvalues of 1000 6-by-6 random matrices with entries uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$, scaled by 10^{15}.

Let B be an n-by-n random matrix with entries that are uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$ for $1 \leq i, j \leq n$ and jointly independent. As $N \to \infty$, for $0 \leq k \leq n$, the probability that A has exactly k real eigenvalues converges to the probability that B has exactly k real eigenvalues. Moreover, the empirical spectral measure of A/N converges weakly in distribution to the empirical spectral measure of B. Finally, the joint distribution of the eigenvalues of A/N converges in distribution to the joint distribution of the eigenvalues of B.

Theorem 5.1.1 is illustrated in Figure 5.2.

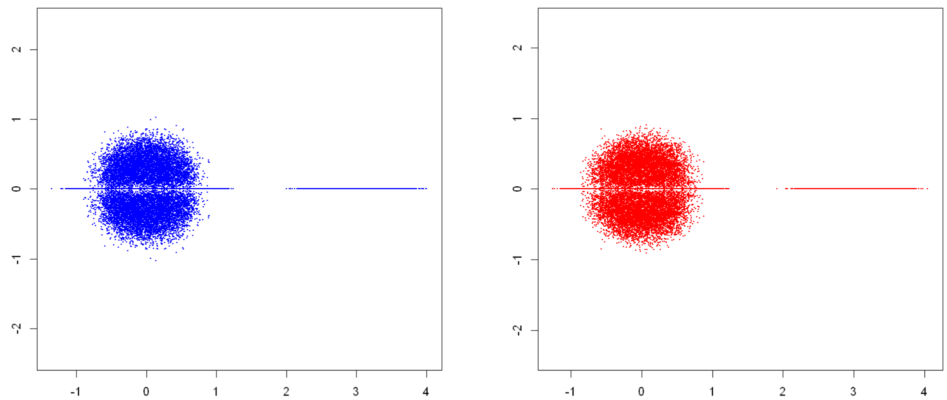

Figure 5.2: These images demonstrate Theorem 5.1.1. The left image shows the eigenvalues of 5000 6-by-6 random matrices with entries independently and uniformly distributed on $[0, 10^{15}] \cap \mathbb{Z}$ and scaled by 10^{15}. The right image shows the eigenvalues of 5000 6-by-6 random matrices with entries independently and uniformly distributed on $(0, 1)$.

In Section 5.5, we consider what happens to the eigenvalues of an n-by-n random matrix A with entries independently and uniformly distributed on the integers $[-N, N]$ when perturbed by another matrix P whose entries are all μ_N, as $N \to \infty$. We consider the three cases where $\mu_N/N \to \infty$, $\mu_N/N \to 0$, and $\mu_N/N \to c$, for $c \in \mathbb{R}$. In general, we will see that if we add very large perturbations, the eigenvalues of $\frac{A+P}{\mu_N}$ converge almost surely to the eigenvalues of the n-by-n all ones matrix J_n. If we add very small perturbations, the empirical spectral measure of $\frac{A+P}{N}$ converges weakly almost surely to the empirical spectral measure the n-by-n matrix whose entries are iid and uniformly distributed on $[-1, 1]$. If we add a perturbation of just the right size, we see that the resulting spectral measure has exactly one outlier. We have the following three main results.

Theorem 5.1.2. *Let A be an n-by-n random matrix whose entries are independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let P be the perturbation matrix whose entries are all $\mu_N \in \mathbb{R}$, with $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$. Normalizing the eigenvalues of $A+P$ by μ_N, we have that as $N \to \infty$, the largest eigenvalue (in magnitude) of $(A+P)/\mu_N$ is real with probability tending to one and converges to n almost surely. As $N \to \infty$, all other eigenvalues of $(A+P)/\mu_N$ converge to 0 almost surely.*

Theorem 5.1.3. *Let A be an n-by-n random matrix whose entries are independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let P be the perturbation matrix whose entries are all μ_N, with $\lim_{N\to\infty} \frac{\mu_N}{N} = 0$. Then the empirical spectral measures of A/N and $(A+P)/N$ both converge weakly in distribution to the empirical spectral measure of the matrix whose entries are iid and uniformly distributed on $[-1, 1]$. Moreover, the joint distribution of eigenvalues of A/N and $(A + P)/N$ both converge in distribution to the joint distribution of eigenvalues of the matrix whose entries are iid and uniformly distributed on $[-1, 1]$.*

Theorem 5.1.3 is illustrated in Figure 5.3.

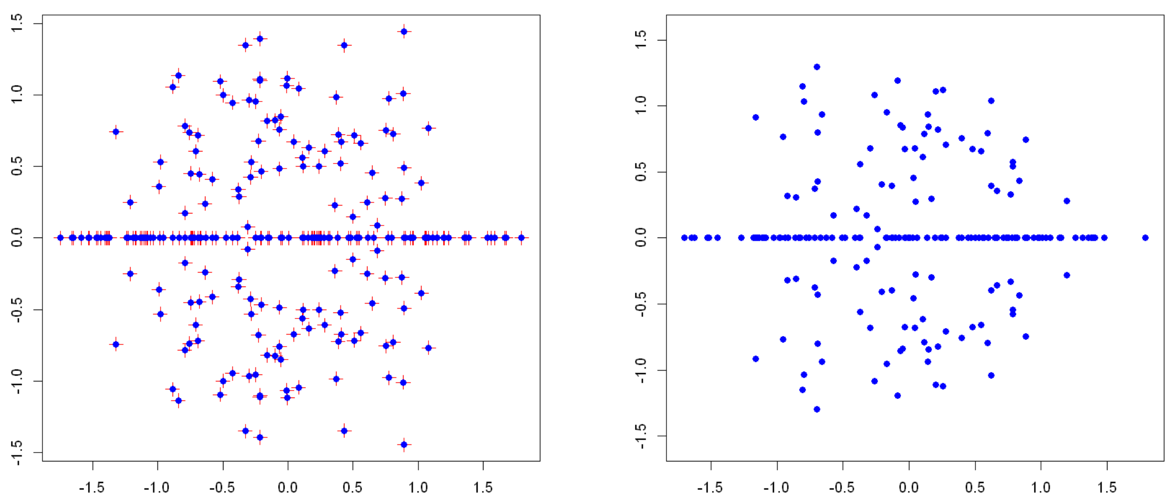

Figure 5.3: The image on the left illustrates Theorem 5.1.3. The blue dots represent the eigenvalues of fifty 4-by-4 random matrices with entries independently and uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$, and scaled by 10^{15}. The red crosses represent the eigenvalues of the same fifty matrices, but with every entry perturbed by $\sqrt{10^{15}}$ and also scaled by 10^{15} For reference, the blue dots in the image on the right represent the eigenvalues of fifty 4-by-4 random matrices with entries independently and uniformly distributed on $[-1, 1]$.

Theorem 5.1.4. *Let A be an n-by-n random matrix whose entries are independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let $c > 0$, and let B be the n-by-n random matrix whose entries are independently and uniformly distributed on $[-1 + c, 1 + c]$. Let P be the perturbation matrix whose* *entries are all* μ_N*, with* $\lim_{N\to\infty}\frac{\mu_N}{N}=c$*. Then as* $N\to\infty$*, the empirical spectral measure of* $(A+P)/N$ *converges weakly in distribution to the empirical spectral measure of* B*.*

Remark 5.1.5. When $\lim_{N\to\infty}\frac{\mu_N}{N}=-\infty$ or when $\lim_{N\to\infty}\frac{\mu_N}{N}=c$ with $c<0$, the theorems above hold with minor adjustments.

In each of Theorems 5.1.2, 5.1.3, and 5.1.4, the matrix P whose entries are all μ_N can be expressed as $P=\mu_N\mathbb{1}\mathbb{1}^T$, where $\mathbb{1}^T=(1,1,\dots,1)$ is a row vector of length n. Letting J_n be the n-by-n matrix of all ones, we can also express this as $P=\mu_N J_n$. Observe that the rank, or number of independent rows, of P is one. Adding P to the matrix A (which has iid entries uniformly distributed on $[-N,N]\cap\mathbb{Z}$ and jointly independent) is therefore an example of an additive rank one perturbation of A.

In 2013, Tao [153] studies outlier eigenvalues created by low rank additive perturbations to certain iid matrices. Tao reminds the readers of the circular law, which he uses along with a finite fourth moment assumption on the matrix entries and a truncation argument to quickly prove asymptotic results regarding the spectral radius and operator norm of the normalized (unperturbed) iid matrix. Essentially, these results show that as the matrix size tends to infinity, the iid random matrices considered by Tao do not have eigenvalues outside of the unit circle with high probability. This is stated precisely in the following theorem.

Theorem 5.1.6 (Theorems 1.3 and 1.4 in [153])**.** *Let* X_n *be an* n*-by-*n *iid random matrix, with complex entries that are normalized to have mean zero and variance one. Then as* $n\to\infty$*,* $\mu_{\frac{1}{\sqrt{n}}X_n}$ *converges almost surely to the circular measure* μ_c*, where* $d\mu_c=\frac{1}{\pi}\mathbb{1}_{|z|\le 1}dz$*. Additionally, if the entries of* X_n *have finite fourth moment, then as* $n\to\infty$*,* $\rho\left(\frac{1}{\sqrt{n}}X_n\right)$ *converges to* 1 *almost surely and for any finite* $m\ge 1$*,* $\|(\frac{1}{\sqrt{n}}X_n)^m\|_{op}$ *converges to* $m+1$ *almost surely.*

Now, let X_n be an n-by-n iid random matrix, with complex entries that have mean zero, variance one, and finite fourth moment. With the asymptotic behavior of the eigenvalues of $\frac{1}{\sqrt{n}}X_n$ established, Tao considers an n-by-n deterministic perturbation matrix P_n, which has both rank and operator norm $O(1)$. He shows that if P_n has eigenvalues which are sufficiently far enough away from the unit disk, then as the matrix size tends to infinity, the matrix $\frac{1}{\sqrt{n}}X_n + P_n$ has the same number of outlier eigenvalues as P_n, and they are within distance $o(1)$ of the corresponding eigenvalues of P_n. The exact statement of this theorem was given earlier, as Theorem 1.3.13.

Theorem 5.1.2 shows that when $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, the largest eigenvalue of $(A+P)/\mu_N = A/\mu_N + J_n$ tends to n, which is the largest eigenvalue of the matrix of J_n, almost surely as $N \to \infty$. Observe that the eigenvalues of J_n are n (with multiplicity one) and 0 (with multiplicity $n-1$), and that J_n has rank one and operator norm n. The hypothesis of our theorem is quite different from Tao's; our atom distribution is very specific and is not normalized to have unit variance, and the dimension n of our matrices stay fixed while Tao takes n to infinity. But interestingly, although Tao's theorem is not directly applicable to our situation, it still accurately predicts the asymptotic location of the outlier eigenvalue of $A/\mu_N + J_n$.

While Tao's result is much more general in the sense that it holds for perturbations of higher rank and for broader classes of atom distributions, our result is useful because it holds for every matrix dimension n. Theorem 5.1.2 is proven by applying Rouché's theorem to appropriately chosen circles around the roots of the characteristic polynomial of $(A+P)/\mu_N$ and the roots of the characteristic polynomial of the all ones matrix. Tao proves Theorem 1.7 by using the following eigenvalue criterion.

Theorem 5.1.7 (Eigenvalue criterion, Lemma 2.1 in [153])**.** *Let z be a complex number that is not an eigenvalue of $\frac{1}{\sqrt{n}}X_n$. Then z is an eigenvalue of $\frac{1}{\sqrt{n}}X_n + P_n$ if and only if*

$$\det\left(I + B_n\left(\frac{1}{\sqrt{n}}X_n - z\right)^{-1} C_n\right) = 0,$$

where $P_n = C_n B_n$.

After applying the eigenvalue criterion, he applies Rouché's theorem to the functions

$$f(z) = \det\left(I + B_n\left(\frac{1}{\sqrt{n}}X_n - z\right)^{-1} C_n\right)$$

and

$$g(z) = \det\left(I + B_n\left(-z\right)^{-1} C_n\right)$$

to conclude the proof.

While this machinery was not necessary for the proof of Theorem 5.1.2, it was precisely these techniques that made it possible for us to investigate the fluctuations of the largest eigenvalues of rank one perturbations of large box model random matrices, discussed in Section 5.6, and lead to a proof of the following theorem.

Theorem 5.1.8. *Let A be an n-by-n random matrix with entries independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let P be the perturbation matrix whose entries are all μ_N, with $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$. Let $\lambda_{\max}$ denote the largest, in magnitude, eigenvalue of $A + P$. Then as $N \to \infty$,*

$$\frac{\lambda_{\max} - n\mu_N}{nN} \overset{d}{\Longrightarrow} \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} X_{ij}}{n^2},$$

where X_{ij} are independently and uniformly distributed on $(-1, 1)$ for $1 \leq i, j \leq n$.

Here, when $\alpha < \beta \in \mathbb{R}$ and $X_i \sim U(\alpha, \beta)$ for $1 \leq i \leq n$ are jointly independent, the continuous probability distribution defined by

$$X = \frac{1}{n} \sum_{i=1}^{n} X_i$$

is known as the **Bates distribution** [12]. In other words, Theorem 5.1.8 shows that when $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, the centered and scaled largest eigenvalue (in magnitude) of $A+P$ converges in distribution to the Bates distribution with n^2 parameters.

Rajagopalan [137] expanded on Tao's work in [153] by examining the fluctuations of the outlier eigenvalues that result from additive perturbations. Under assumptions similar to Tao's, Rajagopalan let X_n be an n-by-n iid complex random matrix whose entries have mean zero and variance one. Let P_n be a perturbation matrix with rank and operator norm bounded by $O(1)$. For a fixed $\epsilon > 0$, Rajagopalan also assumes between $4 + \epsilon$ and $8 + \epsilon$ bounded moments for the atom distribution of X_n, depending on the eigenvectors of the perturbation matrix. Let $\lambda_{\max}$ denote the largest (in magnitude) eigenvalue of $X_n/\sqrt{n} + P_n$.

In the special case where the perturbation matrix is given by

$$P_N = \frac{\mu_N}{N} \left(\frac{1}{\sqrt{n}} \mathbb{1} \right) \left(\frac{1}{\sqrt{n}} \mathbb{1} \right)^T,$$

where $|\mu_N/N| > 1$ and $4+\epsilon$ bounded moments for X_{11} are assumed, Rajagopalan shows as a discussion to his main theorem that as $n \to \infty$, the normalized fluctuation $\sqrt{n}(\lambda_{\max} - \mu_N/N)$ converges in distribution to a complex Gaussian g with mean zero, $\mathbb{E}g^2 = \frac{|\mu_N/N|^2 \mathbb{E}X_{11}^2}{|\mu_N/N|^2 - \mathbb{E}X_{11}^2}$, and $\mathbb{E}|g|^2 = \frac{|\mu_N/N|^2}{|\mu_N/N|^2-1}$.

As before with our comparison to Tao, the entries of our matrix X_n do not have variance one, and we are not considering the limit as the matrix dimension tends to infinity. On the contrary, our result for finite dimension n in this case looks quite different than Rajagopalan's for limiting n: Theorem 5.1.8 says when the random matrix A is drawn from the large box model and $P = \mu_N \mathbb{1}\mathbb{1}^T$ with $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, the fluctuation of the largest eigenvalue $\lambda_{\max}$ of $A+P$ given by $\frac{1}{n}\left(\lambda_{\max}/N - n\mu_N/N\right)$ converges in distribution $\frac{\sum_{j=1}^n \sum_{i=1}^n X_{ij}}{n^2}$, where each $X_{ij} \sim U(-1,1)$ is iid and jointly independent, as $N \to \infty$. Of course, we note that as $n \to \infty$, the central limit theorem (see for instance Theorem 3.4.1 in [48]) says that when the iid random variables X_{ij} are uniformly distributed on $[-\sqrt{3}, \sqrt{3}]$ (so that they are normalized to have variance one) and jointly independent, then $\frac{\sum_{i=1}^n \sum_{j=1}^n X_{ij}}{n}$ converges in distribution to a standard Gaussian.

Of course, Rajagopalan's main theorem is much more general than ours, able to handle perturbations of rank $O(1)$, rather than just rank one. However, his theorem only applies as $n \to \infty$, and it is interesting to see that for large box model random matrices and for finite n, the limiting distribution is in terms a Bates distribution, rather than a Gaussian distribution.

In another special case of his main theorem, Rajagopalan provides another example for rank one perturbations. Let $|\theta| > 1$ and let P_n be the perturbation matrix whose $(1,1)$ entry is θ and all other entries are zero. Assume $8+\epsilon$ bounded moments for the atom distribution X_{11}. Observe that θ is the only non-zero eigenvalue of P_n. Let $\lambda_{\max}$ be the largest eigenvalue of $X_n/\sqrt{n} + P_n$. Rajagopalan finds that as $n \to \infty$, the centered and properly normalized outlier fluctuations converge in distribution to $X_{11} + g$, where g is a complex Gaussian with $\mathbb{E}(g) = 0, \mathbb{E}(g^2) = 0$, and where $\mathbb{E}(|g|^2)$ has the distribution of complex Gaussian random variable with mean zero and variance $\frac{1}{|\theta|^2-1}$. While we did not consider perturbations of this kind, we suspect it would be possible to apply the proof methods of this thesis to this case as well, yielding more information about eigenvalue fluctuations for finite n.

The proof of Rajagopalan's main result relies on the Lidskii-Vishik-Lyusternik theorem, which he states in Appendix A. This theorem provides an algorithm for finding the spectrum and fluctuations of the eigenvalues of an n-by-n deterministic matrix when perturbed by a sequence of perturbation matrices whose entries are all bounded by $o(1)$.

Another very recent finding regarding eigenvalue fluctuations is one by Banerjee, Mukherjee, and Pal [10], who considered the largest eigenvalue of symmetric random matrices X_n where the entries are standard Gaussians which also satisfy $\sup_{(i,j)\neq(i',j')} \left|\mathbb{E}[X_{ij}X_{i'j'}]\right| = O(n^{-(1+\epsilon)})$, for any $\epsilon > 0$. As the matrix size tends to infinity, they prove that the largest eigenvalue of $X_n/\sqrt{n}$ converges to 2 almost surely by considering the formula

$$\mathrm{tr}[(n^{-1/2}X_n)^k] = \frac{1}{n^{k/2}} \sum_{i_1,i_2,\ldots,i_k} X_{i_1 i_2} \ldots X_{i_k,i_1},$$

and giving a combinatorial argument which counts these powers. They then use Wick's formula which calculates $\mathbb{E}(X_1X_2\cdots X_k)$ to bound $\mathbb{E}(\mathrm{tr}[n^{-1/2}X_n)^k])$, and then apply Markov's inequality to bound the probability that the largest eigenvalue of $X_n/\sqrt{n}$ is greater than $2+\nu$, for some $\nu > 0$. Summing over the contribution from the trace powers, they complete the proof by an application of the Borel-Cantelli lemma.

Recalling that $\sup_{(i,j)\neq(i',j')} \left|\mathbb{E}[X_{ij}X_{i'j'}]\right| = O(n^{-(1+\epsilon)})$, when $\epsilon \geq 1$, Banerjee, Mukherjee, and Pal also consider perturbations of the form $P_n = \frac{\lambda}{\sqrt{n}}\mathbb{1}\mathbb{1}^T$, where $\lambda \gg n^{1/4}$. In this case, they find that the properly normalized largest eigenvalue fluctuation converges in distribution to $\sqrt{2}Z$, where $Z \sim N(0,1)$. They accomplish this by first finding another representation of the largest eigenvalue, and then by showing that as $n \to \infty$, $\frac{1}{n}\sum_{1\leq i,j\leq n} X_{ij}$ has a Gaussian distribution with variance 2. Although we did not consider symmetric matrices with Gaussian entries or let our matrix dimension n tend to infinity, it is interesting to us that in yet another case, the normalized largest eigenvalue converges to an integer almost surely as $n \to \infty$, and that the properly normalized fluctuations tend to a Gaussian random variable, as in the result of Rajagopalan.

5.2 Expectation of characteristic polynomial for matrix with iid entries

We begin with an example that suggests it may be reasonable to assume that a random matrix with particularly nice entries has a high probability, or even probability one, of having all real eigenvalues.

Let I_n denote the n-by-n identity matrix. Recall that the characteristic polynomial of a matrix $A \in M_n(\mathbb{C})$ is given by

$$p_A(t) = \det(tI_n - A).$$

Consider a random matrix $A \in M_n(\mathbb{R})$ with independent and identically distributed entries. We show that the deterministic polynomial whose coefficients are given by the expected value of the coefficients of the characteristic polynomial of A has all real roots.

We begin by introducing principal minors, the sums of which will comprise the coefficients of the characteristic polynomials. The following definition is due to Horn [79].

Definition 5.2.1 (See Section 0.7.1 in [79])**.** Let $A \in M_n(\mathbb{R})$. For index sets $\alpha \subseteq \{1, \ldots, n\}$ and $\beta \subseteq \{1, \ldots, n\}$ we denote by $A[\alpha, \beta]$ the submatrix of entries that lie in the rows of A indexed by α and the columns indexed by β. If $\alpha = \beta$, $A[\alpha, \alpha]$ is a **principal submatrix** of A. The determinant of a principal submatrix of A is a **principal minor** of A.

Let $E_k(A)$ denote the sum of the principal minors of A of size k. Since a principal submatrix of size k is formed by keeping the same k rows and k columns of the original matrix A, and there are exactly $\binom{n}{k}$ possible ways to choose the rows (and hence also the columns), there are $\binom{n}{k}$ principal minors of A of size k.

The characteristic polynomial of A can be written as

$$p_A(t) = t^n - E_1(A)t^{n-1} + \cdots + (-1)^{n-1}E_{n-1}(A)t + (-1)^n E_n(A), \tag{5.1}$$

for instance, see (1.2.13) in [79].

Theorem 5.2.2. *Suppose that for $1 \leq i, j \leq n$, the entries a_{ij} of $A \in M_n(\mathbb{R})$ are independently and identically distributed with $\mathbb{E}|a_{11}| < \infty$. Suppose further that the mean of each entry is $\mu \in \mathbb{R}$.*

Then

$$\begin{aligned}\mathbb{E}(p_A(t)) :&= \mathbb{E}(\det(tI - A))\\ &= t^n - \mathbb{E}(E_1(A))t^{n-1} + \cdots + (-1)^{n-1}\mathbb{E}(E_{n-1}(A))t + (-1)^n\mathbb{E}(E_n(A))\\ &= t^n - n\mu t^{n-1}.\end{aligned}$$

In particular, the deterministic polynomial $\mathbb{E}(p_A(t))$ *has all real roots.*

Proof. Observe that $E_1(A)$ is the sum of all principal minors of size one of A. This is the sum of all entries that lie in the k-th row and k-th column of A, for each $1 \leq k \leq n$. This is simply the trace of A. We have

$$\begin{aligned}\mathbb{E}(\mathrm{tr}(A)) &= \mathbb{E}\left(\sum_{i=1}^{n} a_{ii}\right)\\ &= \sum_{i=1}^{n}\mathbb{E}(a_{ii})\\ &= n\mu\end{aligned}$$

by the linearity of expectation and since each entry of A has mean μ.

Now consider the expected value of $E_k(A)$, for each $1 < k \leq n$. For $k = 2$, we know that $E_2(A)$ is the sum of the $\binom{n}{2}$ determinants of the 2-by-2 principal submatrices of A. The determinant of each 2-by-2 submatrix of A (and hence also the principal minors of A of size 2) can be written as

$$\det\begin{pmatrix} a_{ij} & a_{ik}\\ a_{lj} & a_{lk}\end{pmatrix} = a_{ij}a_{lk} - a_{lj}a_{ik},$$

for some $1 \leq i < l \leq n$ and $1 \leq j < k \leq n$. By linearity of expectation, the independence of the entries of A, and the fact that the identically distributed entries satisfy $\mathbb{E}|a_{11}| < \infty$, we compute by Theorem 2.1.13 in [48] that

$$\begin{aligned}\mathbb{E}\left(a_{ij}a_{lk} - a_{lj}a_{ik}\right) &= \mathbb{E}(a_{ij})\mathbb{E}(a_{lk}) - \mathbb{E}(a_{lj})\mathbb{E}(a_{ik})\\ &= \mu^2 - \mu^2\\ &= 0.\end{aligned}$$

This shows that the expected value of the determinant of each 2-by-2 submatrix of A is zero. In particular, the expected value of each of the $\binom{n}{2}$ principal minors in the sum of $E_2(A)$ is zero. From this and the linearity of expectation, we conclude that $\mathbb{E}(E_2(A)) = 0$ as well.

We now proceed by induction to show that the remaining coefficients of $\mathbb{E}(p_A(t))$ are zero. Assume $k \geq 2$, and let our inductive hypothesis be that the expected value of the determinant of every k-by-k submatrix of A is zero. In particular, this means that the $\binom{n}{k}$ summands that comprise $E_k(A)$ have expectation zero, implying that $\mathbb{E}(E_k(A)) = 0$ for k. Now consider $\mathbb{E}(E_{k+1}(A))$. Every principal minor in the sum of $E_{k+1}(A)$ is a determinant of a $(k+1)$-by-$(k+1)$ submatrix of A. For the determinant of any $(k+1)$-by-$(k+1)$ submatrix of A, consider the cofactor expansion along its first row. Each cofactor in this determinant is multiplied by the determinant of a k-by-k submatrix of A, which has an expectation of zero by our inductive hypothesis. By Theorem 2.1.13 in [48], the expectation of the product of the cofactor and the k-by-k determinant also has an expected value of zero. Since this is true for every cofactor appearing in the first row, the expectation of this summand is zero. Hence, by the principle of mathematical induction, the expected value of the determinant of every $(k+1)$-by-$(k+1)$ submatrix of A if zero. In particular, the expectation of $E_{k+1}(A)$ is also zero.

This shows that

$$\mathbb{E}(p_A(t)) = t^n - n\mu t^{n-1}.$$

Factoring this polynomial into

$$t^{n-1}(t - n\mu),$$

we see that it has a root at $n\mu$ with multiplicity one, and a root at 0 with multiplicity $n-1$. These roots, and hence the eigenvalues of A, are clearly all real. □

Consider now the special case where each entry of the n-by-n matrix A is independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Then for every fixed N,

$$\mathbb{E}(p_A(t)) = t^n,$$

so therefore

$$\lim_{N\to\infty} \mathbb{E}(p_A(t)) = t^n$$

as well, perhaps suggesting a high number of real eigenvalues. Unfortunately, we will see in our investigation that it is highly unusual for a matrix, even with iid integral entries, to have all real eigenvalues.

5.3 Definitions and preliminary results

This section contains all definitions, tools, and lemmas needed to prove the main theorems.

Definition 5.3.1 (Chapter 5, Section 5.6 in [79]). A function $\|\cdot\| : M_n(\mathbb{C}) \to \mathbb{R}$ is a **matrix norm** if, for all $A, B \in M_n(\mathbb{C})$, it satisfies the following five axioms:

(1) $\|A\| \geq 0$

(2) $\|A\| = 0$ if and only if $A = 0$

(3) $\|cA\| = |c|\|A\|$ for all $c \in \mathbb{C}$

(4) $\|A + B\| \leq \|A\| + \|B\|$

(5) $\|AB\| \leq \|A\|\|B\|$.

Not all authors require a matrix norm to satisfy the submultiplicativity condition above, but we have decided to adopt Horn's convention. We will need the following two definitions, both from [79], for bounds in our proofs.

Definition 5.3.2. The **maximum magnitude of matrix entries** $\|\cdot\|_{\max}$ is defined on $M_n(\mathbb{C})$ by

$$\|A\|_{\max} = \max_{1\leq i,j\leq n} |a_{ij}|.$$

We do not consider it a matrix norm because it is not submultiplicative.

Definition 5.3.3. The **maximum column sum matrix norm** $\|\cdot\|_1$ is defined on $M_n(\mathbb{C})$ by

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|.$$

Example 5.6.4 in [79] shows that this is indeed submultiplicative, and therefore also a matrix norm.

The following lemma is used to bound the differences of the coefficients of the characteristic polynomials of two matrices of the same size. It is similar to Lemma 12 in [127], which establishes a slightly different bound for $|\det(R) - \det(S)|$, where R an S are n-by-n matrices.

Lemma 5.3.4. *Let R and S be n-by-n matrices. Then for any $1 \leq k \leq n$,*

$$|E_k(R) - E_k(S)| \leq \frac{n!}{(n-k)!}\|R-S\|_{\max}\left(\|R\|_{\max} + \|S\|_{\max}\right)^{k-1}.$$

Proof. We will proceed by induction. For $k = 1$, we have that

$$E_1(R) = \operatorname{tr}(R) = R_{11} + \cdots + R_{nn}$$

and

$$E_1(S) = \operatorname{tr}(S) = S_{11} + \cdots + S_{nn}.$$

From the triangle inequality, it follows that

$$\begin{aligned}|E_1(R) - E_1(S)| &= |R_{11} - S_{11} + \cdots + R_{nn} - S_{nn}| \\ &\leq \sum_{i=1}^{n} |R_{ii} - S_{ii}| \\ &\leq n\|R-S\|_{\max}.\end{aligned}$$

We will also show all of the details for the $k = 2$ case. This will help the reader understand the origin of the formula in the general case. Observe that $E_2(R)$ is the sum of $\binom{n}{2}$ principal minors of size two. Each one of these principal minors has the form

$$(R_{i_1j_1}R_{i_2j_2} - R_{i_1j_2}R_{i_2j_1})$$

for some $1 \leq i_1 < i_2 \leq n$ and $1 \leq j_1 < j_2 \leq n$. From here, we see that every summand of $|E_2(R) - E_2(S)|$ has the form

$$|(R_{i_1j_1}R_{i_2j_2} - S_{i_1j_1}S_{i_2j_2}) + (S_{i_1j_2}S_{i_2j_1} - R_{i_1j_2}R_{i_2j_1})|$$

which we may bound above with the triangle inequality by

$$|R_{i_1j_1}R_{i_2j_2} - S_{i_1j_1}S_{i_2j_2}| + |S_{i_1j_2}S_{i_2j_1} - R_{i_1j_2}R_{i_2j_1}|. \tag{5.2}$$

Consider the first summand in (5.2). We add and subtract the same term to obtain the bound

$$\begin{aligned}
&|R_{i_1j_1}R_{i_2j_2} - S_{i_1j_1}S_{i_2j_2}| \\
=\ &|R_{i_1j_1}R_{i_2j_2} - R_{i_1j_1}S_{i_2j_2} + R_{i_1j_1}S_{i_2j_2} - S_{i_1j_1}S_{i_2j_2}| \\
\leq\ &|R_{i_1j_1}\left(R_{i_2j_2} - S_{i_2j_2}\right)| + |\left(R_{i_1j_1} - S_{i_1j_1}\right)S_{i_2j_2}| \\
\leq\ &\|R\|_{\max}\|R - S\|_{\max} + \|R - S\|_{\max}\|S\|_{\max}.
\end{aligned}$$

By symmetry, this shows that an upper bound for (5.2) is

$$2\|R - S\|_{\max}\left(\|R\|_{\max} + \|S\|_{\max}\right).$$

Since there are $\binom{n}{2}$ total terms in $|E_2(R) - E_2(S)|$ that have the form of (5.2), altogether we have the bound

$$|E_2(R) - E_2(S)| \leq \frac{n!}{(n-2)!}\|R - S\|_{\max}\left(\|R\|_{\max} + \|S\|_{\max}\right).$$

Now consider $|E_k(R) - E_k(S)|$ for $2 \leq k < n$. We know that each principal minor of size k of R is a determinant of a k-by-k submatrix of R. Using sigma notation where S_k is the symmetric group of k elements and σ is a permutation in S_k, and relabeling the rows and columns that appear in the principal submatrix to go from 1 to k, we can express each of the $\binom{n}{k}$ terms of $|E_k(R) - E_k(S)|$ as

$$\begin{aligned}
&\left|\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) R_{1\sigma(1)} \cdots R_{k\sigma(k)} - \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) S_{1\sigma(1)} \cdots S_{k\sigma(k)}\right| \\
\leq &\sum_{\sigma \in S_k} \left|R_{1\sigma(1)} \cdots R_{k\sigma(k)} - S_{1\sigma(1)} \cdots S_{k\sigma(k)}\right|.
\end{aligned}$$

Now consider only a single permutation, which we call σ, of the above expansion. Let our inductive hypothesis be that for $\sigma \in S_{k-1}$ with the appropriate relabeling,

$$\left|R_{1\sigma(1)} \cdots R_{k-1\sigma(k-1)} - S_{1\sigma(1)} \cdots S_{k-1\sigma(k-1)}\right| \leq \|R-S\|_{\max} \left(\|R\|_{\max} + \|S\|_{\max}\right)^{k-2},$$

which trivially holds when $k = 2$ and has been shown to hold for the base case when $k = 3$. Then for $\sigma \in S_k$, by adding and subtracting the same term and applying the triangle inequality, we have that

$$\begin{aligned}
&\left|R_{1\sigma(1)} \cdots R_{k\sigma(k)} - S_{1\sigma(1)} \cdots S_{k\sigma(k)}\right| \\
\leq &\left|R_{1\sigma(1)} \cdots R_{k\sigma(k)} - R_{1\sigma(1)} \cdots R_{k-1\sigma(k-1)} S_{k\sigma(k)}\right| \\
&\quad + \left|R_{1\sigma(1)} \cdots R_{k-1\sigma(k-1)} S_{k\sigma(k)} - S_{1\sigma(1)} \cdots S_{k\sigma(k)}\right| \\
= &\left|\left(R_{1\sigma(1)} \cdots R_{k-1\sigma(k-1)}\right)\left(R_{k\sigma(k)} - S_{k\sigma(k)}\right)\right| \\
&\quad + \left|\left(R_{1\sigma(1)} \cdots R_{k-1\sigma(k-1)} - S_{1\sigma(1)} \cdots S_{k-1\sigma(k-1)}\right) S_{k\sigma(k)}\right| \\
&\leq \|R\|_{\max}^{k-1} \|R-S\|_{\max} + \|S\|_{\max} \|R-S\|_{\max} (\|R\|_{\max} + \|S\|_{\max})^{k-2} \\
&= \|R-S\|_{\max} \left(\|R\|_{\max}^{k-1} + \|S\|_{\max} (\|R\|_{\max} + \|S\|_{\max})^{k-2}\right) \\
&\leq \|R-S\|_{\max} \left(\|R\|_{\max} + \|S\|_{\max}\right)^{k-1}
\end{aligned}$$

where we used the inductive hypothesis to bound the differences of the products of $k-1$ terms and the binomial theorem to obtain the last line. Since there are $k!$ permutations in S_k and $\binom{n}{k}$ such terms to consider for $|E_k(R) - E_k(S)|$, we obtain the final bound

$$|E_k(R) - E_k(S)| \leq \frac{n!}{(n-k)!} \|R-S\|_{\max} \left(\|R\|_{\max} + \|S\|_{\max}\right)^{k-1}.$$

□

Lemma 5.3.5. *Let $\alpha_{ij} < \beta_{ij} \in \mathbb{R}$ for each $1 \leq i, j \leq n$. Let B be an n-by-n random matrix with entries b_{ij} that are independently and uniformly distributed on $[\alpha_{ij}, \beta_{ij}]$ for each $1 \leq i, j \leq n$. Let A be the random matrix where for each $1 \leq i, j \leq n$ and $1 \leq l \leq \lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1$,*

$$a_{ij} = \begin{cases} \lceil \alpha_{ij} N \rceil + (l-1), & \text{if } b_{ij} \in \left[\alpha_{ij} + \frac{(\beta_{ij}-\alpha_{ij})(l-1)}{\lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1}, \alpha_{ij} + \frac{(\beta_{ij}-\alpha_{ij})l}{\lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1}\right) \\ \lfloor \beta_{ij} N \rfloor + 1, & \text{if } b_{ij} = \beta_{ij} \end{cases}. \tag{5.3}$$

Then

$$|E_k(A/N) - E_k(B)| = O(1/N)$$

almost surely for each $1 \le k \le n$.

Proof. We will apply Lemma 5.3.4 with $R = A/N$ and $S = B$. For the coupled entries a_{ij} and b_{ij}, we begin by bounding $|a_{ij}/N - b_{ij}|$.

If $b_{ij} = \beta_{ij}$, we have that

$$|a_{ij}/N - b_{ij}| = \left| \frac{\lfloor \beta_{ij} N \rfloor + 1}{N} - \beta_{ij} \right| \le \left| \frac{\beta_{ij} N + 1}{N} - \beta_{ij} \right| \le \frac{1}{N} = O(1/N).$$

If $b_{ij} \ne \beta_{ij}$, we examine the distance of a_{ij}/N from both endpoints of the interval containing b_{ij}. A similar computation shows that for any $1 \le l \le \lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1$,

$$\left| \frac{\lceil \alpha_{ij} N \rceil + (l-1)}{N} - \left(\alpha_{ij} + \frac{(\beta_{ij} - \alpha_{ij})(l-1)}{\lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1} \right) \right| = O(1/N)$$

and

$$\left| \frac{\lceil \alpha_{ij} N \rceil + (l-1)}{N} - \left(\alpha_{ij} + \frac{(\beta_{ij} - \alpha_{ij}) l}{\lfloor \beta_{ij} N \rfloor - \lceil \alpha_{ij} N \rceil + 1} \right) \right| = O(1/N).$$

Altogether, we have

$$|a_{ij}/N - b_{ij}| = O(1/N)$$

as well. Since all of the entries of A/N and B have been coupled in this manner, it follows that

$$\|A/N - B\|_{\max} = O(1/N) \tag{5.4}$$

almost surely, where the implicit constant depends on α_{ij}, β_{ij}, for $1 \le i, j \le n$. Furthermore,

$$\|A/N\|_{\max} \le \max_{1 \le i,j \le n} \{|\alpha_{ij}|, |\beta_{ij}|\}$$

and

$$\|B\|_{\max} \le \max_{1 \le i,j \le n} \{|\alpha_{ij}|, |\beta_{ij}|\}.$$

Finally, applying Lemma 5.3.4 gives the bound

$$\begin{aligned} |E_k(A/N) - E_k(B)| &\leq \frac{n!}{(n-k)!}\|A/N - B\|_{\max}(||A/N||_{\max} + ||B||_{\max})^{k-1} \\ &\leq \frac{n!}{(n-k)!}\|A/N - B\|_{\max}\left(2\max_{1\leq i,j\leq n}\{|\alpha_{ij}|, |\beta_{ij}|\}\right)^{k-1} \\ &= O(1/N) \end{aligned}$$

almost surely. □

Theorem 5.3.6 (Continuity Theorem, Theorem 1.3.1 in [136]). *Let*

$$f(z) = \sum_{i=0}^{n} a_i z^i = \prod_{j=1}^{k}(z - z_j)^{m_j} \quad (m_1 + \cdots + m_k = n)$$

be a monic polynomial of degree n *with distinct zeros* $z_1, \ldots, z_k$ *of multiplicities* $m_1, \ldots, m_k$*. Then, given a positive* $\epsilon < \min_{1\leq i<j\leq k}|z_i - z_j|/2$*, there exists a* $\delta > 0$ *so that any monic polynomial* $g(z) = \sum_{i=0}^{n} b_i z^i$ *whose coefficients satisfy* $|b_i - a_i| < \delta$*, for* $i = 1, \ldots, n-1$*, has exactly* m_j *zeros in the disc of radius* ϵ *centered at* z_j *for each* $1 \leq j \leq k$*.*

Theorem 5.3.7 (Theorem 29.4 in [22]). *For random vectors* $X_n = (X_{n1}, \ldots, X_{nk})$ *and* $Y = (Y_1, \ldots, Y_k)$*, a necessary and sufficient condition for* $X_n \overset{d}{\Rightarrow} Y$ *is that* $\sum_{u=1}^{k} t_u X_{nu} \overset{d}{\Rightarrow} \sum_{u=1}^{k} t_u Y_u$ *for each* $(t_1, \ldots, t_k)$ *in* $\mathbb{R}^k$*.*

Definition 5.3.8 (Definition 1.2.9 in [79]). Let $A \in M_n(\mathbb{C})$. The **spectral radius** of A is $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$, where $\sigma(A)$ denotes the set of all eigenvalues of A.

Definition 5.3.9 (Definition 5.6.6 in [79]). Recall that the singular values of $A \in M_n(\mathbb{C})$ are the square roots of the eigenvalues of A^*A, indexed in decreasing order, $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A)$. The **spectral norm** $\|\cdot\|_2$ is defined on $M_n(\mathbb{C})$ by

$$\|A\|_2 = \sigma_1(A), \text{ the largest singular value of } A.$$

Theorem 5.3.10 (Theorem 8.1.22 in [79]). *Let* $A \in M_n(\mathbb{C})$ *be nonnegative, i.e.,* $a_{ij} \geq 0$ *for every* $1 \leq i, j, \leq n$*. Then*

$$\min_{1\leq i\leq n}\sum_{j=1}^{n} A_{ij} \leq \rho(A) \leq \max_{1\leq i\leq n}\sum_{j=1}^{n} A_{ij}$$

and

$$\min_{1\leq j\leq n}\sum_{i=1}^{n} A_{ij} \leq \rho(A) \leq \max_{1\leq j\leq n}\sum_{i=1}^{n} A_{ij}$$

We will need the following theorem to show that the eigenvalues of our coupled matrices can be made arbitrarily close.

Theorem 5.3.11 (Gelfand's formula, Corollary 5.6.14 in [79])**.** *Let* $||\cdot||$ *be a matrix norm on* $M_n(\mathbb{C})$ *and let* $A \in M_n(\mathbb{C})$*. Then* $\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$.

The following lemma shows that a matrix whose entries are independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$ and scaled by N has all of its eigenvalues contained in the circle centered at the origin of radius n. In this case, we say that the matrix has no outlier eigenvalues.

Lemma 5.3.12. *Let* $A \in M_n(\mathbb{R})$*, with every entry independently and uniformly distributed on* $[-N, N] \cap \mathbb{Z}$*. Then*

$$\rho\left(A/N\right) \leq n$$

with probability one.

Proof. Let A be the matrix where every entry is independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Consider the maximum column sum matrix norm on A/N,

$$\|A/N\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^{n} |a_{ij}|/N,$$

From the submultiplicativity of this matrix norm, it follows that $\|(A/N)^k\|_1 \leq \|A/N\|_1^k$ for every $k \in \mathbb{N}$.

Well,

$$\begin{aligned}\|A/N\|_1 &= \max_{1\leq j\leq n}\sum_{i=1}^{n} |a_{ij}|/N \\ &\leq \max_{1\leq j\leq n}\sum_{i=1}^{n} N/N \\ &= n.\end{aligned}$$

Therefore, by Theorem 5.3.11,

$$\begin{aligned}\rho(A) &= \lim_{k\to\infty} \|(A/N)^k\|_1^{1/k} \\ &\le \lim_{k\to\infty} (\|A/N\|_1^k)^{1/k} \\ &\le \|A/N\|_1 \\ &\le n.\end{aligned}$$

□

The following theorem from Brailovskaya and van Handel [28] and the fact that Theorem 5.3.11 implies that the spectral radius is bounded above by the spectral norm gives us a bound on the probability of outlier eigenvalues for matrices whose entries are iid and uniformly distributed on $[-1,1]$.

Theorem 5.3.13 (Theorem 1.4 in [28])**.** *Let X be an n-by-m matrix with $X_{ij} = b_{ij}\xi_{ij}$, where $b_{ij} \ge 0$ are arbitrary scalars and ξ_{ij} are independent symmetrically distributed real random variables with $\mathbb{E}(\xi_{ij}^{2p}) \le \mathbb{E}(g^{2p})$ for all i, j, and $p \in \mathbb{N}$ (here $g \sim N(0,1)$). Define*

$$\nu_1^2 = \max_{j\le m} \sum_{i\le n} b_{ij}^2, \quad \nu_2^2 = \max_{i\le n} \sum_{j\le m} b_{ij}^2, \quad \nu_*^2 = \max_{\substack{i\le n\\ j\le m}} b_{ij}^2.$$

Suppose that $\nu_1 \le \nu_2$. Then

$$\mathbb{P}\left(\|X\|_2 > \nu_1 + \nu_2 + \nu_*^{4/3}\nu_1^{-1/3}t\right) \le \frac{n\nu_*^2}{C\nu_1^2} e^{-Ct^{3/2}}$$

for all $0 \le t \le \frac{\nu_1^{1/3}\nu_2}{\nu_*^{4/3}}$, where C is a universal constant.

Let us apply this theorem to the matrices with iid entries that are uniformly distributed on $[-1,1]$. In this case, we have that $\nu_1^2 = \nu_2^2 = n$ and $\nu_*^2 = 1$. Then Theorem 5.3.13 says that

$$\mathbb{P}\left(\|X\|_2 > 2\sqrt{n} + tn^{-1/6}\right) \le \frac{e^{-Ct^{3/2}}}{C}$$

for all $0 \le t \le n^{2/3}$. In particular, letting $t = n^{2/3}$, this yields

$$\mathbb{P}\left(\|X\|_2 > 3\sqrt{n}\right) \le \frac{e^{-Cn}}{C}.$$

While we do not know the precise value of the universal constant C, we see that outlier eigenvalues should be unlikely, at least for n sufficiently large. Since we are working with finite n, we will use the spectral radius bound of n given by Lemma 5.3.12, which holds with probability one. When the dimension n of an iid random matrix X_n with entries that have mean zero, variance one, and bounded fourth moment tends to infinity, Theorem 1.4 in [153] shows that the spectral radius of $X_n/\sqrt{n}$ tends to 1 almost surely. From simulations, it seems like the spectral radius of our matrices should be closer to $\sqrt{n}$ with high probability, though we were unable to prove this. A bound of this order would allow us to make more precise statements about outlier eigenvalues.

5.4 Equivalence of generalized large box model and generalized bounded height model random matrices

We are ready to prove Theorem 5.1.1, which shows that the eigenvalues of the generalized large box model random matrices and the generalized bounded height model random matrices behave similarly, allowing us to work with the simplified continuously distributed matrix entries. As with the polynomial model, Rouché's theorem is the key ingredient of our proof.

Proof of Theorem 5.1.1. Begin by coupling the entries of the random matrices A and B with the coupling described in Lemma 5.3.5; Lemma 3.3.2 and Remark 3.3.3 show that this is indeed a coupling of the entries of A and B. We refer to the coupled matrices as A' and B'.

If x is an eigenvector of A with eigenvalue λ, then

$$\left(\frac{A}{N}\right)x = \frac{1}{N}(Ax) = \frac{\lambda}{N}x,$$

showing that λ/N is an eigenvalue of A/N. This shows that the eigenvalues of A/N are precisely the eigenvalues of A scaled by a factor of N. In particular, A and A/N have the same number of real eigenvalues. Therefore, it suffices to prove the first part of the theorem for A/N and B. Our approach will be to use Theorem 2.6.5 on the characteristic polynomials of the coupled matrices A'/N and B', and then use that $A \sim A'$ and $B \sim B'$ to deduce the desired result.

Suppose that the entries of the matrix B' are fixed so that B' does not have any repeated eigenvalues, which happens with probability one by Lemma 4.5.5. Fixing the entries of B' also fixes the entries of A' due to the assumed coupling. Let $\lambda_1', \dots, \lambda_n'$ denote the eigenvalues of B'; order the eigenvalues first by increasing magnitude and then by increasing argument in $[0, 2\pi)$. Note that these eigenvalues are exactly the roots of the characteristic polynomial of B'. For any $\epsilon > 0$ and $1 \le j \le n$, let $C_{\lambda_j', \epsilon}$ denote the circle of radius ϵ centered at λ_j'. Since the eigenvalues of B' are all distinct, there exists an $r > 0$ sufficiently small so that the circles $C_{\lambda_1', r}, \dots, C_{\lambda_n', r}$ are all disjoint. Since the entries of B' are all real, the coefficients of the characteristic polynomial of B' are also all real, implying that any complex eigenvalues of B' must occur in pairs. Therefore, if λ_j' is a complex eigenvalue of B', the circle $C_{\lambda_j', r}$ is contained entirely in the upper or lower half-plane; otherwise, $C_{\lambda_j', r}$ would intersect the circle of radius r centered at the complex conjugate of λ_j', a contradiction to the circles being disjoint.

Let $0 < \epsilon < r/2$ be the radius of the circles centered at the eigenvalues of B'. Let f be the characteristic polynomial of A'/N and let g be the characteristic polynomial of B'. For every $1 \le j \le n$ and every $z \in C_{\lambda_j', \epsilon}$

$$|g(z)| = |z - \lambda_1'| \cdots |z - \lambda_n'| \ge \epsilon r^{n-1} > \epsilon^n \tag{5.5}$$

by the construction of the circles.

By (5.1), the triangle inequality, and Lemma 5.3.5, and Lemma 5.3.12, for every $1 \le j \le n$ and every $z \in C_{\lambda_j', \epsilon}$, we may bound

$$\begin{aligned} |f(z) - g(z)| &\le |z^n - z^n| + \sum_{k=1}^{n} |z|^{n-k} \left| E_k(A'/N) - E_k(B') \right| \\ &\le \sum_{k=1}^{n} (|\lambda_j'| + \epsilon)^{n-k} \left| E_k(A'/N) - E_k(B') \right| \\ &= O(1/N) \end{aligned}$$

almost surely. This means that we can choose N sufficiently large so that

$$|f(z) - g(z)| < \epsilon^n < |g(z)|,$$

for all $z \in C_{\lambda'_j,\epsilon}$. By Theorem 2.6.5, A'/N and B' (and hence A' and B') have the same number of eigenvalues in $C_{\lambda'_j,\epsilon}$ for every $1 \leq j \leq n$.

This implies that the coupled matrices A' and B' have exactly the same number of real eigenvalues. Indeed, for every real eigenvalue λ'_j of B', we know from Rouché's theorem that the circle $C_{\lambda'_j,\epsilon}$ contains exactly one eigenvalue of A'. This eigenvalue must be real; otherwise $C_{\lambda'_j,\epsilon}$, whose center is on the real axis, would capture the eigenvalue of A' and its complex conjugate, a contradiction to the number of eigenvalues in $C_{\lambda'_j,\epsilon}$. On the other hand, if λ'_j is a complex eigenvalue of B' and $\bar{\lambda'}_j$ is its complex conjugate, by assumption ϵ is sufficiently small so that $C_{\lambda'_j,\epsilon} \cap C_{\bar{\lambda'}_j,\epsilon} = \emptyset$. This means that the circles cannot intersect the real axis, so the corresponding eigenvalue of A' in each circle must also be complex. Hence, for N sufficiently large, A' has the same number of real eigenvalues as B'. Since the entries of B' were fixed to have no repeated eigenvalues, which happens with probability one, this result holds for almost every realization of B'. Hence, as $N \to \infty$, the probability that A' has exactly k real eigenvalues converges to the probability that B' has exactly k real eigenvalues. Since $A \sim A'$ and $B \sim B'$, this implies that as $N \to \infty$, the probability that A has exactly k real eigenvalues converges to the probability that B has exactly k real eigenvalues.

Next, we show that the empirical spectral measure of A'/N converges weakly almost surely to the empirical spectral measure of B'. Recall that the eigenvalues of B', denoted $\lambda'_1, \ldots, \lambda'_n$ have been ordered first by increasing magnitude and then by increasing argument in $[0, 2\pi)$. For any fixed realization of B' and coupled matrix A'/N, and for every $0 < \epsilon < r/2$, we have seen that each circle $C_{\lambda'_j,\epsilon}$ contains precisely one eigenvalue of B' and precisely one eigenvalue of A'/N for all N sufficiently large. For each $1 \leq j \leq n$, label the eigenvalue of A'/N in the circle $C_{\lambda'_j,\epsilon}$ by λ'^N_j/N. This implies that for the fixed matrix B', coupled matrix A'/N, and for any $\epsilon > 0$,

$$\left|\lambda'^N_j/N - \lambda'_j\right| < \epsilon$$

for each $1 \leq j \leq n$ and for all N sufficiently large.

Next, stop fixing the entries of B', but continue to condition on B' not having any repeated eigenvalues (which happens with probability one) and continue to couple the matrix A'/N with B'.

We wish to show that

$$\mathbb{P}\left(\bigcap_{1\le j\le n}\left\{\lim_{N\to\infty}\left|\lambda'^N_j/N-\lambda'_j\right|=0\right\}\right)=1.$$

Let $\epsilon = 1/k$ for $k\in\mathbb{N}$ and define the events

$$E_{j,k}=\left\{\limsup_{N\to\infty}\left|\lambda'^N_j/N-\lambda'_j\right|<1/k.\right\}$$

From the discussion above, we know that $\mathbb{P}(E_{j,k}) = 1$ for each $1\le j\le n$ and each $1\le k<\infty$ whenever N is sufficiently large. Define

$$E_j=\bigcap_{k=1}^{\infty}E_{j,k},\qquad \text{for } 1\le j\le n.$$

Observe that $E_{j,1}\supseteq E_{j,2}\supseteq\ldots$ is a nested and decreasing sequence of events. Therefore, by continuity from above of the probability measure,

$$\begin{aligned}\mathbb{P}(E_j)&=\mathbb{P}\left(\bigcap_{k=1}^{\infty}E_{j,k}\right)\\&=\lim_{k\to\infty}\mathbb{P}(E_{j,k})\\&=\inf_{k\ge1}\mathbb{P}(E_{j,k})\\&=1.\end{aligned}$$

Let

$$E=\bigcap_{1\le j\le n}E_j.$$

Then clearly

$$\mathbb{P}(E)=1$$

as well. This shows that

$$\mathbb{P}\left(\bigcap_{1\le j\le n}\left\{\limsup_{N\to\infty}\left|\lambda'^N_j/N-\lambda'_j\right|=0\right\}\right)=1.$$

On the other hand, clearly

$$\mathbb{P}\left(\bigcap_{1\le j\le n}\left\{\liminf_{N\to\infty}\left|\lambda'^N_j/N-\lambda'_j\right|\ge0\right\}\right)=1.$$

Hence, the limit exists, and

$$\mathbb{P}\left(\bigcap_{1\le j\le n}\left\{\lim_{N\to\infty}\left|\lambda'^N_j/N-\lambda'_j\right|=0\right\}\right)=1,$$

so $\lambda'^N_j/N \longrightarrow \lambda'_j$ almost surely, for each $1\le j\le n$. Let $f:\mathbb{C}\to\mathbb{C}$ be a bounded and continuous function. By the continuous mapping theorem,

$$\int_{\mathbb{C}} f\,d\mu_{A'/N}=\frac{1}{n}\sum_{j=1}^{n} f\left(\lambda'^N_j/N\right)\longrightarrow\frac{1}{n}\sum_{j=1}^{n} f(\lambda'_j)=\int_{\mathbb{C}} f\,d\mu_{B'},$$

almost surely as well, proving the weak convergence of the empirical spectral measure of A'/N to the empirical spectral measure of B' almost surely. Since $A'\sim A$ and $B'\sim B$, this means that $\mu_{A'/N}\sim\mu_{A/N}$ and $\mu_{B'}\sim\mu_B$, so it follows that $\mu_{A/N}$ converges weakly in distribution to μ_B.

We now show that the joint distribution of the eigenvalues of A'/N given by

$$(\lambda'^N_1/N,\ldots,\lambda'^N_n/N)$$

converges in probability to the joint distribution of the eigenvalues of B' given by $(\lambda'_1,\ldots,\lambda'_n)$. Since $\lambda'^N_j/N\longrightarrow\lambda'_j$ almost surely for each $1\le j\le n$, we know that $\lambda'^N_j/N\longrightarrow\lambda'_j$ in probability for each $1\le j\le n$ as well. Let $\epsilon>0$. Then

$$\begin{aligned}\lim_{N\to\infty}\mathbb{P}\left(\left\|(\lambda'^N_1/N,\ldots,\lambda'^N_n/N)-(\lambda'_1,\ldots,\lambda'_n)\right\|>\epsilon\right)&\le\lim_{N\to\infty}\mathbb{P}\left(\sum_{j=1}^{n}\left|\lambda'^N_j/N-\lambda'_j\right|>\epsilon\right)\\&\le\lim_{N\to\infty}\sum_{j=1}^{n}\mathbb{P}\left(\left|\lambda'^N_j/N-\lambda'_j\right|>\epsilon/n\right)\\&=0.\end{aligned}$$

In particular, by Theorem 2.2.6, this implies that

$$(\lambda'^N_1/N,\ldots,\lambda'^N_n/N)\overset{\mathrm{d}}{\Longrightarrow}(\lambda'_1,\ldots,\lambda'_n).$$

Since $A\sim A'$ and $B\sim B'$, we conclude that

$$({\lambda_1}^N/N,\ldots,{\lambda_n}^N/N)\overset{\mathrm{d}}{\Longrightarrow}(\lambda_1,\ldots,\lambda_n)$$

as well. $\square$

5.5 Rank one perturbations of large box model matrices

Let A be an n-by-n random matrix with entries independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let P be the n-by-n rank one perturbation matrix whose entries are all μ_N. We now describe the limiting empirical spectral measure of $(A+P)/N$ in the cases where $\mu_N/N \to \infty$, $\mu_N/N \to 0$, or $\mu_N/N \to c$, for $c \in \mathbb{R}$, as $N \to \infty$.

5.5.1 Rank one perturbations of large box model matrices when $\lim_{N\to\infty} \mu_N/N = \infty$

We begin with the easiest case, where the perturbation matrix dominates the original random matrix. The following lemma will be needed to find the eigenvalues of the large perturbation matrix.

Lemma 5.5.1. *The eigenvalues of the n-by-n matrix of all ones*

$$J_n = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

are n with multiplicity 1 and 0 with multiplicity $n-1$. The characteristic polynomial of J_n is given by

$$p_{J_n}(t) = t^n - nt^{n-1}.$$

Proof. Observe that the rank of J_n is 1, so the nullity of J_n must be $n-1$. By inspection, we see that n is an eigenvalue of J_n with corresponding eigenvector $(1, \dots, 1)$. From the fundamental theorem of linear maps (see Theorem 3.22 in [6]), we see that the remaining $n-1$ eigenvalues must all be zero. Once the eigenvalues are known, the characteristic polynomial of J_n can be constructed as

$$p_{J_n}(t) = t^{n-1}(t-n) = t^n - nt^{n-1}.$$

□

We are now ready to prove the main result of this subsection.

Proof of Theorem 5.1.2. First, consider the n-by-n matrix of all ones

$$J_n = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

By Lemma 5.5.1, the characteristic polynomial of J_n is given by

$$p_{J_n}(t) = t^n - nt^{n-1}.$$

The characteristic polynomial $\frac{A+P}{\mu_N}$ can be written in terms of the principal minors of $A+P$ as in (5.1), which is given by

$$p_{\frac{A+P}{\mu_N}}(t) = t^n - E_1\left(\frac{A+P}{\mu_N}\right)t^{n-1} + \cdots + (-1)^{n-1}E_{n-1}\left(\frac{A+P}{\mu_N}\right)t + (-1)^n E_n\left(\frac{A+P}{\mu_N}\right).$$

We will apply Theorem 2.6.5 to the characteristic polynomials of $\frac{A+P}{\mu_N}$ and J_n as $N \to \infty$. This will show that as $N \to \infty$, the eigenvalues of $\frac{A+P}{\mu_N}$ converge almost surely to those of J_n. We begin with some bounds on the coefficients of the difference of the characteristic polynomials of $\frac{A+P}{\mu_N}$ and J_n, i.e.,

$$\begin{aligned}&\left|p_{\frac{A+P}{\mu_N}}(t) - p_{J_n}(t)\right| \\ &= \left|\left(-E_1\left(\frac{A+P}{\mu_N}\right) + n\right)t^{n-1} + E_2\left(\frac{A+P}{\mu_N}\right)t^{n-2} + \cdots + (-1)^n E_n\left(\frac{A+P}{\mu_N}\right)\right|.\end{aligned}$$

For each $1 \leq k \leq n$, we see that $E_k\left(\frac{A+P}{\mu_N}\right) = \left(\frac{1}{\mu_N}\right)^k E_k(A+P)$ by recalling that each E_k is the sum of determinants, which satisfy the property that for a scalar c and n-by-n matrix M, $\det(cM) = c^n \det(M)$.

For $k = 1$, we have that $|E_1(A+P)| = |\mathrm{tr}(A) + n\mu_N|$ and therefore

$$\left|-E_1\left(\frac{A+P}{\mu_N}\right) + n\right| = \left|\frac{\mathrm{tr}(A) + n\mu_N}{\mu_N} - n\right| = \left|\frac{\mathrm{tr}(A)}{\mu_N}\right| \leq \left|\frac{nN}{\mu_N}\right|. \tag{5.6}$$

Since $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, we see that

$$\lim_{N\to\infty}\left|E_1\left(\frac{A+P}{\mu_N}\right) - n\right| \leq \lim_{N\to\infty}\left|\frac{nN}{\mu_N}\right| = 0 \tag{5.7}$$

almost surely. Now consider $|E_k(A+P)|$ for $k \geq 2$. We know that each principal minor of size k of $A+P$ is a determinant of a k-by-k submatrix of $A+P$, and that there are $\binom{n}{k}$ such principal minors. Using sigma notation where S_k is the symmetric group of k elements and σ is a permutation in S_k, and relabeling the rows and columns that appear in the principal submatrix to go from 1 to k, we have that each of the $\binom{n}{k}$ terms in the sum for $|E_k(A+P)|$ can be written as

$$\begin{aligned} &\left|\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)(A+P)_{1\sigma(1)} \cdots (A+P)_{k\sigma(k)}\right| \\ &= \left|\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)(A_{1\sigma(1)} + \mu_N) \cdots (A_{k\sigma(k)} + \mu_N)\right|. \end{aligned}$$

The symmetric group S_k has $k!$ permutations. This number is even for $k \geq 2$, and exactly half of the permutations are even and half are odd. This means that $\operatorname{sgn}(\sigma) = 1$ for half of the permutations and $\operatorname{sgn}(\sigma) = -1$ for the other half. Consider the expansion of the product $(A_{1\sigma(1)} + \mu_N) \cdots (A_{k\sigma(k)} + \mu_N)$. Each permutation with $\operatorname{sgn}(\sigma) = 1$ contributes μ_N^k and each permutation with $\operatorname{sgn}(\sigma) = -1$ contributes $-\mu_N^k$ to the leading term in μ_N. Since there are the same number of even and odd permutations, these cancel out, and each term of $|E_k(A+P)|$ is at most

$$\sum_{\sigma \in S_k} \left(\left|(A_{1\sigma(1)} + \mu_N) \cdots (A_{k\sigma(k)} + \mu_N)\right| - |\mu_N|^k \right).$$

Altogether, we can bound

$$|E_k(A+P)| \leq \binom{n}{k} k! \left(|N + \mu_N|^k - |\mu_N|^k \right).$$

By the binomial expansion, we have

$$|N + \mu_N|^k \leq (N + |\mu_N|)^k = \sum_{l=0}^{k} \binom{k}{l} N^{k-l} |\mu_N|^l.$$

Subtracting $|\mu_N|^k$, multiplying by the $k!$ permutations in S_k, and accounting for the $\binom{n}{k}$ principal minors, we have the overall bound

$$\begin{aligned} |E_k(A+P)| &\leq \binom{n}{k} k! \left(\left(\sum_{l=0}^{k} \binom{k}{l} N^{k-l} |\mu_N|^l \right) - |\mu_N|^k \right) \\ &= \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} N^{k-l} |\mu_N|^l. \end{aligned}$$

Therefore, for $k \geq 2$,

$$\frac{|E_k(A+P)|}{|\mu_N|^k} \leq \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{N}{|\mu_N|}\right)^{k-l}. \tag{5.8}$$

Hence,

$$\lim_{N\to\infty} \left| E_k\left(\frac{A+P}{\mu_N}\right)\right| = 0 \tag{5.9}$$

almost surely, since $\lim_{N\to\infty} \frac{N}{\mu_N} = 0$.

Let $0 < \epsilon < 1/4$ and let $C_{n,\epsilon}$ be the circle of radius ϵ around the eigenvalue n of J_n. Note that this circle has been chosen in a way that avoids all of the eigenvalues of J_n, which are 0 (with multiplicity $n-1$) and n (with multiplicity 1).

Our goal is to show that

$$\mathbb{P}\left(\limsup_{N\to\infty} \left|p_{\frac{A+P}{\mu_N}}(t) - p_{J_n}(t)\right| < |p_{J_n}(t)|, \text{ for all } t \in C_{n,\epsilon}\right) = 1. \tag{5.10}$$

Applying the triangle inequality to the difference of the characteristic polynomials of $\frac{A+P}{\mu_N}$ and J_n, and using (5.6) and (5.8), we have

$$\begin{aligned} \left|p_{\frac{A+P}{\mu_N}}(t) - p_{J_n}(t)\right| &\leq |t^n - t^n| + \left|E_1\left(\frac{A+P}{\mu_N}\right) - n\right| |t|^{n-1} + \sum_{k=2}^{n} |t|^{n-k} \left|E_k\left(\frac{A+P}{\mu_N}\right)\right| \\ &\leq \left|\frac{nN}{\mu_N}\right| |t|^{n-1} + \sum_{k=2}^{n} |t|^{n-k} \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{N}{|\mu_N|}\right)^{k-l}. \end{aligned} \tag{5.11}$$

We also have that

$$|p_{J_n}(t)| = |t^n - nt^{n-1}| = |t|^{n-1}|t-n|. \tag{5.12}$$

For every $t \in C_{n,\epsilon}$, we can bound (5.11) above by

$$\left|\frac{nN}{\mu_N}\right| |n+\epsilon|^{n-1} + \sum_{k=2}^{n} |n+\epsilon|^{n-k} \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{N}{|\mu_N|}\right)^{k-l} \tag{5.13}$$

and we can bound (5.12) below by

$$\begin{aligned} |t|^{n-1}|t-n| &\geq |n-\epsilon|^{n-1}|\epsilon| \\ &> |n-2\epsilon|^{n-1}|\epsilon|. \end{aligned} \tag{5.14}$$

By (5.7) and (5.9), we can choose N sufficiently large so that, for all $t \in C_{n,\epsilon}$, (5.13) is smaller than (5.14), i.e., for all $t \in C_{n,\epsilon}$ and N sufficiently large,

$$\left| p_{\frac{A+P}{\mu_N}}(t) - p_{J_n}(t) \right| < |p_{J_n}(t)|,$$

with probability one, showing that (5.10) also holds.

Let $\lambda_1^N/\mu_N, \ldots, \lambda_n^N/\mu_N$ denote the eigenvalues of $\frac{A+P}{\mu_N}$, which are ordered first by increasing magnitude and then by increasing argument in $[0, 2\pi)$.

By Theorem 2.6.5, we conclude that as $N \to \infty$, the characteristic polynomial of the scaled matrix $\frac{A+P}{\mu_N}$ has exactly one root in $C_{n,\epsilon}$, for each $0 < \epsilon < 1/4$. We will see shortly that the other eigenvalues of $\frac{A+P}{\mu_N}$ converge almost surely to zero, so this eigenvalue must be λ_n^N/μ_N. Hence, for any $\epsilon > 0$,

$$\mathbb{P}\left(\limsup_{N\to\infty} \left| \lambda_n^N/\mu_N - n \right| < \epsilon \right) = 1,$$

showing that as $N \to \infty$, the eigenvalue λ_n^N/μ_N of $\frac{A+P}{\mu_N}$ converges to n almost surely.

A similar argument also shows that the other eigenvalues of $(A+P)/\mu_N$ converge to 0 almost surely as $N \to \infty$. Let $0 < \epsilon < 1/4$, and consider $t \in C_{0,\epsilon}$. Note again that this circle has been chosen to avoid the $n-1$ eigenvalues of J_n at the origin and the eigenvalue at n.

For every $t \in C_{0,\epsilon}$, we can bound (5.11) above by

$$\begin{aligned} &\left| \frac{nN}{\mu_N} \right| |\epsilon|^{n-1} + \sum_{k=2}^{n} |\epsilon|^{n-k} \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{N}{|\mu_N|} \right)^{k-l} \\ \leq &\frac{nN}{|\mu_N|} + \sum_{k=2}^{n} \frac{n!}{(n-k)!} \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{N}{|\mu_N|} \right)^{k-l} \end{aligned} \tag{5.15}$$

and we can bound (5.12) below by

$$\begin{aligned} |t|^{n-1}|t-n| &\geq |\epsilon|^{n-1}|n-\epsilon| \\ &> |\epsilon|^{n-1}|n-2\epsilon|. \end{aligned} \tag{5.16}$$

Since we can again choose N sufficiently large so that (5.15) is smaller than (5.16) for all $t \in C_{0,\epsilon}$, this shows that

$$\mathbb{P}\left(\limsup_{N\to\infty} \left| p_{\frac{A+P}{\mu_N}}(t) - p_{J_n}(t) \right| < |p_{J_n}(t)|, \text{ for all } t \in C_{0,\epsilon} \right) = 1.$$

By Theorem 2.6.5, we conclude that as $N \to \infty$, the characteristic polynomial of the scaled perturbed matrix $\frac{A+P}{\mu_N}$ has exactly $n-1$ roots in $C_{0,\epsilon}$, for each $0 < \epsilon < 1/4$. Hence, for any $\epsilon > 0$, and eigenvalues λ_j^N/μ_N for $1 \leq j \leq n-1$,

$$\mathbb{P}\left(\bigcap_{1\leq j\leq n-1}\left\{\limsup_{N\to\infty}|\lambda_j^N/\mu_N| < \epsilon\right\}\right) = 1,$$

showing that the remaining $n-1$ eigenvalues of $\frac{A+P}{\mu_N}$ converge to 0 almost surely. □

The rate of the convergence of the eigenvalues in Theorem 5.1.2 is discussed in Section 5.6.

5.5.2 Rank one perturbations of large box model matrices when $\lim_{N\to\infty}\mu_N/N = 0$

Next, we show that if the matrix is perturbed by adding a matrix with sufficiently small entries, the normalized empirical spectral measure of the perturbed matrix converges weakly in distribution to the same limiting measure as that of the normalized unperturbed matrix. We are ready to prove the main result.

Proof of Theorem 5.1.3. Let B be the matrix whose entries are iid and uniformly distributed on $[-1,1]$. By Theorem 5.1.1, it is clear that the empirical spectral measure of A/N converges weakly in distribution to the empirical spectral measure of B. We now show that the empirical spectral measure of $(A+P)/N$ has the same limiting behavior.

We will perform Rouché's theorem on the characteristic polynomials of the matrices

$$A/N = \begin{pmatrix} a_{11}/N & \dots & a_{1n}/N \\ \vdots & \ddots & \vdots \\ a_{n1}/N & \dots & a_{nn}/N \end{pmatrix}$$

and

$$\frac{A+P}{N} = \begin{pmatrix} (a_{11}+\mu_N)/N & \dots & (a_{1n}+\mu_N)/N \\ \vdots & \ddots & \vdots \\ (a_{n1}+\mu_N)/N & \dots & (a_{nn}+\mu_N)/N \end{pmatrix}.$$

We have that

$$p_{\frac{A}{N}}(t) = t^n - E_1\left(\frac{A}{N}\right)t^{n-1} + \cdots + (-1)^{n-1}E_{n-1}\left(\frac{A}{N}\right)t + (-1)^n E_n\left(\frac{A}{N}\right)$$

and

$$p_{\frac{A+P}{N}}(t) = t^n - E_1\left(\frac{A+P}{N}\right)t^{n-1} + \cdots + (-1)^{n-1}E_{n-1}\left(\frac{A+P}{N}\right)t + (-1)^n E_n\left(\frac{A+P}{N}\right).$$

We begin by bounding $\left|E_k\left(\frac{A+P}{N}\right) - E_k\left(\frac{A}{N}\right)\right|$ for $1 \leq k \leq n$ by applying Lemma 5.3.4 with $R = \frac{A+P}{N}$ and $S = \frac{A}{N}$. Observe that

$$\left\|\frac{A+P}{N}\right\|_{\max} \leq 1 + |\mu_N|/N$$

$$\left\|\frac{A}{N}\right\|_{\max} \leq 1$$

and

$$\left\|\frac{A+P}{N} - \frac{A}{N}\right\|_{\max} \leq |\mu_N|/N.$$

By Lemma 5.3.4,

$$\begin{aligned}\left|E_k\left(\frac{A+P}{N}\right) - E_k\left(\frac{A}{N}\right)\right| &\leq \frac{n!}{(n-k)!}\left\|\frac{A+P}{N} - \frac{A}{N}\right\|_{\max}\left(\left\|\frac{A+P}{N}\right\|_{\max} + \left\|\frac{A}{N}\right\|_{\max}\right)^{k-1}. \\ &\leq \frac{n!}{(n-k)!}\left(\frac{|\mu_N|}{N}\right)\left(\frac{|\mu_N|}{N} + 1\right)^{k-1}.\end{aligned}$$

This leads to the bound

$$\left|p_{\frac{A+P}{N}}(t) - p_{\frac{A}{N}}(t)\right| \leq |t^n - t^n| + \sum_{k=1}^{n} |t|^{n-k}\frac{n!}{(n-k)!}\left(\frac{|\mu_N|}{N}\right)\left(\frac{|\mu_N|}{N} + 1\right)^{k-1}. \tag{5.17}$$

Let $\lambda_1(A), \ldots, \lambda_n(A)$ denote the eigenvalues of A; then $\lambda_1(A)/N, \ldots, \lambda_n(A)/N$ denote the eigenvalues of A/N by the previous discussion of the scaling of eigenvalues. Similarly, let $\lambda_1(A+P)/N, \ldots, \lambda_n(A+P)/N$ denote the eigenvalues of $(A+P)/N$. These will be the labels of our eigenvalues, though we will reorder them momentarily. Consider the factorization

$$\left|p_{\frac{A}{N}}(t)\right| = |t - \lambda_1(A)/N| \cdots |t - \lambda_n(A)/N|. \tag{5.18}$$

As $N \to \infty$, we wish to show that (5.17) is strictly less than (5.18) for all t in appropriately chosen circles with probability one.

Recall that B is the n-by-n random matrix whose entries are iid and uniformly distributed on $[-1, 1]$. Let the entries of the matrix A be coupled with the entries of the matrix B as in the statement of Lemma 5.3.5 and couple the random matrices A and B as in Remark 3.3.3. Refer to the coupled matrices as A' and B'. Let $\lambda_1(B'), \ldots, \lambda_n(B')$ denote the eigenvalues of B', ordered first by increasing magnitude and then by increasing argument in $[0, 2\pi)$. Fix a realization of B' that does not have any repeated eigenvalues, which we can do with probability one by Lemma 4.5.5. Note that for every $N \in \mathbb{N}$, this coupling also fixes A'/N and $(A' + P)/N$. For this fixed realization of B', note that all eigenvalues are at least some small positive distance apart; call this distance δ. Since the characteristic polynomials of A'/N and B' are both monic, and since $|E_k(A'/N) - E_k(B')| = O(1/N)$ by Lemma 5.3.5, Theorem 5.3.6 implies that we can take N large enough so that for $1 \leq j \leq n$, each circle of radius $0 < \epsilon < \delta/4$ around the eigenvalue $\lambda_j(B')$ contains exactly one eigenvalue of A'/N; denote this eigenvalue of A'/N by $\lambda_j(A')/N$. Then for every $1 \leq j \leq n$, the eigenvalues of A'/N and B' satisfy the inequalities

$$\left| \lambda_j(B') - \frac{\lambda_j(A')}{N} \right| < \delta/4$$

for all N sufficiently large. From this, for each $j \neq k$ and N sufficiently large, we can see that the difference between two eigenvalues of A'/N is at least

$$\begin{aligned}
\left| \frac{\lambda_j(A')}{N} - \frac{\lambda_k(A')}{N} \right| &= \left| \left(\lambda_j(B') - \lambda_k(B')\right) + \left(\frac{\lambda_j(A')}{N} - \lambda_j(B') \right) + \left(\lambda_k(B') - \frac{\lambda_k(A')}{N} \right) \right| \\
&\geq \left| \left|\lambda_j(B') - \lambda_k(B')\right| - \left| \left(\frac{\lambda_j(A')}{N} - \lambda_j(B') \right) + \left(\lambda_k(B') - \frac{\lambda_k(A')}{N} \right) \right| \right| \\
&\geq \left|\lambda_j(B') - \lambda_k(B')\right| - \left| \left(\frac{\lambda_j(A')}{N} - \lambda_j(B') \right) + \left(\lambda_k(B') - \frac{\lambda_k(A')}{N} \right) \right| \\
&\geq \left|\lambda_j(B') - \lambda_k(B')\right| - \left| \frac{\lambda_j(A')}{N} - \lambda_j(B') \right| - \left| \lambda_k(B') - \frac{\lambda_k(A')}{N} \right| \\
&> \delta - \delta/4 - \delta/4 \\
&= \delta/2,
\end{aligned}$$

where the second line follows from the reverse triangle inequality, the third line follows from the definition of absolute value, and the fourth line follows from the triangle inequality.

Now, for some $1 \le j \le n$, consider the eigenvalue $\lambda_j(A')/N$ of A'/N. Let $0 < \epsilon < \delta/8$ and consider $t \in C_{\lambda_j(A')/N,\epsilon}$. We have the following lower bound for (5.18).

$$\begin{aligned} \left|p_{\frac{A'}{N}}(t)\right| &= |t - \lambda_1(A')/N| \cdots |t - \lambda_n(A')/N| \\ &\ge \epsilon|\delta/2 - \delta/8|^{n-1} \\ &> \epsilon|\delta/4|^{n-1} \end{aligned} \tag{5.19}$$

for all $t \in C_{\lambda_j(A')/N,\epsilon}$. By Lemma 5.3.12, the eigenvalues $\lambda_1(A')/N, \dots, \lambda_n(A')/N$ are all contained in $C_{0,n}$ almost surely. Hence for all $t \in C_{\lambda_j(A')/N,\epsilon}$, we can bound $|t| \le |\lambda_j(A')/N| + \epsilon \le n + \delta/8$, so an upper bound for (5.17) is

$$\sum_{k=1}^{n} \frac{n!}{(n-k)!}|n + \delta/8|^{n-k} \left(\frac{|\mu_N|}{N}\right)\left(\frac{|\mu_N|}{N} + 1\right)^{k-1}. \tag{5.20}$$

Since $\lim_{N\to\infty} \frac{\mu_N}{N} = 0$ and n is fixed, we have that (5.20) is less than (5.19) for N sufficiently large for the fixed realization B'. By 2.6.5, for each $1 \le j \le n$, A'/N and $(A' + P)/N$ both have exactly one eigenvalue in $C_{\lambda_j(A')/N,\epsilon}$, whenever N is sufficiently large. Label this corresponding eigenvalue of $(A' + P)/N$ by $\lambda_j(A' + P)/N$, for $1 \le j \le n$. This implies that for the fixed matrix B', coupled matrices A'/N and $(A' + P)/N$, and for any $\epsilon > 0$,

$$\left|\lambda_j(A')/N - \lambda_j(A' + P)/N\right| < \epsilon$$

for all N sufficiently large. Next, we stop fixing the matrix B', but continue to condition that B' does not have any repeated eigenvalues (which happens with probability one) and continue to couple the matrix A'/N (and hence $(A' + P)/N$) with B'. From the discussion at the end of the proof of Theorem 5.1.1, we see that

$$\mathbb{P}\left(\bigcap_{1\le j\le n}\left\{\lim_{N\to\infty}\left|\lambda_j(A')/N - \lambda_j(A' + P)/N\right| = 0\right\}\right) = 1,$$

and so we conclude that the empirical spectral measure of $(A' + P)/N$ converges weakly almost surely to the same limit as the empirical spectral measure of A'/N. Since $A' \sim A$ and $B' \sim B$, we conclude that the empirical spectral measures of A/N and $(A+P)/N$ must both converge weakly in distribution to that of B. Continuing to proceed as in the proof of Theorem 5.1.1, we see that the joint distribution of eigenvalues of A/N and $(A+P)/N$ both converge in distribution to the joint distribution of eigenvalues of B. □

5.5.3 Rank one perturbations of large box model matrices when $\lim_{N\to\infty} \mu_N/N = c$

We now consider the case when $\lim_{N\to\infty} \mu_N/N = c$, for some $c > 0 \in \mathbb{R}$.

Proof of Theorem 5.1.4. Similarly to Lemma 5.3.5, consider the coupling of matrix entries where for each $1 \leq k \leq 2N+1$, we let

$$a_{ij} = \begin{cases} -N+(k-1), & \text{if } b_{ij} \in \left[-1+c+\frac{2(k-1)}{2N+1}, -1+c+\frac{2k}{2N+1}\right) \\ N, & \text{if } b_{ij} = 1+c \end{cases}. \tag{5.21}$$

For the coupled entries a_{ij} and b_{ij}, we begin by bounding $|(a_{ij}+\mu_N)/N - b_{ij}|$. If $b_{ij} = 1+c$, we have that

$$|(a_{ij}+\mu_N)/N - b_{ij}| = \left|\frac{N+\mu_N}{N} - (1+c)\right| \leq \left|\frac{\mu_N}{N} - c\right| = o(1)$$

since $\lim_{N\to\infty} \frac{\mu_N}{N} = c$ with an unspecified rate of convergence. If $b_{ij} \neq 1+c$, we examine the distance of $(a_{ij}+\mu_N)/N$ from both endpoints of the interval containing b_{ij}. A similar computation shows that in both cases, for $1 \leq l \leq 2N+1$,

$$|(a_{ij}+\mu_N)/N - b_{ij}| = o(1)$$

as well. Since all entries of A and B have been coupled in this way, it follows that

$$\|(A+\mu_N)/N - B\|_{\max} = o(1)$$

almost surely, where the implicit constant depends on the rate of convergence of μ_N/N. Furthermore,

$$\|(A+\mu_N)/N\|_{\max} \leq |\mu_N|/N + 1$$

and

$$\|B\|_{\max} \leq c + 1.$$

Applying Lemma 5.3.4 gives the bound

$$\begin{aligned}|E_k((A+\mu_N)/N) - E_k(B)| &\leq \frac{n!}{(n-k)!}\|(A+\mu_N)/N - B\|_{\max}(||(A+\mu_N)/N||_{\max} + ||B||_{\max})^{k-1} \\ &\leq \frac{n!}{(n-k)!}\|(A+\mu_N)/N - B\|_{\max}\left(|\mu_N|/N + c + 2\right)^{k-1} \\ &= o(1)\end{aligned}$$

almost surely.

Refer to the coupled matrices as $A' + P$ and B'. Now fix the matrix B' so that B' does not have any repeated eigenvalues, which happens with probability one by Lemma 4.5.5. This also fixes the entries of $A' + P$ due to the assumed coupling. Let $\delta > 0$ be the minimum distance between the eigenvalues of B'. Now proceed exactly as in the proof of Theorem 5.1.1, but replacing A' with $A' + P$ and using the $o(1)$ bound obtained above when bounding $|E_k((A+\mu_N)/N) - E_k(B)|$. Then we see that for every $0 < \epsilon < \delta/4$, we can find N large enough so that the circles of radius ϵ centered at the eigenvalues of B' all contain exactly one eigenvalue of $(A'+P)/N$. Continuing as in the proof of Theorem 5.1.1, as $N \to \infty$, we see that the empirical spectral measure of $(A' + P)/N$ converges weakly almost surely to the empirical spectral measure B'. Moreover, since $A \sim A'$ and $B \sim B'$, the empirical spectral measure of $(A + P)/N$ converges weakly in distribution to the empirical spectral measure of B. □

The above result shows that as $N \to \infty$, the outlier eigenvalues of $(A+P)/N$ behave the same as the outlier eigenvalues of B. The following theorem briefly shows that for certain values of c, the matrices with entries independently and uniformly distributed on $[-1+c, 1+c]$ are guaranteed to have outliers. In Section 5.6, we will investigate the outliers of $(A+P)/N$ more thoroughly (see for instance Theorem 5.6.5), and show that under certain conditions, exactly one outlier eigenvalue exists.

Theorem 5.5.2. *Let $c \geq 2$. Then every n-by-n random matrix A with entries independently and uniformly distributed on $(-1+c, 1+c)$ has at least one eigenvalue outside of $B(0,n)$ with probability one.*

Proof. This follows immediately from Theorem 5.3.10. We have

$$\min_{1\leq i\leq n} \sum_{j=1}^{n} A_{ij} \leq \rho(A) \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} A_{ij}.$$

When $c \geq 2$, we have that $A_{ij} > 1$ with probability one, and so

$$n = \sum_{j=1}^{n} 1 < \min_{1\leq i\leq n} \sum_{j=1}^{n} A_{ij} \leq \rho(A)$$

with probability one. □

5.6 Eigenvalue fluctuations for rank one perturbations

In this section, we discuss the fluctuations of the largest eigenvalue of rank one perturbations of large box model random matrices. In particular, for the n-by-n random matrix A with iid entries uniformly distributed on $[-N,N] \cap \mathbb{Z}$, we show that the centered and scaled largest eigenvalue (in magnitude) of $A+P$ converges in distribution to $\frac{1}{n^2}\sum_{i=1}^{n^2} X_i$ where each X_i is uniformly distributed on $(-1,1)$ and jointly independent; this is a Bates distribution with n^2 parameters.

We begin now with the 2-by-2 case, which demonstrates how we first noticed this phenomenon. The 2-by-2 case is also particularly nice because we are able to find the joint fluctuations of the eigenvalues.

Theorem 5.6.1. *Let A be a 2-by-2 random matrix with entries independently and uniformly distributed on $[-N,N] \cap \mathbb{Z}$. Let P be the perturbation matrix whose entries are all μ_N, with $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$. Let λ_+ denote the largest, in magnitude, eigenvalue of $A+P$; let λ_- denote the smallest, in magnitude, eigenvalue of $A+P$. Then as $N \to \infty$,*

$$\left(\frac{\lambda_+ - 2\mu_N}{2N}, \frac{\lambda_-}{2N}\right) \stackrel{d}{\Rightarrow} \left(\frac{W+X+Y+Z}{4}, \frac{(W+Z)-(X+Y)}{4}\right).$$

where W, X, Y, Z are independently and uniformly distributed on $(-1,1)$.

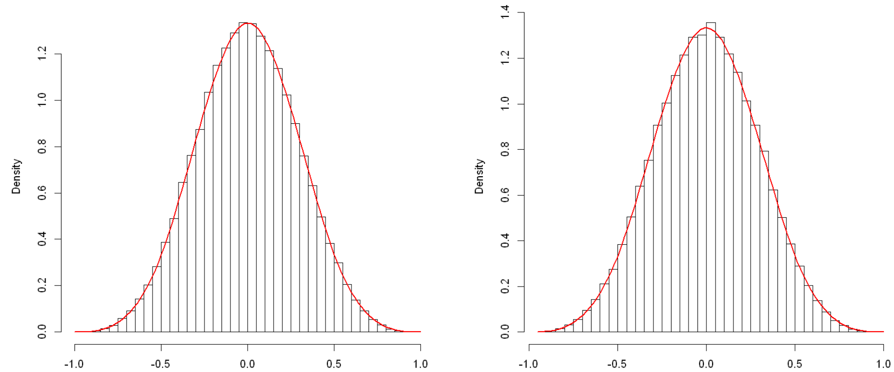

Figure 5.4: These images demonstrate Theorem 5.6.1. The left image shows a histogram of the centered and scaled largest eigenvalues of 50000 2-by-2 random matrices with entries independently and uniformly distributed on $[-10^{10}, 10^{10}] \cap \mathbb{Z}$ and perturbed by $(10^{10})^{(1.5)}$. The eigenvalues were centered by -2×10^{10} and scaled by 2×10^{10}. The right image shows a histogram of the scaled smallest eigenvalues of the same matrices; these eigenvalues were scaled by 2×10^{10}. The red lines in both histograms show the pdf of $\frac{\sum_{i=1}^{4} X_i}{4}$, where each X_i is independently and uniformly distributed on $(-1, 1)$.

Theorem 5.6.1 is illustrated in Figure 5.4.

Proof. We want to find the joint fluctuations of the eigenvalues of the matrix

$$A + P = \begin{pmatrix} A_{11} + \mu_N & A_{12} + \mu_N \\ A_{21} + \mu_N & A_{22} + \mu_N \end{pmatrix}.$$

The characteristic polynomial of $A + P$ is given by

$$t^2 - (A_{11} + A_{22} + 2\mu_N)t + (A_{11} + \mu_N)(A_{22} + \mu_N) - (A_{12} + \mu_N)(A_{21} + \mu_N).$$

From the quadratic formula, we know that the eigenvalues λ_+ and λ_- of $A + P$ are given by

$$\lambda_+ = \frac{(A_{11} + A_{22} + 2\mu_N) + \sqrt{(A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N + \mu_N^2)}}{2}$$

and

$$\lambda_- = \frac{(A_{11} + A_{22} + 2\mu_N) - \sqrt{(A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N + \mu_N^2)}}{2}.$$

Centering λ_+ by $-2\mu_N$ and normalizing by $2N$, we have

$$\frac{\lambda_+ - 2\mu_N}{2N} = \frac{(A_{11} + A_{22})}{4N} + \left(-\frac{\mu_N}{2N} + \frac{\sqrt{(A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N + \mu_N^2)}}{4N} \right).$$

Multiplying and dividing the second term of the above summation by its conjugate and simplifying, we have that

$$\frac{\lambda_+ - 2\mu_N}{2N} = \frac{(A_{11} + A_{22})}{4N} + \frac{-\left((A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N)\right)/(16N^2)}{\left(-\frac{\mu_N}{2N} - \frac{\sqrt{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}}{4N}\right)}.$$

Now multiplying and dividing the second term by $\frac{N}{\mu_N}$, we get

$$\begin{aligned}\frac{\lambda_+ - 2\mu_N}{2N} &= \frac{(A_{11} + A_{22})}{4N} + \frac{-\left((A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N)\right)/(16N\mu_N)}{(N/\mu_N)\left(-\frac{\mu_N}{2N} - \frac{\sqrt{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}}{4N}\right)} \\ &= \frac{(A_{11} + A_{22})}{4N} + \frac{-\left((A_{11} - A_{22})^2 + 4(A_{12}A_{21} + A_{12}\mu_N + A_{21}\mu_N)\right)/(16N\mu_N)}{\left(-\frac{1}{2} - \sqrt{\frac{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}{16\mu_N^2}}\right)}.\end{aligned}$$

A similar process shows that for the smallest eigenvalue λ_-, we have that

$$\begin{aligned}
\frac{\lambda_-}{2N} &= \frac{(A_{11}+A_{22})}{4N} \\
&\quad + \left(\frac{\mu_N}{2N} - \frac{\sqrt{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}}{4N}\right) \\
&= \frac{(A_{11}+A_{22})}{4N} \\
&\quad + \frac{-\left((A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N)\right)/(16N^2)}{\left(\frac{\mu_N}{2N}+\frac{\sqrt{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}}{4N}\right)} \\
&= \frac{(A_{11}+A_{22})}{4N} \\
&\quad + \frac{-\left((A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N)\right)/(16N\mu_N)}{(N/\mu_N)\left(\frac{\mu_N}{2N}+\frac{\sqrt{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}}{4N}\right)} \\
&= \frac{(A_{11}+A_{22})}{4N} \\
&\quad + \frac{-\left((A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N)\right)/(16N\mu_N)}{\left(\frac{1}{2}+\sqrt{\frac{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}{16\mu_N^2}}\right)}.
\end{aligned}$$

Now consider the random vector $\left(\frac{\lambda_+-2\mu_N}{2N}, \frac{\lambda_-}{2N}\right)$ and let $(s,t)\in\mathbb{R}^2$.

By Theorem 5.3.6 and Theorem 2.2.5, we see that

$$\begin{aligned}
&s\left(\frac{\lambda_+-2\mu_N}{2N}\right)+t\left(\frac{\lambda_-}{2N}\right) \\
&= s\left(\frac{(A_{11}+A_{22})}{4N}\right) \\
&\quad + s\left(\frac{-\left((A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N)\right)/(16N\mu_N)}{\left(-\frac{1}{2}-\sqrt{\frac{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}{16\mu_N^2}}\right)}\right) \\
&+ t\left(\frac{(A_{11}+A_{22})}{4N}\right) \\
&\quad + t\left(\frac{-\left((A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N)\right)/(16N\mu_N)}{\left(\frac{1}{2}+\sqrt{\frac{(A_{11}-A_{22})^2+4(A_{12}A_{21}+A_{12}\mu_N+A_{21}\mu_N+\mu_N^2)}{16\mu_N^2}}\right)}\right)
\end{aligned}$$

converges in distribution to

$$s\left(\frac{W+X+Y+Z}{4}\right)+t\left(\frac{(W+Z)-(X+Y)}{4}\right),$$

where W, X, Y, Z are independently and uniformly distributed on $(-1, 1)$. Theorem 5.3.7 allows us to conclude that as $N \to \infty$,

$$\left(\frac{\lambda_+ - 2\mu_N}{2N}, \frac{\lambda_-}{2N}\right) \overset{\text{d}}{\Rightarrow} \left(\frac{W+X+Y+Z}{4}, \frac{(W+Z)-(X+Y)}{4}\right).$$

$\square$

We wish to prove a similar theorem for a matrix of any size. We will need to adapt our method, because explicit formulas for eigenvalues of higher dimensional matrices are either unmanageable or do not exist.

We will need the following two theorems.

Theorem 5.6.2 (Sylvester's Determinant Theorem, see [3, 153])**.** *Let A be an n-by-k matrix and B be a k-by-n matrix. Then*

$$\det(I_n + AB) = \det(I_k + BA). \tag{5.22}$$

When $k = 1$ in the above theorem, the right-hand side is simply the determinant of a number. This will be extremely useful when we are considering rank one perturbations. This identity was used by Tao [153] when considering small rank perturbations.

Theorem 5.6.3 (Exercise 5.6.P26 in [79])**.** *If $A \in M_n(\mathbb{C})$ and $\rho(A) < 1$, then $\sum_{i=0}^{\infty} A^i = (I - A)^{-1}$, where $\sum_{i=0}^{\infty} A^i$ is called the **Neumann series** of A.*

Making use of the Neumann series to analyze small rank perturbations was another technique used by Tao in [153].

Lemma 5.6.4. *Let A be an n-by-n random matrix with entries independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let A_{ij}^k denote the ij entry of A^k. Then for any $k \geq 1$,*

$$\sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij}^k \leq n^{k+1} N^k$$

almost surely.

Proof. Consider the n-by-n all ones matrix J_n. We will first show by induction that

$$J_n^k = \begin{pmatrix} n^{k-1} & \cdots & n^{k-1} \\ \vdots & \ddots & \vdots \\ n^{k-1} & \cdots & n^{k-1} \end{pmatrix}.$$

When $k = 1$, the claim is clearly true and the base case is established. Let $k \geq 1$, and let our inductive hypothesis be that

$$J_n^{k-1} = \begin{pmatrix} n^{k-2} & \cdots & n^{k-2} \\ \vdots & \ddots & \vdots \\ n^{k-2} & \cdots & n^{k-2} \end{pmatrix}.$$

Then

$$\begin{aligned} J_n^k &= J_n^{k-1} J_n \\ &= \begin{pmatrix} n^{k-2} & \cdots & n^{k-2} \\ \vdots & \ddots & \vdots \\ n^{k-2} & \cdots & n^{k-2} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n n^{k-2} & \cdots & \sum_{i=1}^n n^{k-2} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n n^{k-2} & \cdots & \sum_{i=1}^n n^{k-2} \end{pmatrix} \\ &= \begin{pmatrix} n^{k-1} & \cdots & n^{k-1} \\ \vdots & \ddots & \vdots \\ n^{k-1} & \cdots & n^{k-1} \end{pmatrix}, \end{aligned}$$

so the statement holds by the principle of mathematical induction. Then

$$\sum_{j=1}^n \sum_{i=1}^n (J_n^k)_{ij} = n^2 n^{k-1} = n^{k+1}.$$

Finally, it follows that

$$\sum_{j=1}^n \sum_{i=1}^n A_{ij}^k \leq \sum_{j=1}^n \sum_{i=1}^n (NJ_n)_{ij}^k \leq \sum_{j=1}^n \sum_{i=1}^n N^k (J_n^k)_{ij} \leq n^{k+1} N^k$$

almost surely. □

We are now ready to prove the main results of this section. Our proofs are an application of the methods used by Tao in [153], described in Section 1 and above.

Proof of Theorem 5.1.8. Suppose that x is an eigenvector of $(A+P)$ with corresponding eigenvalue λ. We see that

$$\left(\frac{A+P}{N}\right)x = \left(\frac{\lambda}{N}\right)x,$$

so it suffices to consider the distribution of largest eigenvalue, in magnitude, of $(A+P)/N$. In terms of the characteristic polynomial, we wish to find the largest z such that

$$\det\left(zI - \frac{A+P}{N}\right) = 0.$$

From Theorem 5.1.2, we know that $(A+P)/\mu_N$ has one outlier eigenvalue, which converges almost surely to n as $N \to \infty$. Hence, for all N sufficiently large, and any eigenvalue λ of $\frac{A+P}{\mu_N}$ with eigenvector x, we have that

$$\left(\frac{A+P}{\mu_N}\right)x = \lambda x$$

implies that

$$\left(\frac{A+P}{N}\right)x = \frac{\mu_N \lambda}{N}x.$$

Since $\frac{\mu_N \lambda}{N} > 2\lambda$ for all N sufficiently large whenever $\lambda \neq 0$, we see that $\left(\frac{A+P}{N}\right)$ has at least one eigenvalue outside of the disk of radius n centered at the origin as $N \to \infty$. On the other hand, from Lemma 5.3.12, we know that all of the eigenvalues of A/N are located inside the disk of radius n centered at the origin. Therefore, in order to locate the eigenvalue of $\left(\frac{A+P}{N}\right)$ that is outside of the disk of radius n centered at the origin, we may consider only $|z| > n$ and analyze the factorization (originally used by Tao in [153] to apply the eigenvalue criterion)

$$\det\left(zI - \frac{A+P}{N}\right) = \det\left(zI - \frac{A}{N}\right)\det\left(I - \left(zI - \frac{A}{N}\right)^{-1}\frac{P}{N}\right).$$

Since the first factor on the right hand side is non-zero for $|z| > n$, it suffices to find $|z| > n$ such that

$$\det\left(I - \left(zI - \frac{A}{N}\right)^{-1}\frac{P}{N}\right) = 0.$$

Letting $\mathbb{1}^T = (1, 1, \ldots, 1)$, we may write

$$\frac{P}{N} = \frac{\mu_N}{N} \mathbb{1}\mathbb{1}^T.$$

We have

$$\begin{aligned}
\det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{P}{N}\right) &= \det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{\mu_N}{N} \mathbb{1}\mathbb{1}^T\right) \\
&= \det\left(I - \mathbb{1}^T \left(zI - \frac{A}{N}\right)^{-1} \frac{\mu_N}{N} \mathbb{1}\right) \\
&= 1 - \left(\mathbb{1}^T \left(zI - \frac{A}{N}\right)^{-1} \frac{\mu_N}{N} \mathbb{1}\right)_{11} \\
&= 1 - \frac{\mu_N}{N} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(zI - \frac{A}{N}\right)^{-1}_{ij}
\end{aligned}$$

where we have used Theorem 5.6.2 to go from the first line to the second line and evaluated the determinant of the 1-by-1 resultant matrix to go from the second line to the third line. The fourth line follows by inspection. Next, observe that

$$\left(zI - \frac{A}{N}\right)^{-1} = \left(z\left(I - \frac{A}{Nz}\right)\right)^{-1} = \frac{1}{z}\left(I - \frac{A}{Nz}\right)^{-1}.$$

Since $|z| > n$ and $\rho\left(\frac{A}{N}\right) \leq n$, it is easy to see with the maximum column sum matrix norm that $\rho\left(\frac{A}{Nz}\right) < 1$, allowing us to apply Theorem 5.6.3. This gives

$$\begin{aligned}
\det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{P}{N}\right) &= 1 - \frac{\mu_N}{N} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(zI - \frac{A}{N}\right)^{-1}_{ij} \\
&= 1 - \frac{\mu_N}{Nz} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(I - \frac{A}{Nz}\right)^{-1}_{ij} \\
&= 1 - \frac{\mu_N}{Nz} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{k=0}^{\infty} \left(\frac{A}{Nz}\right)^{k}_{ij} \\
&= 1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^{k}_{ij},
\end{aligned}$$

where we have rewritten $\left(I - \frac{A}{Nz}\right)^{-1}_{ij}$ in terms of its Neumann series to go from line two to line three.

Let

$$f(z) = 1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^{k}_{ij}$$

and

$$g(z) = 1 - \frac{n\mu_N}{Nz} - \frac{\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}{N^2 z^2}.$$

Observe that $f(z)$ and $g(z)$ are both meromorphic functions, and that the zeros of $f(z)$ are eigenvalues of $(A+P)/N$. We will quickly find the zeros of $g(z)$. Suppose that $z \neq 0$. Then

$$z^2 g(z) = z^2 - \frac{n\mu_N}{N} z - \frac{\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}{N^2}. \tag{5.23}$$

By the quadratic formula, we can see that the zeros of $z^2 g(z)$, for $z \neq 0$, are given by

$$\frac{\mu_N n \pm \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N}.$$

Whenever N is sufficiently large, the distance between these two roots is

$$\left| \frac{\sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{N} \right| \geq \left| \frac{\sqrt{\mu_N^2 n^2 - 4n^2 \mu_N N}}{N} \right| = \left| \frac{\mu_N n}{N} \sqrt{1 - \frac{4N}{\mu_N}} \right|. \tag{5.24}$$

This tends to infinity as $N \to \infty$ since $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$. In particular, this means that any circle C of finite radius $\epsilon > 0$ centered around the root $\frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N}$ of $g(z)$ contains only that root, whenever N is sufficiently large.

Now, let $0 < \epsilon < 1$ and let C be the circle of radius ϵ centered at $\frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N}$. We will show that as $N \to \infty$,

$$|f(z) - g(z)| < |g(z)|$$

for all $z \in C$ with probability tending to 1.

Observe that

$$|z|^2 |f(z) - g(z)| = |z|^2 \left| \frac{\mu_N}{Nz} \sum_{k=2}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz} \right)^k_{ij} \right|$$

and

$$|z|^2 |g(z)|$$

$$= \left| z - \frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N} \right| \left| z - \frac{\mu_N n - \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N} \right|.$$

We will bound these quantities. Observe first that whenever N is sufficiently large,

$$\begin{aligned}\sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}} &\geq \sqrt{\mu_N^2 n^2 - 4n^2\mu_N N}\\ &= \mu_N n\sqrt{1-\frac{4N}{\mu_N}},\end{aligned}$$

which is real with probability tending to 1 as $N \to \infty$ since $\lim_{N\to\infty} \frac{N}{\mu_N} = 0$. For $z \in C$, we have

$$\begin{aligned}|z|^2|g(z)| &\geq |\epsilon|\left|\frac{\sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}}}{N} - \epsilon\right|\\ &> \left|\frac{\epsilon}{2}\right|\left|\frac{\sqrt{\mu_N^2 n^2 - 4\mu_N n^2 N}}{N} - \epsilon\right|\\ &\geq \left|\frac{\epsilon}{2}\right|\left|\frac{\mu_N n}{N}\sqrt{1-\frac{4N}{\mu_N}} - \epsilon\right|\\ &\geq \left|\frac{\epsilon}{2}\right|\left|\frac{\mu_N}{2N} - \epsilon\right|\\ &\geq \left|\frac{\epsilon}{2}\right||1-\epsilon|\end{aligned}$$

for N sufficiently large, since $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$.

By Lemma 5.6.4, $\sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}^k \leq n^{k+1}N^k$. For $k \geq 2$, we have

$$\begin{aligned}|z|^2\left|\frac{\mu_N}{Nz}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A}{Nz}\right)^k_{ij}\right| &\leq \left(\frac{|z|^2|\mu_N|}{N|z|}\right)\left(\frac{n^{k+1}N^k}{N^k|z|^k}\right)\\ &= \frac{n^{k+1}|\mu_N|}{N|z|^{k-1}}\\ &\leq \frac{n^{k+1}|\mu_N|}{N\left(\left|\frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}}}{2N}\right| - \epsilon\right)^{k-1}}\\ &\leq \frac{n^{k+1}}{\frac{N}{|\mu_N|}\left(\left|\frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}}}{2N}\right| - \epsilon\right)^{k-1}}\end{aligned}$$

which equals

$$n^2\left(\frac{n}{\left(\left|\frac{\mu_N^{(1-1/(k-1))}n + \mu_N^{(1/2-1/(k-1))}\sqrt{\mu_N n^2 + 4\sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}}}{2N^{(1-1/(k-1))}}\right| - \epsilon\left(\frac{N}{\mu_N}\right)^{1/(k-1)}\right)}\right)^{k-1}.$$

Since $-n^2N \leq \sum_{j=1}^n \sum_{i=1}^n A_{ij} \leq n^2N$ and $\lim_{N\to\infty} \frac{\mu_N}{N} = \infty$, we see that the term in the absolute value of the denominator in the expression above tends to infinity. Therefore, $k \geq 2$,

$$\lim_{N\to\infty} \frac{n}{\left(\left|\frac{\mu_N^{(1-1/(k-1))}n+\mu_N^{(1/2-1/(k-1))}\sqrt{\mu_N n^2+4\sum_{j=1}^n\sum_{i=1}^n A_{ij}}}{2N^{(1-1/(k-1))}}\right| - \epsilon\left(\frac{N}{\mu_N}\right)^{1/(k-1)}\right)} = 0.$$

Thus, we can choose N sufficiently large so that the value inside the limit is less than any $0 < \delta < 1$. Then

$$\begin{aligned} |z|^2|f(z) - g(z)| &= |z|^2 \left| \frac{\mu_N}{Nz} \sum_{k=2}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^k_{ij} \right| \\ &\leq \sum_{k=2}^{\infty} n^2 \delta^{k-1} \\ &= \frac{n^2\delta}{1-\delta}. \end{aligned}$$

Choosing δ small enough so that $\frac{n^2\delta}{1-\delta} < |\epsilon/2||1-\epsilon|$ shows that as $N \to \infty$,

$$|z|^2|f(z) - g(z)| < |z|^2|g(z)|$$

for all $z \in C$ with probability tending to 1, which proves by Theorem 2.6.5 that $f(z)$ and $g(z)$ both have exactly one root in C with probability tending to 1 as $N \to \infty$; in particular, this root is within ϵ of an eigenvalue of $(A+P)/N$.

Let us now argue that the eigenvalue we have found is $\lambda_{\max}$. From the discussion at the beginning of the proof, we see that for any eigenvalue λ of $\frac{A+P}{\mu_N}$, we have that $\frac{\mu_N \lambda}{N}$ is an eigenvalue of $\frac{A+P}{N}$. From Theorem 5.1.2, we know that all but one eigenvalue of $(A+P)/\mu_N$ converge almost surely to zero as $N \to \infty$. We may assume for that sufficiently large N, all of these non-outlier eigenvalues of $(A+P)/\mu_N$ are located in the circle of radius $n/4$ centered at the origin, since they converge to zero almost surely. This implies that for sufficiently large N, the eigenvalues of $\frac{A+P}{N}$ are all located in the circle of radius $\frac{\mu_N n}{4N}$, except for one. The eigenvalue that we found from Rouché's theorem is located within ϵ of

$$\frac{\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N}.$$

Hence,

$$\frac{\left|\mu_N n + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}\right|}{2N} \geq \frac{\left|\mu_N n \left(1 + \sqrt{1 - \frac{4N}{\mu_N}}\right)\right|}{2N}$$
$$\geq \frac{\mu_N n}{2N}$$
$$> \frac{\mu_N n}{4N}$$

for N sufficiently large, showing that this eigenvalue has magnitude larger than $\frac{\mu_N n}{4N}$, so this eigenvalue must be the outlier $\lambda_{\max}$.

Since ϵ was arbitrary, we have shown that

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{\lambda_{\max}}{N} - \frac{\mu_N + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N}\right| > \epsilon\right) = 0.$$

In particular, this implies that

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\left(\frac{\lambda_{\max}}{N} - \frac{n\mu_N}{N}\right) - \left(\frac{\mu_N + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N} - \frac{n\mu_N}{N}\right)\right| > \epsilon\right) = 0$$

as well. Next, observe that as $N \to \infty$,

$$\begin{aligned}
&\frac{\mu_N + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N} - \frac{n\mu_N}{N} \\
&= \frac{-n\mu_N + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{2N} \\
&= \frac{-4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}{2N\left(-n\mu_N - \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}\right)} \\
&= \frac{-4 \sum_{j=1}^n \sum_{i=1}^n A_{ij}}{2N\left(-n - \sqrt{n^2 + \frac{4}{\mu_N} \sum_{j=1}^n \sum_{i=1}^n A_{ij}}\right)} \\
&\overset{\mathrm{d}}{\Longrightarrow} \frac{\sum_{j=1}^n \sum_{i=1}^n X_{ij}}{n},
\end{aligned}$$

by Lemma 2.2.1, Theorem 2.2.3, and Theorem 2.2.5.

By part (4) of Theorem 2.2.6, we conclude that

$$\frac{\lambda_{\max} - n\mu_N}{N} \overset{\mathrm{d}}{\Longrightarrow} \frac{\sum_{j=1}^n \sum_{i=1}^n X_{ij}}{n}.$$

Applying the Theorem 2.2.3, we obtain the proper scaling for the Bates distribution, i.e.,

$$\frac{\lambda_{\max} - n\mu_N}{nN} \xrightarrow{\text{d}} \frac{\sum_{j=1}^{n}\sum_{i=1}^{n} X_{ij}}{n^2}.$$

□

Theorem 5.6.5. *Let A be an n-by-n random matrix with entries independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let $\lim_{N\to\infty} \frac{\mu_N}{N} > 12$ and let P be the perturbation matrix whose entries are all μ_N. Then as $N \to \infty$, $\frac{A+P}{N}$ contains exactly one real outlier eigenvalue in the interval $\left[\frac{\mu_N n}{2N}, \frac{\mu_N n(1+\sqrt{2})}{2N}\right]$ with probability tending to one and the remaining eigenvalues of $\frac{A+P}{N}$ are contained in $B(0, 4n)$ with probability tending to one.*

Theorem 5.6.5 is illustrated in Figure 5.5. Figure 5.6 shows what happens when $\lim_{N\to\infty} \frac{\mu_N}{N} = 3$. We conjecture that Theorem 5.6.5 holds whenever $\lim_{N\to\infty} \frac{\mu_N}{N} > 1$.

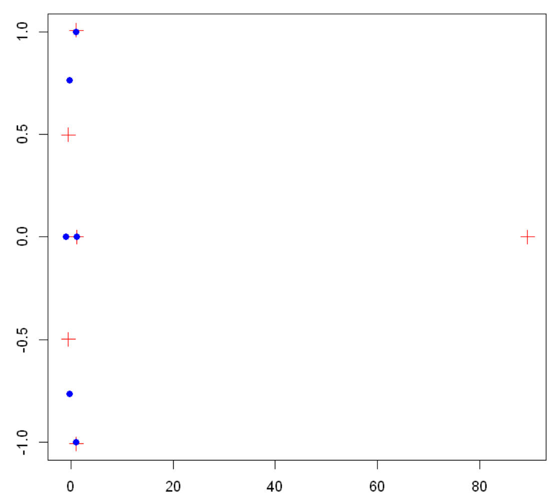

Figure 5.5: This figure illustrates Theorem 5.6.5. The blue circles denote the eigenvalues of a 6-by-6 random matrix with entries uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$, scaled by 10^{15}. The red crosses denote the eigenvalues of the same random matrix where every entry is perturbed by 15×10^{15}, and also scaled by 10^{15}

Proof. We proceed similarly to the proof of Theorem 5.1.8. Suppose that x is an eigenvector of

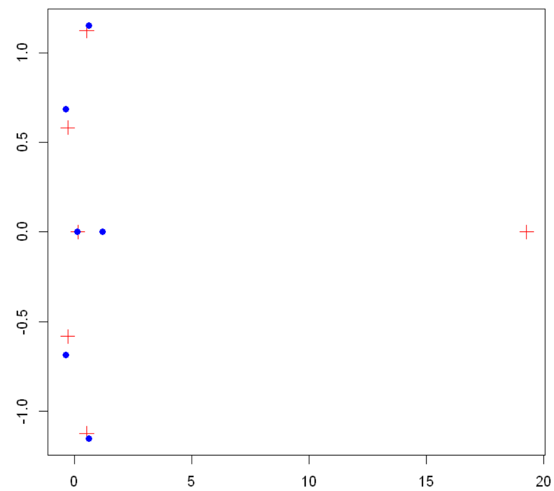

Figure 5.6: This figure illustrates an example not covered by Theorem 5.6.5. The blue circles denote the eigenvalues of a 6-by-6 random matrix with entries uniformly distributed on $[-10^{15}, 10^{15}] \cap \mathbb{Z}$, scaled by 10^{15}. The red crosses denote the eigenvalues of the same random matrix where every entry is perturbed by 3×10^{15}, and also scaled by 10^{15}.

$(A + P)$ with corresponding eigenvalue λ. We see that

$$\left(\frac{A+P}{N}\right) x = \left(\frac{\lambda}{N}\right) x,$$

so it suffices to consider the distribution of largest eigenvalue, in magnitude, of $(A + P)/N$. In terms of the characteristic polynomial, we wish to find the largest z such that

$$\det\left(zI - \frac{A+P}{N}\right) = 0.$$

Our first goal is to show that $(A+P)/N$ has only one outlier eigenvalue. Again, consider the factorization

$$\det\left(zI - \frac{A+P}{N}\right) = \det\left(zI - \frac{A}{N}\right) \det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{P}{N}\right).$$

From Lemma 5.3.12, we know that all of the eigenvalues of A/N (and hence zeros of $\det\left(zI - \frac{A}{N}\right)$) are located inside the disk of radius n centered at the origin. Therefore, any z such that $|z| > n$

that is an eigenvalue of $\frac{A+P}{N}$ (and hence a zero of $\det\left(zI - \frac{A+P}{N}\right)$) must also be a solution to

$$\det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{P}{N}\right) = 0.$$

We will first find the correct z that satisfies this equation, and then show that $|z| > \frac{\mu_N n}{2N}$ as $N \to \infty$. We will then show that the remaining eigenvalues of $\frac{A+P}{N}$ are located in $B(0, 4n)$ whenever N is sufficiently large, implying that the z found is the unique outlier eigenvalue. Now, supposing that $|z| > n$ and proceeding exactly as in the proof of Theorem 5.1.8, we can write this determinant in terms of its Neumann series to get

$$\det\left(I - \left(zI - \frac{A}{N}\right)^{-1} \frac{P}{N}\right) = 1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^k_{ij}.$$

Let

$$f(z) = 1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^k_{ij}$$

and

$$g(z) = 1 - \frac{n\mu_N}{Nz} - \frac{\mu_N \sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij}}{N^2 z^2}.$$

Observe that $f(z)$ and $g(z)$ are both meromorphic functions and when $|z| > n$, the zeros of $f(z)$ are eigenvalues of $(A+P)/N$.

Let $\epsilon = n$ and let C be the circle of radius ϵ centered at the root $\frac{n\mu_N + \sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij}}}{2N}$ of $g(z)$, which we found in the proof of Theorem 5.1.8. We will show that as $N \to \infty$

$$|f(z) - g(z)| < |g(z)|$$

for all $z \in C$ with probability tending to one. We will see momentarily that g has one root in C with probability tending to one as $N \to \infty$. Then Theorem 2.6.5 allows us to conclude that $f(z)$ and $g(z)$ both have precisely one zero in C, with probability tending to one as $N \to \infty$. Let us now investigate the roots of $g(z)$. As before, we have that

$$|z|^2 |f(z) - g(z)| = |z|^2 \left| \frac{\mu_N}{Nz} \sum_{k=2}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A}{Nz}\right)^k_{ij} \right|$$

and

$$\begin{aligned}&|z|^2|g(z)|\\&=\left|z-\frac{n\mu_N+\sqrt{\mu_N^2n^2+4\mu_N\sum_{j=1}^n\sum_{i=1}^n A_{ij}}}{2N}\right|\left|z-\frac{n\mu_N-\sqrt{\mu_N^2n^2+4\mu_N\sum_{j=1}^n\sum_{i=1}^n A_{ij}}}{2N}\right|.\end{aligned}$$

We begin by bounding the location of the roots of $z^2g(z)$, which are

$$\lambda_+=\frac{n\mu_N+\sqrt{\mu_N^2n^2+4\mu_N\sum_{j=1}^n\sum_{i=1}^n A_{ij}}}{2N}$$

and

$$\lambda_-=\frac{n\mu_N-\sqrt{\mu_N^2n^2+4\mu_N\sum_{j=1}^n\sum_{i=1}^n A_{ij}}}{2N}$$

by the quadratic formula.

Since $\lim_{N\to\infty}\frac{\mu_N}{N}>12$, for N sufficiently large, we may bound

$$\frac{\mu_N n}{2N}\le\frac{\mu_N n\left(1+\sqrt{1-\frac{4N}{\mu_N}}\right)}{2N}\le\lambda_+\le\frac{\mu_N n\left(1+\sqrt{1+\frac{4N}{\mu_N}}\right)}{2N}\le\frac{\mu_N n\left(1+\frac{2}{\sqrt{3}}\right)}{2N}$$

and

$$\frac{\mu_N n\left(1-\sqrt{1+\frac{4N}{\mu_N}}\right)}{2N}\le\lambda_-\le\frac{\mu_N n\left(1-\sqrt{1-\frac{4N}{\mu_N}}\right)}{2N}.$$

This leads to

$$\lambda_-\le\frac{\mu_N n\left(1-\sqrt{1-\frac{4N}{\mu_N}}\right)\left(1+\sqrt{1-\frac{4N}{\mu_N}}\right)}{2N\left(1+\sqrt{1-\frac{4N}{\mu_N}}\right)}\le\frac{6n}{3+\sqrt{6}}.$$

and

$$\lambda_-\ge\frac{\mu_N n\left(1-\sqrt{1+\frac{4N}{\mu_N}}\right)\left(1+\sqrt{1+\frac{4N}{\mu_N}}\right)}{2N\left(1+\sqrt{1+\frac{4N}{\mu_N}}\right)}\ge\frac{-2n}{\left(1+\sqrt{1+\frac{4N}{\mu_N}}\right)}\ge -n.$$

As $N\to\infty$, this shows that

$$\lambda_-\le\frac{6n}{3+\sqrt{6}}\le 2n<6n\le\lambda_+.$$

Hence, whenever N is sufficiently large, a circle of radius n around the root λ_+ of $g(z)$ contains only the root λ_+ of $g(z)$. For $z \in C$, we have

$$\begin{aligned}|z|^2|g(z)| &= |z-\lambda_+||z-\lambda_-| \\ &\geq |n|\left|\frac{\sqrt{\mu_N^2 n^2 + 4\mu_N \sum_{j=1}^n \sum_{i=1}^n A_{ij}}}{N} - n\right| \\ &\geq n\left|\frac{\mu_N n\sqrt{1-\frac{4N}{\mu_N}}}{N} - n\right|.\end{aligned}$$

Since $c > 12$, whenever N is sufficiently large, we have that

$$|z|^2|g(z)| > \frac{\mu_N n^2\sqrt{\frac{2}{3}}}{N} - n^2.$$

By a similar calculation, we see that for $z \in C$, whenever N is sufficiently large,

$$|z| \geq \frac{\mu_N n}{2N}\left(1+\sqrt{\frac{2}{3}}\right) - n \geq \frac{\mu_N n}{2N} - n.$$

On the other hand, by Lemma 5.6.4, $\sum_{j=1}^n \sum_{i=1}^n A_{ij}^k \leq n^{k+1}N^k$, so for $k \geq 2$, whenever N is sufficiently large, we may bound

$$\begin{aligned}|z|^2\left|\frac{\mu_N}{Nz}\sum_{j=1}^n\sum_{i=1}^n\left(\frac{A}{Nz}\right)^k_{ij}\right| &\leq \left(\frac{|z|^2|\mu_N|}{N|z|}\right)\left(\frac{n^{k+1}N^k}{N^k|z|^k}\right) \\ &= \frac{n^{k+1}|\mu_N|}{N|z|^{k-1}} \\ &\leq \frac{n^{k+1}|\mu_N|}{N\left(\frac{|\mu_N|n}{2N} - n\right)^{k-1}} \\ &\leq \frac{n^2|\mu_N|}{N}\left(\frac{n}{\frac{|\mu_N|n}{2N} - n}\right)^{k-1} \\ &\leq \frac{n^2|\mu_N|}{N}\left(\frac{1}{\frac{|\mu_N|}{2N} - 1}\right)^{k-1}.\end{aligned}$$

Summing from $2 \leq k < \infty$, whenever N is sufficiently large, we have

$$\begin{aligned}|z|^2|f(z)-g(z)| &= |z|^2\left|\frac{\mu_N}{Nz}\sum_{k=2}^\infty\sum_{j=1}^n\sum_{i=1}^n\left(\frac{A}{Nz}\right)^k_{ij}\right| \\ &\leq \frac{n^2|\mu_N|}{N}\sum_{k=2}^\infty\left(\frac{1}{\frac{|\mu_N|}{2N} - 1}\right)^{k-1} \\ &= \frac{n^2|\mu_N|}{N}\left(\frac{2}{\frac{|\mu_N|}{N} - 3}\right).\end{aligned}$$

We see that

$$\lim_{N\to\infty} \frac{n^2|\mu_N|}{N}\left(\frac{2}{\frac{|\mu_N|}{N}-3}\right) < \lim_{N\to\infty} \frac{\mu_N n^2\sqrt{2/3}}{N} - n^2$$

whenever $n > 0$ and $\lim_{N\to\infty} \frac{\mu_N}{N} = c > \frac{(6+3\sqrt{6}+\sqrt{90+12\sqrt{6}})}{4} \approx 6.0688$. Since we assumed $\lim_{N\to\infty} \frac{\mu_N}{N} > 12$, we have that as $N \to \infty$,

$$|z|^2|f(z) - g(z)| < |z|^2|g(z)|$$

for all $z \in C$ with probability tending to one. By Theorem 2.6.5, this proves that as $N \to \infty$, $f(z)$ and $g(z)$ both have exactly one root in C with probability tending to one and this root is within n of an eigenvalue of $(A + P)/N$,

We note that at this point that we could have chosen the circle of radius $n/3$ around λ_+, rather than the circle of radius n, and the inequalities would have still worked out in the end. However, for a circle of radius smaller than $n/4$, we would have had to increase c. For instance, the circle of radius $n/4$ requires that $c > 12.4526$ and the circle of radius $n/100$ requires that $c > 204.02$.

We will now locate the other eigenvalues of $(A+P)/N$. We will use the earlier bounds for λ_+ and λ_- given by $|\lambda_+| \geq \frac{\mu_N n}{2N}$ and$|\lambda_-| \leq \frac{6n}{3+\sqrt{6}} < 2n$ whenver N is sufficiently large, while keeping in mind that λ_+ and λ_- are both real with probability tending to 1 as $N \to \infty$. As before, let

$$f(z) = 1 - \frac{\mu_N}{Nz}\sum_{k=0}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A}{Nz}\right)^k_{ij}$$

and

$$g(z) = 1 - \frac{n\mu_N}{Nz} - \frac{\mu_N \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}}{N^2 z^2}$$

and consider the functions $|z|^2|f(z) - g(z)|$ and $|z|^2|g(z)|$. Similarly to before, since $|z| > n$, we may bound

$$|z|^2|f(z)-g(z)| = |z|^2\left|\frac{\mu_N}{Nz}\sum_{k=2}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A}{Nz}\right)^k_{ij}\right|$$
$$\leq \sum_{k=2}^{\infty}\frac{n^{k+1}|\mu_N|}{N|z|^{k-1}}$$
$$\leq \frac{n^2|\mu_N|}{N}\sum_{k=2}^{\infty}\left(\frac{n}{|z|}\right)^{k-1}$$
$$\leq \frac{n^3|\mu_N|}{N||z|-n|}.$$

Hence, we must find the values for z so that whenever N is sufficiently large,

$$\frac{n^3\mu_N}{N||z|-n|} < |z-\lambda_-|\,|z-\lambda_+|\,. \tag{5.25}$$

Let R be the rectangle of height $4n$ centered on the real axis with a lower bound for its real coordinates given by $\frac{\mu_N n}{2N}-2n$ and upper bound for its real coordinates given by $\frac{\mu_N n(1+\sqrt{2})}{2N}+2n$. Let $\mathring{R}$ denote all points enclosed by R. First, consider any $z\in\mathbb{C}$ such that $|z|\geq 4n$ and z is outside of the interior of R, i.e., $z\in\mathbb{C}\setminus(\mathring{R}\cup(B(0,4n))$.

Since $|z|\geq 4n$ for all $z\in\mathbb{C}\setminus(\mathring{R}\cup(B(0,4n))$ whenever N sufficiently large, we have that

$$\frac{n^3\mu_N}{N||z|-n|}\leq\frac{n^3\mu_N}{N||4n|-n|}=\frac{\mu_N n^2}{3N}.$$

Let $z=x+iy$ be the decomposition of z into its real and imaginary parts. Whenever λ_+ and λ_- are real, which happens with probability tending to 1 as $N\to\infty$, we see that

$$|z-\lambda_+|=\sqrt{(x-\lambda_+)^2+y^2}\geq|x-\lambda_+|$$

and similarly

$$|z-\lambda_-|\geq\sqrt{(x-\lambda_-)^2+y^2}\geq|x-\lambda_-|.$$

We now consider four different locations for $z\in\mathbb{C}\setminus(\mathring{R}\cup B(0,4n))$.

Observe that whenever N is sufficiently large, $\frac{\mu_N n}{2N}-2n>4n$, since $\lim_{N\to\infty}\frac{\mu_N}{N}=c>12$. If $|z|\geq 4n$ and $4n\leq x\leq\frac{\mu_N n}{2N}-2n$, then since $\lambda_-\leq 2n$ and $\lambda_+\geq\frac{\mu_N n}{2N}$ whenever N is sufficiently large, we wish to minimize

$$\begin{aligned} |z-\lambda_-|\,|z-\lambda_+| &\geq |x-\lambda_-|\,|x-\lambda_+| \\ &= (x-\lambda_-)(\lambda_+-x) \\ &\geq (x-2n)\left(\frac{\mu_N n}{2N}-x\right). \end{aligned}$$

From elementary calculus, it is clear that the concave down parabola in the variable x from the third line above is minimized at one of the endpoints for $x \in \left[4n, \frac{\mu_N n}{2N} - 2n\right]$. In this case, both endpoints yield the same value, and we obtain the bound

$$|z-\lambda_-|\,|z-\lambda_+| \geq (2n)\left(\frac{\mu_N n}{2N}-4n\right),$$

whenever N is sufficiently large.

If $|z| \geq 4n$ and $-\infty < x < 4n$, then we see that the distances are minimized when $z = 4n$, and we have

$$\begin{aligned} |z-\lambda_-|\,|z-\lambda_+| &\geq |x-\lambda_-|\,|x-\lambda_+| \\ &\geq (2n)\left(\frac{\mu_N n}{2N}-4n\right) \end{aligned}$$

whenever N is sufficiently large.

If $|z| \geq 4n$ and $\frac{\mu_N n}{2N} - 2n < x \leq \frac{\mu_N n(1+\sqrt{2})}{2N} + 2n$ and $|y| \geq 2n$, we have

$$\begin{aligned} |z-\lambda_-|\,|z-\lambda_+| &\geq |x-\lambda_-||y-\lambda_+| \\ &\geq \left(\frac{\mu_N n}{2N}-4n\right)(2n) \end{aligned}$$

whenever N is sufficiently large.

Finally, if $|z| \geq 4n$ and $\frac{\mu_N n(1+\sqrt{2})}{2N} + 2n < x < \infty$,

$$\begin{aligned} |z-\lambda_-|\,|z-\lambda_+| &\geq |x-\lambda_-|\,|x-\lambda_+| \\ &\geq \left(\frac{\mu_N n(1+\sqrt{2})}{2N}\right)\left(\frac{\mu_N n}{2N}\left(\sqrt{2}-2/\sqrt{3}\right)\right) \\ &\geq \left(\frac{\mu_N n(1+\sqrt{2})}{2N}\right)(2n) \\ &\geq \left(\frac{\mu_N n}{2N}-4n\right)(2n) \end{aligned}$$

whenever N is sufficiently large.

In all four cases, we arrive at

$$|z-\lambda_-||z-\lambda_+| \geq \left(\frac{\mu_N n}{2N} - 4n\right)(2n)$$

whenever N is sufficiently large.

Observe that

$$\frac{\mu_N n^2}{3N} < \left(\frac{\mu_N n}{2N} - 4n\right)(2n)$$

whenever $\frac{\mu_N}{N} > 12$, which occurs whenever N is sufficiently large.

Let Γ by any circle in the complex plane contained in $\mathbb{C}\setminus(\mathring{R}\cup B(0,4n))$. We have shown that

$$|z|^2|f(z)-g(z)| < |z|^2|g(z)|$$

for all $z \in \Gamma$ as $N \to \infty$ with probability tending to 1. By Theorem 2.6.5, we conclude that as $N \to \infty$, $z^2 f(z)$ and $z^2 g(z)$ have the same number of roots in Γ with probability tending to one; this must be exactly zero roots, since we have avoided the roots of $z^2 g(z)$. By varying the circle Γ, we can include every point in $\mathbb{C}\setminus(\mathring{R}\cup(B(0,4n))$. This implies that $z^2 f(z)$, and hence $f(z)$, cannot have any zeros in these regions, with probability tending to 1 as $N \to \infty$. Therefore, all of the zeros of $f(z)$ must be in $\overline{B(0,4n)}$ or $\overline{R}$.

Next, we show that the zeros of $f(z)$ cannot be in $\overline{R}$. Let $\epsilon > 0$ be small. Let W be the rectangle of height $4n+2\epsilon$ centered on the real axis whose real coordinates have lower bound $\frac{\mu_N n}{2N} - 2n - \epsilon$ and whose real coordinates have upper bound $\frac{\mu_N n(1+\sqrt{2})}{2N} + 2n + \epsilon$. Note that every side of W is at least length ϵ away from the corresponding side of R; clearly $\bar{R} \subset \bar{W}$. For $z \in W$,

$$|z| \geq \left|\frac{\mu_N n}{2N} - 2n - \epsilon\right|,$$

and so

$$\frac{n^3\mu_N}{N||z|-n|} \leq \frac{n^3\mu_N}{N\left|\left|\frac{\mu_N n}{2N} - 2n - \epsilon\right| - n\right|} \leq \frac{n^3\mu_N}{N\left(\frac{\mu_N n}{2N} - 3n - \epsilon\right)}$$

whenever N is sufficiently large. On the other hand, W was chosen in such a way so that $|z-\lambda_+| \geq \frac{\mu_N n\left(\sqrt{2}-2/\sqrt{3}\right)}{2N} \geq 2n+\epsilon$ and $|z-\lambda_-| \geq \frac{\mu_N n}{2N} - 4n - \epsilon$ for all $z \in R$ whenever N is sufficiently large.

Therefore, we wish to show

$$\frac{\mu_N n^3}{N\left(\frac{\mu_N n}{2N} - 3n - \epsilon\right)} < (2n+\epsilon)\left(\frac{\mu_N n}{2N} - 4n - \epsilon\right),$$

with probability tending to 1 as $N \to \infty$. Letting $\epsilon \to 0$, the inequality holds whenever $\frac{\mu_N}{N} > 12$, which happens when N is sufficiently large. By Theorem 2.6.5, we conclude that $z^2 f(z)$ and $z^2 g(z)$ both contain exactly one root in W with probability tending to one as $N \to \infty$; these roots must be the real roots already found on the real line. Putting everything together, we see that the remaining roots of $z^2 f(z)$, and hence $f(z)$, must be contained in $\overline{B(0, 4n)}$, with probability tending to 1 as $N \to \infty$. □

Lemma 5.6.6. *Suppose that the n-by-n random matrices A and B with all entries having modulus at most one have been coupled in such a way that $|A_{ij} - B_{ij}| < \frac{C}{N}$ for all $1 \le i, j \le n$, where $C > 0$ is a constant. Then for any $k \in \mathbb{N}$,*

$$\left|A^k_{ij} - B^k_{ij}\right| \le \frac{kn^k C}{N}.$$

Proof. Consider the matrix factorization for the difference of powers

$$A^k - B^k = \sum_{i=0}^{k-1} A^i (A - B) B^{k-1-i},$$

which can be proven by expanding the sum and noticing that it is a telescoping series. Furthermore, observe that the ij entry of any matrix is bounded above by the maximum column sum norm. Applying the (submultiplicative) maximum column sum norm, we have that

$$\|A^k - B^k\|_1 \le \sum_{i=0}^{k-1} \|A^i\|_1 \|A - B\|_1 \|B^{k-1-i}\|_1.$$

We may bound

$$\|A^i\|_1 \le \|A\|_1^i \le n^i, \quad \|B^{k-1-i}\|_1 \le \|B\|_1^{k-1-i} \le n^{k-1-i},$$

since all matrix entries of A and B have modulus at most one. Moreover,

$$\|A - B\|_1 \le \frac{nC}{N}$$

by the assumed coupling. Thus,

$$\begin{aligned}\|A^k - B^k\|_1 &\leq \sum_{i=0}^{k-1} n^i \left(\frac{nC}{N}\right) n^{k-1-i} \\ &= \left(\frac{nC}{N}\right) \sum_{i=0}^{k-1} n^{k-1} \\ &= \frac{kn^k C}{N}.\end{aligned}$$

Since

$$\left|A_{ij}^k - B_{ij}^k\right| = \left|(A^k - B^k)_{ij}\right| \leq \|A^k - B^k\|_1,$$

the result follows. □

Theorem 5.6.7. *Let A be an n-by-n random matrix with entries independently and uniformly distributed on $[-N, N] \cap \mathbb{Z}$. Let $\lim_{N\to\infty} \frac{\mu_N}{N} = c > 12$ and let P be the perturbation matrix whose entries are all μ_N. Let B be an n-by-n random matrix with entries independently and uniformly distributed on $(-1, 1)$. Let $\lambda_{\max}$ denote the largest eigenvalue, in magnitude, of $\frac{A+P}{N}$. Let Z be the random variable that is the unique solution to the equation*

$$\sum_{j=1}^{n} \sum_{i=1}^{n} (zI - B)_{ij}^{-1} = \frac{1}{c}.$$

outside of $B(0, 4n)$. Then

$$\lambda_{\max} \overset{d}{\Rightarrow} Z.$$

Proof. Let A and B be coupled as in Lemma 5.3.5; Lemma 3.3.2 and Remark 3.3.3 show that this is indeed a coupling of the matrices A and B. Refer to the coupled matrices as A' and B'. In Theorem 5.6.5, it was shown that $\frac{A+P}{N}$ (and hence $\frac{A'+P}{N}$) has a unique real outlier eigenvalue contained in the interval $\left[\frac{\mu_N n}{2N}, \frac{\mu_N n(1+\sqrt{2})}{2N}\right]$ with probability tending to one as $N \to \infty$, and that all other eigenvalues of $\frac{A+P}{N}$ (and hence $\frac{A'+P}{N}$) are contained in $B(0, 4n)$ with probability tending to one as $N \to \infty$; call the outlier eigenvalue $\lambda_{\max}$. Let C be the perturbation matrix whose entries are all c. We will now compare the eigenvalues of $\frac{A'+P}{N}$ and the eigenvalues of $B' + C$ via the factorizations

$$\det\left(zI - \frac{A' + P}{N}\right) = \det\left(zI - \frac{A'}{N}\right) \det\left(I - \left(zI - \frac{A'}{N}\right)^{-1} \frac{P}{N}\right)$$

and

$$\det\left(zI - \left(B' + C\right)\right) = \det\left(zI - B'\right) \det\left(I - \left(zI - B'\right)^{-1} C\right).$$

Since we are interested in eigenvalues near $\lambda_{\max}$, consider $|z| > 4n$. Expanding in terms of a Neumann series as in the proof of Theorem 5.1.8, let

$$\begin{aligned} f(z) &= \det\left(I - \left(zI - \frac{A'}{N}\right)^{-1} \frac{P}{N}\right) \\ &= 1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \end{aligned}$$

and

$$\begin{aligned} g(z) &= \det\left(I - \left(zI - B'\right)^{-1} C\right) \\ &= 1 - \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{B'}{z}\right)_{ij}^{k}. \end{aligned}$$

Then for $|z| > 4n$ and N sufficiently large, we have

$$\begin{aligned} |f(z) - g(z)| &= \left| \left(1 - \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k}\right) - \left(1 - \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{B'}{z}\right)_{ij}^{k}\right) \right| \\ &= \left| \left(\frac{c}{z} - \frac{\mu_N}{Nz}\right) \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} + \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\left(\frac{B'}{z}\right)_{ij}^{k} - \left(\frac{A'}{Nz}\right)_{ij}^{k}\right) \right| \\ &\le \left| \left(c - \frac{\mu_N}{N}\right) \frac{1}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \right| + \left| \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\left(\frac{B'}{z}\right)_{ij}^{k} - \left(\frac{A'}{Nz}\right)_{ij}^{k}\right) \right| \\ &\le \left| \left(c - \frac{\mu_N}{N}\right) \frac{1}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \right| + \left| \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{kn^k M}{z^k N} \right| \\ &\le \left| \left(c - \frac{\mu_N}{N}\right) \frac{1}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \right| + \left| \frac{cnM}{N} \sum_{k=0}^{\infty} k \left(\frac{n}{z}\right)^{k+1} \right| \\ &\le \left| \left(c - \frac{\mu_N}{N}\right) \frac{1}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \right| + \left| \frac{cn^3 M}{N(n-z)^2} \right|, \end{aligned} \tag{5.26}$$

where we have applied Lemma 5.6.6 to go from line three to line four; here, $M > 0$ is the constant satisfying $|A'_{ij}/N - B'_{ij}| < M/N$ for each $1 \le i, j \le n$, which exists due to Lemma 5.3.5. We also have that

$$|f(z)| = \left| \frac{\mu_N}{Nz} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{A'}{Nz}\right)_{ij}^{k} \right|. \tag{5.27}$$

Since $A^0 = I$,

$$\left|\frac{\mu_N}{Nz}\sum_{k=0}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A'}{Nz}\right)^k_{ij}\right| = \left|\left(\frac{\mu_N}{Nz}\right)\left(n+\sum_{k=1}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A'}{Nz}\right)^k_{ij}\right)\right|$$
$$\geq \left|\frac{\mu_N}{Nz}\right|\left|n-\left|\sum_{k=1}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A'}{Nz}\right)^k_{ij}\right|\right|,$$

where the second line follows by the reverse triangle inequality. By arguments similar to Lemma 5.6.4, we can bound the sum

$$\left|\sum_{k=1}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A'}{Nz}\right)^k_{ij}\right| \leq \sum_{k=1}^{\infty}\frac{n^{k+1}N^k}{N^k|z|^k} \leq \frac{n}{3}$$

since $|z| > 4n$. Hence, a lower bound for (5.27) is given by

$$\left|\frac{\mu_N}{Nz}\right|(n-n/3) \geq \left|\frac{2n\mu_N}{3Nz}\right|.$$

Now let $\epsilon > 0$ and consider $z \in C_{\lambda_{\max},\epsilon}$. Let us compare (5.26) and (5.27). Since $\lim_{N\to\infty}\frac{\mu_N}{N} = c$ and $|z| > 4n$, using Lemma 5.6.4 and the geometric series, we have

$$\left|\left(c-\frac{\mu_N}{N}\right)\frac{1}{z}\sum_{k=0}^{\infty}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{A'}{Nz}\right)^k_{ij}\right| \leq \left|c-\frac{\mu_N}{N}\right|\left|\frac{1}{z}\right|\left|\sum_{k=0}^{\infty}\frac{n^{k+1}N^k}{N^kz^k}\right|$$
$$\leq \left|c-\frac{\mu_N}{N}\right|\left|\frac{1}{z}\right|\left|\frac{nz}{z-n}\right|$$
$$< \left|\frac{1}{3}\left(c-\frac{\mu_N}{N}\right)\right|.$$

Since $|z| > 4n$, we can bound

$$\left|\frac{cn^3M}{N(n-z)^2}\right| < \left|\frac{cnM}{3N}\right|.$$

Combining these pieces, since

$$\lim_{N\to\infty}\left|\frac{1}{3}\left(c-\frac{\mu_N}{N}\right)\right| + \left|\frac{cnM}{3N}\right| = 0,$$

we see that

$$\mathbb{P}\left(\limsup_{N\to\infty}|f(z)-g(z)| < |f(z)|,\ \text{for all } z \in C_{\lambda_{\max},\epsilon}\right) = 1.$$

By Theorem 2.6.5, we see that both $f(z)$ and $g(z)$ have the same number of zeros in $C_{\lambda_{\max},\epsilon}$ with probability tending to one, as $N \to \infty$. To see that this zero is unique, observe first that $\lambda_{\max}$

must be a zero of $f(z)$. On the other hand, any zero of $f(z)$ is also an eigenvalue of $(A'+P)/N$, and we know from Theorem 5.6.5 that the remaining eigenvalues are contained in $B(0,4n)$ with probability tending to one as $N \to \infty$. Since

$$1 - \frac{c}{z} \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{B'}{z} \right)_{ij}^{k} = 1 - c \sum_{j=1}^{n} \sum_{i=1}^{n} \left(zI - B' \right)_{ij}^{-1},$$

let Z be the random variable that is the unique solution to the equation

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \left(zI - B' \right)_{ij}^{-1} = \frac{1}{c}$$

outside of $B(0,4n)$. To see that this solution actually exists and is unique, one can repeat the proof of Theorem 5.6.5, but applied to the matrix $B+C$ rather than $(A+P)/N$, and replacing μ_N/N by c. Since $\lim_{N\to\infty} \frac{\mu_N}{N} = c$, the bounds and algebra involved are very similar. Then we have showed that

$$\lim_{N\to\infty} \mathbb{P}\left(|\lambda_{\max} - Z| < 2\epsilon \right) = 0,$$

which implies that

$$\lambda_{\max} \overset{\mathrm{d}}{\Rightarrow} Z.$$

In particular, since $A \sim A'$ and $B \sim B'$, the result holds. $\square$

Bibliography

[1] Gernot Akemann and Eugene Kanzieper. Integrable structure of Ginibre's ensemble of real random matrices and a Pfaffian integration theorem. J. Stat. Phys., 129(5-6):1159–1231, 2007.

[2] Shigeki Akiyama and Attila Pethő. On the distribution of polynomials with bounded roots, I. Polynomials with real coefficients. J. Math. Soc. Japan, 66(3):927–949, 2014.

[3] Alkiviadis G. Akritas, Evgenia K. Akritas, and Genadii I. Malaschonok. Various proofs of Sylvester's (determinant) identity. Mathematics and Computers in Simulation, 42(4):585–593, 1996. Sybolic Computation, New Trends and Developments.

[4] Theresa C. Anderson, Ayla Gafni, Robert J. Lemke Oliver, David Lowry-Duda, George Shakan, and Ruixiang Zhang. Quantitative Hilbert irreducibility and almost prime values of polynomial discriminants. Int. Math. Res. Not. IMRN, 2023(3):2188–2214, 2021.

[5] Jürgen Angst, Thibault Pautrel, and Guillaume Poly. Real zeros of random trigonometric polynomials with dependent coefficients. Trans. Amer. Math. Soc., 375(10):7209–7260, 2022.

[6] Sheldon Axler. Linear algebra done right. Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.

[7] Z. D. Bai. Circular law. Ann. Probab., 25(1):494–529, 1997.

[8] Zhidong Bai and Jack W. Silverstein. Spectral analysis of large dimensional random matrices. Springer Series in Statistics. Springer, New York, second edition, 2010.

[9] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. Ann. Probab., 33(5):1643–1697, 2005.

[10] Debapratim Banerjee, Soumendu Sundar Mukherjee, and Dipranjan Pal. Edge spectra of gaussian random symmetric matrices with correlated entries. Available at arXiv:2409.11381, 2024.

[11] Lior Bary-Soroker, Or Ben-Porath, and Vlad Matei. Probabilistic galois theory – the square discriminant case. Available at arXiv:2207.12493v2, 2024.

[12] Grace E. Bates. Joint distributions of time intervals for the occurrence of successive accidents in a generalized Pólya scheme. Ann. Math. Statist., 26:705–720, 1955.

[13] Serban Belinschi, Charles Bordenave, Mireille Capitaine, and Guillaume Cébron. Outlier eigenvalues for non-Hermitian polynomials in independent i.i.d. matrices and deterministic matrices. Electron. J. Probab., 26:Paper No. 100, 37, 2021.

[14] F. Benaych-Georges, A. Guionnet, and M. Maida. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. Electron. J. Probab., 16:no. 60, 1621–1662, 2011.

[15] F. Benaych-Georges, A. Guionnet, and M. Maida. Large deviations of the extreme eigenvalues of random deformations of matrices. Probab. Theory Related Fields, 154(3-4):703–751, 2012.

[16] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. Adv. Math., 227(1):494–521, 2011.

[17] Florent Benaych-Georges and Jean Rochet. Outliers in the single ring theorem. Probab. Theory Related Fields, 165(1-2):313–363, 2016.

[18] Csanád Bertók, Lajos Hajdu, and Attila Pethő. On the distribution of polynomials with bounded height. J. Number Theory, 179:172–184, 2017.

[19] Manjul Bhargava. A proof of van der Waerden's conjecture on random Galois groups of polynomials. Pure Appl. Math. Q., 19(1):45–60, 2023.

[20] Manjul Bhargava, Arul Shankar, and Xiaoheng Wang. Squarefree values of polynomial discriminants I. Invent. Math., 228(3):1037–1073, 2022.

[21] A. T. Bharucha-Reid and M. Sambandham. Random polynomials. Probability and Mathematical Statistics. Academic Press, Inc., Orlando, FL, 1986.

[22] Patrick Billingsley. Probability and measure. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[23] Patrick Billingsley. Convergence of probability measures. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[24] Dario Andrea Bini. Numerical computation of polynomial zeros by means of Aberth's method. Numer. Algorithms, 13(3-4):179–200, 1996.

[25] A. Bloch and G. Pólya. On the roots of certain algebraic equations. Proceedings of the London Mathematical Society, s2-33(1):102–114, 1932.

[26] Christian Borst, Evan Boyd, Claire Brekken, Samantha Solberg, Melanie Matchett Wood, and Philip Matchett Wood. Irreducibility of random polynomials. Exp. Math., 27(4):498–506, 2018.

[27] Jean Bourgain, Van H. Vu, and Philip Matchett Wood. On the singularity probability of discrete random matrices. J. Funct. Anal., 258(2):559–603, 2010.

[28] Tatiana Brailovskaya and Ramon van Handel. Extremal random matrices with independent entries and matrix superconcentration inequalities. Available at arXiv:2401.06284, 2024.

[29] M. Capitaine, C. Donati-Martin, D. Féral, and M. Février. Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices. Electron. J. Probab., 16:no. 64, 1750–1792, 2011.

[30] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. Ann. Probab., 37(1):1–47, 2009.

[31] R. Chela. Reducible polynomials. J. London Math. Soc., 38:183–188, 1963.

[32] Sam Chow and Rainer Dietmann. Enumerative Galois theory for cubics and quartics. Adv. Math., 372:107282, 37, 2020.

[33] Sam Chow and Rainer Dietmann. Towards van der Waerden's conjecture. Trans. Amer. Math. Soc., 376(4):2739–2785, 2023.

[34] S. D. Cohen. The distribution of the Galois groups of integral polynomials. Illinois J. Math., 23(1):135–152, 1979.

[35] Natalie Coston, Sean O'Rourke, and Philip Matchett Wood. Outliers in the spectrum for products of independent random matrices. Ann. Inst. Henri Poincaré Probab. Stat., 56(2):1284–1320, 2020.

[36] Natalie A. Coston. Spectral properties of products of independent random matrices. Doctoral dissertation, University of Colorado, 2018.

[37] J. H. Curtiss. On the distribution of the quotient of two chance variables. Ann. Math. Statistics, 12:409–421, 1941.

[38] Amir Dembo, Bjorn Poonen, Qi-Man Shao, and Ofer Zeitouni. Random polynomials having few or no real zeros. J. Amer. Math. Soc., 15(4):857–892, 2002.

[39] Rainer Dietmann. Probabilistic Galois theory. Bull. Lond. Math. Soc., 45(3):453–462, 2013.

[40] Xiucai Ding and Hong Chang Ji. Spiked multiplicative random matrices and principal components. Stochastic Process. Appl., 163:25–60, 2023.

[41] Yen Do, Hoi Nguyen, and Van Vu. Real roots of random polynomials: expectation and repulsion. Proc. Lond. Math. Soc. (3), 111(6):1231–1260, 2015.

[42] Yen Do, Hoi H. Nguyen, and Oanh Nguyen. Random trigonometric polynomials: universality and non-universality of the variance for the number of real roots. Ann. Inst. Henri Poincaré Probab. Stat., 58(3):1460–1504, 2022.

[43] K. Dörge. Abschätzung der Anzahl der reduziblen Polynome. Math. Ann., 160:59–63, 1965.

[44] Artūras Dubickas. On the number of monic integer polynomials with given signature. Arch. Math. (Basel), 110(4):333–342, 2018.

[45] Artūras Dubickas and Min Sha. Counting and testing dominant polynomials. Exp. Math., 24(3):312–325, 2015.

[46] Artūras Dubickas and Min Sha. Positive density of integer polynomials with some prescribed properties. J. Number Theory, 159:27–44, 2016.

[47] D.S. Dummit and R.M. Foote. Abstract Algebra. Wiley, 2003.

[48] Rick Durrett. Probability—theory and examples, volume 49 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fifth edition, 2019.

[49] F. J. Dyson. A brownian motion model for the eigenvalues of a random matrix. J. Math. Phys., 3:1191, 1962.

[50] F. J. Dyson. Statistical theory of the energy levels of complex systems, i. J. Math. Phys., 3:140, 1962.

[51] F. J. Dyson. Statistical theory of the energy levels of complex systems, ii. J. Math. Phys., 3:157, 1962.

[52] F. J. Dyson. Statistical theory of the energy levels of complex systems, iii. J. Math. Phys., 3:166, 1962.

[53] F. J. Dyson. The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics. J. Math. Phys., 3:1200, 1962.

[54] Alan Edelman. The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. J. Multivariate Anal., 60(2):203–232, 1997.

[55] Alan Edelman and Eric Kostlan. How many zeros of a random polynomial are real? Bull. Amer. Math. Soc. (N.S.), 32(1):1–37, 1995.

[56] Alan Edelman, Eric Kostlan, and Michael Shub. How many eigenvalues of a random matrix are real? J. Amer. Math. Soc., 7(1):247–267, 1994.

[57] Tamás Erdélyi. Extensions of the Bloch-Pólya theorem on the number of real zeros of polynomials. J. Théor. Nombres Bordeaux, 20(2):281–287, 2008.

[58] P. Erdös and P. Turán. On the distribution of roots of polynomials. Ann. of Math. (2), 51:105–119, 1950.

[59] Paul Erdös and A. C. Offord. On the number of real roots of a random algebraic equation. Proc. London Math. Soc. (3), 6:139–160, 1956.

[60] E. A. Evans. On the number of real roots of a random algebraic equation. Proc. London Math. Soc. (3), 15:731–749, 1965.

[61] Kambiz Farahmand. Topics in random polynomials, volume 393 of Pitman Research Notes in Mathematics Series. Longman, Harlow, 1998.

[62] Delphine Féral and Sandrine Péché. The largest eigenvalue of rank one deformation of large Wigner matrices. Comm. Math. Phys., 272(1):185–228, 2007.

[63] P. J. Forrester, N. C. Snaith, and J. J. M. Verbaarschot. Developments in random matrix theory. J. Phys. A, 36(12):R1–R10, 2003.

[64] Peter J. Forrester and S. Ole Warnaar. The importance of the selberg integral. Bulletin of the American Mathematical Society, 45:489–489, 2008.

[65] P.J. Forrester. Log-Gases and Random Matrices (LMS-34). Princeton University Press, 2010.

[66] Michael Fried. On Hilbert's irreducibility theorem. J. Number Theory, 6:211–231, 1974.

[67] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. Combinatorica, 1(3):233–241, 1981.

[68] P. X. Gallagher. The large sieve and probabilistic Galois theory. In Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), volume Vol. XXIV of Proc. Sympos. Pure Math., pages 91–101. Amer. Math. Soc., Providence, RI, 1973.

[69] Michel Gaudin. Sur la loi limite de l'espacement des valeurs propres d'une matrice ale ́atoire. Nuclear Physics, 25:447–458, 1961.

[70] Promit Ghosal and Sumit Mukherjee. Universality of persistence of random polynomials. Available at arXiv: 1208.2382, 2024.

[71] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. J. Mathematical Phys., 6:440–449, 1965.

[72] V. L. Girko. The circular law. Teor. Veroyatnost. i Primenen., 29(4):669–679, 1984.

[73] V. L. Girko. Theory of random determinants, volume 45 of Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1990. Translated from the Russian.

[74] Gene H. Golub and Charles F. Van Loan. Matrix computations. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[75] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. Ann. Probab., 38(4):1444–1491, 2010.

[76] Allan Gut. Probability: a graduate course. Springer Texts in Statistics. Springer, New York, 2005.

[77] Andrew J. Hetzel, Jay S. Liew, and Kent E. Morrison. The probability that a matrix of integers is diagonalizable. Amer. Math. Monthly, 114(6):491–499, 2007.

[78] P. Holgate. Studies in the history of probability and statistics xli waring and sylvester on random algebraic equations. Biometrika, 73(1):228–231, 1986.

[79] Roger A. Horn and Charles R. Johnson. Matrix analysis. Cambridge University Press, Cambridge, second edition, 2013.

[80] I. A. Ibragimov and N. B. Maslova. The average number of zeros of random polynomials. Vestnik Leningrad. Univ., 23(19):171–172, 1968.

[81] I. A. Ibragimov and N. B. Maslova. The average number of real roots of random polynomials. Dokl. Akad. Nauk SSSR, 199:13–16, 1971.

[82] I. A. Ibragimov and N. B. Maslova. The mean number of real zeros of random polynomials. I. Coefficients with zero mean. Teor. Verojatnost. i Primenen., 16:229–248, 1971.

[83] I. A. Ibragimov and N. B. Maslova. The mean number of real zeros of random polynomials. II. Coefficients with a nonzero mean. Teor. Verojatnost. i Primenen., 16:495–503, 1971.

[84] R.S. Irving. Integers, Polynomials, and Rings: A Course in Algebra. Undergraduate Texts in Mathematics. Springer New York, 2004.

[85] A. T. James. Normal multivariate analysis and the orthogonal group. Ann. Math. Statistics, 25:40–75, 1954.

[86] Shane T. Jensen and George P. H. Styan. Some Comments and a Bibliography on the Laguerre-Samuelson Inequality with Extensions and Applications in Statistics and Matrix Theory, pages 151–181. Springer Netherlands, Dordrecht, 1999.

[87] Zakhar Kabluchko and Dmitry Zaporozhets. Universality for zeros of random analytic functions. Available at arXiv:1205.5355, 2012.

[88] M. Kac. A correction to "On the average number of real roots of a random algebraic equation.". Bull. Amer. Math. Soc., 49:938, 1943.

[89] M. Kac. On the average number of real roots of a random algebraic equation. Bull. Amer. Math. Soc., 49:314–320, 1943.

[90] M. Kac. On the average number of real roots of a random algebraic equation. II. Proc. London Math. Soc. (2), 50:390–408, 1949.

[91] Eugene Kanzieper and Gernot Akemann. Statistics of real eigenvalues in Ginibre's ensemble of random real matrices. Phys. Rev. Lett., 95(23):230201, 4, 2005.

[92] Eugene Kanzieper, Mihail Poplavskyi, Carsten Timm, Roger Tribe, and Oleg Zaboronski. What is the probability that a large random matrix has no real eigenvalues? Ann. Appl. Probab., 26(5):2733–2753, 2016.

[93] Sankeerth Rao Karingula and Shachar Lovett. Singularity of random integer matrices with large entries. In Approximation, randomization, and combinatorial optimization. Algorithms and techniques, volume 207 of LIPIcs. Leibniz Int. Proc. Inform., pages Art. No. 33, 16. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2021.

[94] Y. R. Katznelson. Singular matrices and a uniform bound for congruence groups of $\mathrm{SL}_n(\mathbf{Z})$. Duke Math. J., 69(1):121–136, 1993.

[95] Yonatan R. Katznelson. Integral matrices of fixed rank. Proc. Amer. Math. Soc., 120(3):667–675, 1994.

[96] Hans-Wilhelm Knobloch. Zum Hilbertschen Irreduzibilitätssatz. Abh. Math. Sem. Univ. Hamburg, 19:176–190, 1955.

[97] Hans-Wilhelm Knobloch. Die Seltenheit der reduziblen Polynome. Jber. Deutsch. Math.-Verein., 59:12–19, 1956.

[98] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's and the sample DF. I. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 32:111–131, 1975.

[99] E. Kostlan. On the distribution of roots of random polynomials. In From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990), pages 419–431. Springer, New York, 1993.

[100] Isabelle Kraus, Marcus Michelen, and Sean O'Rourke. Sums of random polynomials with differing degrees. Trans. Amer. Math. Soc., 377(5):3325–3355, 2024.

[101] Gerald Kuba. On the distribution of reducible polynomials. Math. Slovaca, 59(3):349–356, 2009.

[102] David C. Kurtz. A sufficient condition for all the roots of a polynomial to be real. Amer. Math. Monthly, 99(3):259–263, 1992.

[103] Edmond Laguerre. Sur une méthode pour obtenir par approximation les racines d'une équation algébrique qui a toutes ses racines réelles. Nouv. Ann., 2(19):193–202, 1880.

[104] John M. Lee. Introduction to smooth manifolds. Graduate Texts in Mathematics. Springer, New York, NY, second edition, August 2012.

[105] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. Journal of the London Mathematical Society, s1-13(4):288–295, 1938.

[106] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. ii. Mathematical Proceedings of the Cambridge Philosophical Society, 35(2):133–148, 1939.

[107] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. III. Rec. Math. [Mat. Sbornik] N.S., 12/54:277–286, 1943.

[108] J. E. Littlewood and A. C. Offord. On the distribution of the zeros and a-values of a random integral function. I. J. London Math. Soc., 20:130–136, 1945.

[109] J. E. Littlewood and A. C. Offord. On the distribution of zeros and a-values of a random integral function. II. Ann. of Math. (2), 49:885–952, 1948.

[110] B. F. Logan and L. A. Shepp. Real zeros of random polynomials. Proc. London Math. Soc. (3), 18:29–35, 1968.

[111] B. F. Logan and L. A. Shepp. Real zeros of random polynomials. II. Proc. London Math. Soc. (3), 18:308–314, 1968.

[112] Kyle Luh. Complex random matrices have no real eigenvalues. Random Matrices Theory Appl., 7(1):1750014, 17, 2018.

[113] I G Macdonald. Symmetric functions and hall polynomials. Oxford Classic Texts in the Physical Sciences. Oxford University Press, London, England, second edition, November 2015.

[114] Greg Martin and Erick Wong. The number of 2×2 integer matrices having a prescribed integer eigenvalue. Algebra Number Theory, 2(8):979–1000, 2008.

[115] Greg Martin and Erick B. Wong. Almost all integer matrices have no integer eigenvalues. Amer. Math. Monthly, 116(7):588–597, 2009.

[116] N. B. Maslova. The distribution of the number of real roots of random polynomials. Teor. Verojatnost. i Primenen., 19:488–500, 1974.

[117] N. B. Maslova. The variance of the number of real roots of random polynomials. Teor. Verojatnost. i Primenen., 19:36–51, 1974.

[118] M. L. Mehta. Random matrices and the statistical theory of energy levels. Academic Press, New York-London, 1967.

[119] Madan Lal Mehta. Random matrices, volume 142 of Pure and Applied Mathematics (Amsterdam). Elsevier/Academic Press, Amsterdam, third edition, 2004.

[120] Maurice Mignotte and Philippe Glesser. Landau's inequality via Hadamard's. J. Symbolic Comput., 18(4):379–383, 1994.

[121] Robb J. Muirhead. Aspects of multivariate statistical theory. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1982.

[122] Hoi Nguyen, Oanh Nguyen, and Van Vu. On the number of real roots of random polynomials. Commun. Contemp. Math., 18(4):1550052, 17, 2016.

[123] Hoi Nguyen and Van Vu. Optimal inverse Littlewood-Offord theorems. Adv. Math., 226(6):5298–5319, 2011.

[124] Sean O'Rourke and Tulasi Ram Reddy. Sums of random polynomials with independent roots. J. Math. Anal. Appl., 495(1):Paper No. 124719, 23, 2021.

[125] Sean O'Rourke and David Renfrew. Low rank perturbations of large elliptic random matrices. Electron. J. Probab., 19:no. 43, 65, 2014.

[126] Sean O'Rourke and Philip Matchett Wood. Low-degree factors of random polynomials. J. Theoret. Probab., 32(2):1076–1104, 2019.

[127] Patryk Pagacz and Michał Wojtylak. Random perturbations of matrix polynomials. J. Theoret. Probab., 35(1):52–88, 2022.

[128] Guangming Pan and Wang Zhou. Circular law, extreme singular values and potential theory. J. Multivariate Anal., 101(3):645–656, 2010.

[129] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. Probab. Theory Related Fields, 134(1):127–173, 2006.

[130] V. V. Petrov. Sums of independent random variables, volume Band 82 of Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown.

[131] Huy Tuan Pham and Max Wenqiang Xu. Irreducibility of random polynomials of bounded degree. Discrete Anal., pages Paper No. 7, 16, 2021.

[132] G. Pólya and G. Szegő. Problems and theorems in analysis. Vol. II. Springer Study Edition. Springer-Verlag, New York-Heidelberg, german edition, 1976. Theory of functions, zeros, polynomials, determinants, number theory, geometry.

[133] C E Porter and N Rosenzweig. Statistical properties of atomic and nuclear spectra. Ann. Acad. Sci. Fennicae. Ser. A VI, Vol: No. 44, 01 1960.

[134] V. V. Prasolov. Polynomials, volume 11 of Algorithms and Computation in Mathematics. Springer, Berlin, 2010.

[135] Richard C. Raffenetti and Klaus Ruedenberg. Parametrization of an orthogonal matrix in terms of generalized eulerian angles. Int. J. Quantum Chem, 4(S):625–634, 1969.

[136] Q. I. Rahman and G. Schmeisser. Analytic theory of polynomials, volume 26 of London Mathematical Society Monographs. New Series. The Clarendon Press, Oxford University Press, Oxford, 2002.

[137] Anand B. Rajagopalan. Outlier eigenvalue fluctuations of perturbed iid matrices. Available at arXiv: 1507.01441, 2015.

[138] Igor Rivin. Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. Duke Math. J., 142(2):353–379, 2008.

[139] Igor Rivin. Galois groups of generic polynomials. Available at arXiv:1511.06446, 2015.

[140] Sebastien Roch. Modern Discrete Probability: An Essential Toolkit. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2024.

[141] Shigui Ruan and Junjie Wei. On the zeros of transcendental functions with applications to stability of delay differential equations with two delays. Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal., 10(6):863–874, 2003.

[142] G. Samal. On the number of real roots of a random algebraic equation. Proc. Cambridge Philos. Soc., 58:433–442, 1962.

[143] G. Samal and M. N. Mishra. On the lower bound of the number of real roots of a random algebraic equation with infinite variance. Proc. Amer. Math. Soc., 33:523–528, 1972.

[144] G. Samal and M. N. Mishra. On the lower bound of the number of real roots of a random algebraic equation with infinite variance. II. Proc. Amer. Math. Soc., 36:557–563, 1972.

[145] Issai Schur. Sitzungsb. Preuss. Akad. Wiss., pages 403–428, 1933.

[146] Bernard Shiffman and Steve Zelditch. Equilibrium distribution of zeros of random polynomials. International Mathematics Research Notices, 2003(1):25–49, 2003.

[147] Theodore Shifrin. Multivariable mathematics: linear algebra, multivariable calculus, and manifolds. Wiley, New Jersey, first edition, January 2004.

[148] Wilhelm Specht. Zur Zahlentheorie der Polynome. S.-B. Math.-Nat. Kl. Bayer. Akad. Wiss., 1951:139–146, 1951.

[149] Melvin D. Springer. The algebra of random variables. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York-Chichester-Brisbane, 1979.

[150] D. C. Stevens. The average number of real zeros of a random polynomial. Comm. Pure Appl. Math., 22:457–477, 1969.

[151] J. J. Sylvester. On an elementary proof and generalisation of Sir Isaac Newton's hitherto undemonstrated Rule for the discovery of imaginary roots. Proceedings of the London Mathematical Society, 1:1–16, 1865. Reprinted (1908) in *Mathematical Papers*, vol. 2, edited by H. F. Baker, pp. 498–513, Cambridge University Press.

[152] Terence Tao. Erratum to: Outliers in the spectrum of iid matrices with bounded rank perturbations [mr3010398]. Probab. Theory Related Fields, 157(1-2):511–514, 2013.

[153] Terence Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. Probab. Theory Related Fields, 155(1-2):231–263, 2013.

[154] Terence Tao and Van Vu. Random matrices: the circular law. Commun. Contemp. Math., 10(2):261–307, 2008.

[155] Terence Tao and Van Vu. From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices. Bull. Amer. Math. Soc. (N.S.), 46(3):377–396, 2009.

[156] Terence Tao and Van Vu. Random matrices: universality of ESDs and the circular law. Ann. Probab., 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur.

[157] Terence Tao and Van Vu. Local universality of zeroes of random polynomials. International Mathematics Research Notices, 2015(13):5053–5139, 2014.

[158] Terence Tao and Van Vu. Random matrices: universality of local spectral statistics of non-Hermitian matrices. Ann. Probab., 43(2):782–874, 2015.

[159] V. Thangaraj and M. Sambandham. On random polynomials-i: A survey. Neural, Parallel, and Scientific Computations, 28(2):106–142, 2020.

[160] Konstantin Tikhomirov. Singularity of random Bernoulli matrices. Ann. of Math. (2), 191(2):593–634, 2020.

[161] E. C. Titchmarsh. The theory of functions. Oxford University Press, Oxford, second edition, 1939.

[162] Isaac Todhunter. A History of the Mathematical Theory of Probability: From the Time of Pascal to that of Laplace. Cambridge Library Collection - Mathematics. Cambridge University Press, 2014.

[163] A. W. van der Vaart. Asymptotic statistics, volume 3 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1998.

[164] B. L. van der Waerden. Die Seltenheit der Gleichungen mit Affekt. Math. Ann., 109(1):13–16, 1934.

[165] B. L. van der Waerden. Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt. Monatsh. Math. Phys., 43(1):133–147, 1936.

[166] You Jing Wang. Bounds on the average number of real roots of a random algebraic equation. Chinese Ann. Math. Ser. A, 4(5):601–605, 1983. An English summary appears in Chinese Ann. Math. Ser. B **4** (1983), no. 4, 527.

[167] Edward Waring. Meditationes algebraicæ. American Mathematical Society, Providence, RI, 1991. Translated from the Latin, edited and with a foreword by Dennis Weeks, With an appendix by Franz X. Mayer, translated from the German by Weeks.

[168] Harold Widom. On the eigenvalues of certain Hermitian operators. Trans. Amer. Math. Soc., 88:491–522, 1958.

[169] E. P. Wigner. On the statistical distribution of the widths and spacings of nuclear resonance levels. Proc. Cam. Phil. Soc., 47:790, 1951.

[170] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math., 62:548, 1955.

[171] E. P. Wigner. Statistical properties of real symmetric matrices with many dimensions. In Can. Math. Congr. Proc., page 174, Toronto, 1957. University of Toronto Press.

[172] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. Ann. of Math. (2), 67:325–327, 1958.

[173] J. Ernest Wilkins, Jr. An asymptotic expansion for the expected number of real zeros of a random polynomial. Proc. Amer. Math. Soc., 103(4):1249–1258, 1988.

[174] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. Biometrika, 20A:32–52, 1928.

[175] David Zywina. Hilbert's irreducibility theorem and the larger sieve. Available at arXiv: 1011.6465v1, 2010.

Appendix A

Pdf of the ratio distribution for continuous uniform random variables

We now investigate the probability density function of the limiting distribution of Z_N, as defined in Corollary 2.1.6. Springer [149] obtains the probability density functions of various random variables in Chapter 4 of his book via the Mellin convolution and the residue theorem. We present here a straightforward geometrical approach for the ratio distribution of two independent uniform random variables.

Lemma A.0.1. *Let $0 \leq \alpha < \beta$ and $0 \leq \gamma < \delta$ and suppose that $X \sim U(\alpha, \beta)$ and $Y \sim U(\gamma, \delta)$ are independent. If $\frac{\delta}{\beta} < \frac{\gamma}{\alpha}$, then $|X/Y|$ has density*

$$f_{|X/Y|}(z) = \begin{cases} \frac{\delta^2 z^2 - \alpha^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, & \textit{if } z \in \left[\frac{\alpha}{\delta}, \frac{\alpha}{\gamma}\right] \\ \frac{\gamma+\delta}{2(\beta-\alpha)}, & \textit{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right] \\ \frac{\beta^2 - z^2\gamma^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, & \textit{if } z \in \left[\frac{\beta}{\delta}, \frac{\beta}{\gamma}\right] \\ 0, & \textit{otherwise} \end{cases}.$$

Otherwise, if $\frac{\delta}{\beta} \geq \frac{\gamma}{\alpha}$, then $|X/Y|$ has density

$$f_{|X/Y|}(z) = \begin{cases} \frac{\delta^2 z^2 - \alpha^2}{2(\beta-\alpha)(\delta-\gamma)z^2}, & \textit{if } z \in \left[\frac{\alpha}{\delta}, \frac{\beta}{\delta}\right] \\ \frac{\alpha+\beta}{2(\delta-\gamma)z^2}, & \textit{if } z \in \left[\frac{\beta}{\delta}, \frac{\alpha}{\gamma}\right] \\ \frac{\beta^2 - \gamma^2 z^2}{2z^2(\beta-\alpha)(\delta-\gamma)}, & \textit{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}\right] \\ 0, & \textit{otherwise} \end{cases}.$$

Proof. Since $\mathbb{P}(X/Y \le z) = 1 - \mathbb{P}(Y < \frac{X}{z})$, we will begin by computing $\mathbb{P}(Y < \frac{X}{z})$ for all $z \in \mathbb{R}$.

Consider the rectangle R in the complex plane defined by $[\alpha, \beta] \times [\gamma, \delta]$. Since X and Y are independently distributed with a joint density that is uniform over R, for any $z \in \mathbb{R}$, the probability that $Y < \frac{X}{z}$ is equal to the area below the line $y = \frac{x}{z}$ and enclosed by the rectangle, divided by the total area of the rectangle. The four corners of the rectangle are given by $(\alpha, \gamma), (\beta, \gamma), (\alpha, \delta)$, and (β, δ). A visual aid is provided in Figure A.1.

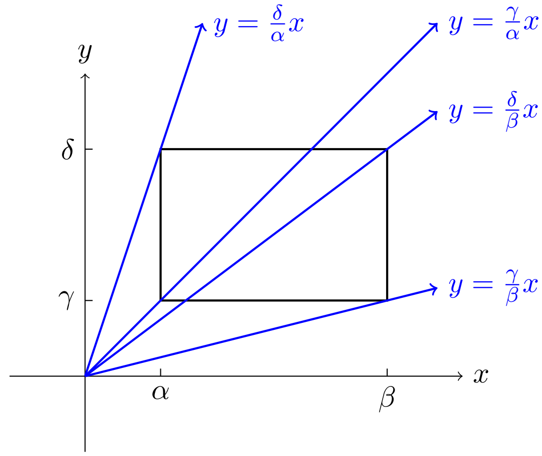

Figure A.1: This is the rectangle given by $[\alpha, \beta] \times [\gamma, \delta]$, when $0 \le \alpha < \beta$ and $0 \le \gamma < \delta$. The figure depicts the lines that go through all four corners of the rectangle. This serves as a visual aid for computing $\mathbb{P}\left(Y < \frac{X}{z}\right)$ in the case where $\frac{\delta}{\beta} < \frac{\gamma}{\alpha}$.

Observe that the lower right corner of the rectangle is given by the point (β, γ). Therefore, the line through the origin with slope $\frac{\gamma}{\beta}$ will only intersect the rectangle at one point. We conclude that for any $z \in \left[\frac{\beta}{\gamma}, \infty\right), \mathbb{P}\left(Y < \frac{X}{z}\right) = 0$. Similarly, observing that the upper left corner of the rectangle is given by the point (α, δ), for any $z \in \left[0, \frac{\alpha}{\delta}\right], \mathbb{P}\left(Y < \frac{X}{z}\right) = 1$. Finally, since the support of the rectangle is contained in the first quadrant, for any $z \in (-\infty, 0), \mathbb{P}\left(Y < \frac{X}{z}\right) = 0$. Differentiating $1 - \mathbb{P}\left(Y < \frac{X}{z}\right)$ and combining these statements yields

$$f_{|X/Y|}(z) = 0, \quad z \in \left(-\infty, \frac{\alpha}{\delta}\right] \cup \left[\frac{\beta}{\gamma}, \infty\right).$$

As the slope of the line through the origin increases, after intersecting the lower right corner (β, γ), the line will intersect either the point (α, γ) or (β, δ) next, depending on whether the quantity

$\frac{\delta}{\beta}$ or $\frac{\gamma}{\alpha}$ is greater. This is why we must consider two cases.

Suppose first that

$$\frac{\delta}{\beta} < \frac{\gamma}{\alpha}.$$

Then for $z \in \left[\frac{\beta}{\delta}, \frac{\beta}{\gamma}\right]$,

$$\begin{aligned}\mathbb{P}\left(Y < \frac{X}{z}\right) &= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \int_{\gamma z}^{\beta} \frac{x}{z} - \gamma \, dx \\ &= \frac{(\beta - \gamma z)^2}{2z(\delta-\gamma)(\beta-\alpha)}.\end{aligned}$$

Differentiating $1 - \mathbb{P}\left(Y < \frac{X}{z}\right)$ yields

$$f_{|X/Y|}(z) = \frac{\beta^2 - z^2\gamma^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, \qquad \text{if } z \in \left[\frac{\beta}{\delta}, \frac{\beta}{\gamma}\right].$$

Then for $z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\delta}, \right]$,

$$\begin{aligned}\mathbb{P}\left(Y < \frac{X}{z}\right) &= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\gamma z}^{\delta z} \frac{x}{z} - \gamma \, dx + \int_{\delta z}^{\beta} \delta - \gamma \, dx\right) \\ &= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\frac{z(\gamma-\delta)^2}{2} + (\delta-\gamma)(\beta-\delta z)\right).\end{aligned}$$

Differentiating $1 - \mathbb{P}\left(Y < \frac{X}{z}\right)$ yields

$$f_{|X/Y|}(z) = \frac{\gamma+\delta}{2(\beta-\alpha)}, \qquad \text{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right].$$

Finally, for $z \in \left[\frac{\alpha}{\delta}, \frac{\alpha}{\gamma}\right]$,

$$\begin{aligned}\mathbb{P}\left(Y < \frac{X}{z}\right) &= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\alpha}^{\delta z} \frac{x}{z} - \gamma \, dx + \int_{\delta z}^{\beta} \delta - \gamma \, dx\right) \\ &= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\frac{(\delta z - \alpha)(\alpha - 2\gamma z + \delta z)}{2z} + (\delta-\gamma)(\beta-\delta z)\right).\end{aligned}$$

Differentiating $1 - \mathbb{P}\left(Y < \frac{X}{z}\right)$ yields

$$f_{|X/Y|}(z) = \frac{\delta^2 z^2 - \alpha^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, \qquad \text{if } z \in \left[\frac{\alpha}{\delta}, \frac{\alpha}{\gamma}\right].$$

When $\frac{\gamma}{\alpha} \leq \frac{\delta}{\beta}$ instead, one can proceed analogously to obtain

$$f_{|X/Y|}(z) = \frac{\alpha+\beta}{2(\delta-\gamma)z^2}, \qquad \text{if } z \in \left[\frac{\beta}{\delta}, \frac{\alpha}{\gamma}\right]$$

$$f_{|X/Y|}(z) = \frac{\beta^2 - \gamma^2 z^2}{2z^2(\beta-\alpha)(\delta-\gamma)}, \quad \text{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}\right].$$
$$f_{|X/Y|}(z) = \frac{\delta^2 z^2 - \alpha^2}{2(\beta-\alpha)(\delta-\gamma)z^2}, \quad \text{if } z \in \left[\frac{\alpha}{\delta}, \frac{\beta}{\delta}\right].$$

□

Remark A.0.2. If $\alpha < \beta \leq 0$ and $\gamma < \delta \leq 0$ (instead of $0 \leq \alpha < \beta$ and $0 \leq \gamma < \delta$) in the hypothesis of Lemma A.0.1, it is not hard to see that if $\frac{\gamma}{\alpha} < \frac{\delta}{\beta}$, then $|X/Y|$ has density

$$f_{|X/Y|}(z) = \begin{cases} \frac{\gamma^2 z^2 - \beta^2}{2z^2(\gamma-\delta)(\alpha-\beta)}, & \text{if } z \in \left[\frac{\beta}{\gamma}, \frac{\beta}{\delta}\right] \\ \frac{\gamma+\delta}{2(\alpha-\beta)}, & \text{if } z \in \left[\frac{\beta}{\delta}, \frac{\alpha}{\gamma}\right] \\ \frac{\alpha^2 - z^2\delta^2}{2z^2(\gamma-\delta)(\alpha-\beta)}, & \text{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\alpha}{\delta}\right] \\ 0, & \text{otherwise} \end{cases}.$$

and if $\frac{\gamma}{\alpha} \geq \frac{\delta}{\beta}$, then $|X/Y|$ has density

$$f_{|X/Y|}(z) = \begin{cases} \frac{\gamma^2 z^2 - \beta^2}{2(\alpha-\beta)(\gamma-\delta)z^2}, & \text{if } z \in \left[\frac{\beta}{\gamma}, \frac{\alpha}{\gamma}\right] \\ \frac{\alpha+\beta}{2(\gamma-\delta)z^2}, & \text{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right] \\ \frac{\alpha^2 - \delta^2 z^2}{2z^2(\alpha-\beta)(\gamma-\delta)}, & \text{if } z \in \left[\frac{\beta}{\delta}, \frac{\alpha}{\delta}\right] \\ 0, & \text{otherwise} \end{cases}.$$

Lemma A.0.3. *Let $\alpha < 0 < \beta$ and $\gamma < 0 < \delta$. Suppose that $X \sim U(\alpha, \beta)$ and $Y \sim U(\gamma, \delta)$ are independent. If $\alpha\delta \geq \beta\gamma$ and $\beta\delta \geq \alpha\gamma$, then X/Y has density*

$$f_{X/Y}(z) = \begin{cases} \frac{\alpha^2+\beta^2}{2z^2(\beta-\alpha)(\delta-\gamma)}, & \textit{if } z \in \left(-\infty, \frac{\beta}{\gamma}\right] \cup \left[\frac{\beta}{\delta}, \infty\right) \\ \frac{\gamma^2+\delta^2}{2(\beta-\alpha)(\delta-\gamma)}, & \textit{if } z \in \left[\frac{\alpha}{\delta}, 0\right) \cup \left[0, \frac{\alpha}{\gamma}\right] \\ \frac{\alpha^2+\delta^2 z^2}{2z^2(\beta-\alpha)(\delta-\gamma)}, & \textit{if } z \in \left[\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right] \\ \frac{\alpha^2+\gamma^2 z^2}{2z^2(\beta-\alpha)(\delta-\gamma)}, & \textit{if } z \in \left[\frac{\beta}{\gamma}, \frac{\alpha}{\delta}\right] \\ 0, & \textit{otherwise} \end{cases}.$$

Proof. Consider the rectangle R defined by $[\alpha, \beta] \times [\gamma, \delta]$, and begin with the simple observation that for $z \in [0, \infty)$, the region that satisfies the inequality $X/Y \leq z$ is given by the intersection of R with the second quadrant, the fourth quadrant, the area above the line $y = \frac{x}{z}$ in the first quadrant, and the area below the line $y = \frac{x}{z}$ in the third quadrant. A visual aid is provided in Figure A.2.

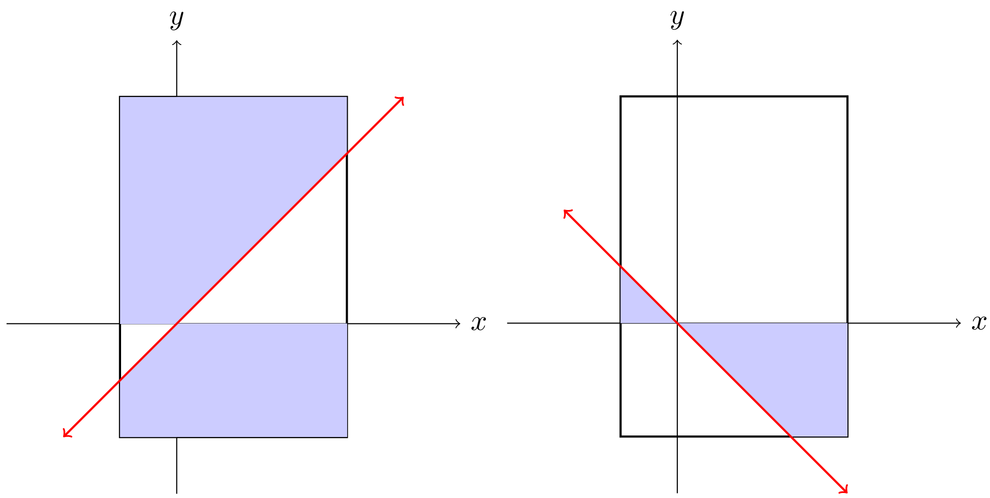

Figure A.2: This is the rectangle given by $[\alpha, \beta] \times [\gamma, \delta]$. The figure on the left depicts, in blue, the points in the rectangle that satisfy $\frac{x}{y} \leq z$ for z positive. The figure on the right depicts, in blue, the points in the rectangle that satisfy $\frac{x}{y} \leq z$ for z negative.

When $z \in (-\infty, 0)$, the region that satisfies the inequality $X/Y \leq z$ is the intersection of R with the area below the line $y = \frac{x}{z}$ in the second quadrant and the area above the line $y = \frac{x}{z}$ in the fourth quadrant. To obtain the probability of these events, we divide the area of these regions by the total area of the rectangle.

We begin by considering positive values for z. Similarly to the proof of Lemma A.0.1, as the slope of the line through the origin increases from 0 towards ∞, the line will intersect either the point (α, γ) or (β, δ) first, depending on whether the quantity $\frac{\delta}{\beta}$ or $\frac{\gamma}{\alpha}$ is greater. This leads to two cases. Consider now the case where $\alpha\delta \geq \beta\gamma$.

When $z \in \left[0, \frac{\alpha}{\gamma}\right]$,

$$\mathbb{P}\left(\frac{X}{Y} \leq z\right) = \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\alpha}^{0} \delta \; dx + \int_{0}^{\delta z} \delta - \frac{x}{z} \; dx + \int_{\gamma z}^{0} \frac{x}{z} - \gamma \; dx + \int_{0}^{\beta} -\gamma \; dx\right)$$
$$= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(-\delta\alpha + \frac{\delta^2 z}{2} + \frac{\gamma^2 z}{2} - \beta\gamma\right).$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\gamma^2+\delta^2}{2(\delta-\gamma)(\beta-\alpha)}, \qquad \text{if } z \in \left[0, \frac{\alpha}{\gamma}\right].$$

When $z \in \left(\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right)$,

$$\mathbb{P}\left(\frac{X}{Y} \leq z\right) = \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\alpha}^{0} \delta \; dx + \int_{0}^{\delta z} \delta - \frac{x}{z} \; dx + \int_{\alpha}^{0} \frac{x}{z} - \gamma \; dx + \int_{0}^{\beta} -\gamma \; dx\right)$$
$$= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(-\delta\alpha + \frac{\delta^2 z}{2} + \alpha\gamma - \frac{\alpha^2}{2z} - \beta\gamma\right).$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\alpha^2+\delta^2 z^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, \qquad \text{if } z \in \left(\frac{\alpha}{\gamma}, \frac{\beta}{\delta}\right).$$

When $z \in \left[\frac{\beta}{\delta}, \infty\right)$,

$$\mathbb{P}\left(\frac{X}{Y} \leq z\right) = \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\alpha}^{0} \delta \; dx + \int_{0}^{\beta} \delta - \frac{x}{z} \; dx + \int_{\alpha}^{0} \frac{x}{z} - \gamma \; dx + \int_{0}^{\beta} -\gamma \; dx\right)$$
$$= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(-\delta\alpha + \beta\delta - \frac{\beta^2}{2z} + \alpha\gamma - \frac{\alpha^2}{2z} - \beta\gamma\right).$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\alpha^2+\beta^2}{2z^2(\delta-\gamma)(\beta-\alpha)}, \qquad \text{if } z \in \left[\frac{\beta}{\delta}, \infty\right).$$

Next, we consider negative values for z. As the slope of the line through the origin decreases from 0 towards $-\infty$, the line will intersect either the point (β, γ) or (α, δ) first, depending on whether the quantity $\frac{\gamma}{\beta}$ or $\frac{\delta}{\alpha}$ is greater. This leads to two cases. Consider now the case where $\beta\delta \geq \alpha\gamma$.

When $z \in \left(-\infty, \frac{\beta}{\gamma}\right]$,

$$\mathbb{P}\left(\frac{X}{Y} \leq z\right) = \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(\int_{\alpha}^{0} \frac{x}{z} \; dx + \int_{0}^{\beta} -\frac{x}{z} \; dx\right)$$
$$= \frac{1}{(\delta-\gamma)(\beta-\alpha)} \left(-\frac{\alpha^2+\beta^2}{2z}\right).$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\alpha^2 + \beta^2}{2z^2(\delta - \gamma)(\beta - \alpha)}, \qquad \text{if } z \in \left(-\infty, \frac{\beta}{\gamma}\right].$$

When $z \in \left(\frac{\beta}{\gamma}, \frac{\alpha}{\delta}\right)$,

$$\begin{aligned}\mathbb{P}\left(\frac{X}{Y} \leq z\right) &= \frac{1}{(\delta - \gamma)(\beta - \alpha)} \left(\int_{\alpha}^{0} \frac{x}{z}\, dx + \int_{0}^{\gamma z} -\frac{x}{z}\, dx + \int_{\gamma z}^{\beta} -\gamma\, dx\right) \\ &= \frac{1}{(\delta - \gamma)(\beta - \alpha)} \left(-\frac{\alpha^2}{2z} - \frac{\gamma^2 z}{2} + \gamma^2 z - \gamma\beta\right).\end{aligned}$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\alpha^2 + \gamma^2 z^2}{2z^2(\delta - \gamma)(\beta - \alpha)}, \qquad \text{if } z \in \left(\frac{\beta}{\gamma}, \frac{\alpha}{\delta}\right).$$

When $z \in \left[\frac{\alpha}{\delta}, 0\right)$,

$$\begin{aligned}\mathbb{P}\left(\frac{X}{Y} \leq z\right) &= \frac{1}{(\delta - \gamma)(\beta - \alpha)} \left(\int_{\alpha}^{\delta z} \delta\, dx + \int_{\delta z}^{0} \frac{x}{z}\, dx + \int_{0}^{\gamma z} -\frac{x}{z}\, dx + \int_{\gamma z}^{\beta} -\gamma\, dx\right) \\ &= \frac{1}{(\delta - \gamma)(\beta - \alpha)} \left(\delta^2 z - \delta\alpha - \frac{\delta^2 z}{2} - \frac{\gamma^2 z}{2} + \gamma^2 z - \gamma\beta\right).\end{aligned}$$

Differentiating gives

$$f_{X/Y}(z) = \frac{\gamma^2 + \delta^2}{2(\delta - \gamma)(\beta - \alpha)}, \qquad \text{if } z \in \left[\frac{\alpha}{\delta}, 0\right).$$

□

In the proof of Lemma A.0.3, we saw that different cases for the density of X/Y must be considered based on whether $\alpha\delta \geq \beta\gamma$ or $\alpha\delta < \beta\gamma$ and whether $\beta\delta \geq \alpha\gamma$ or $\beta\delta < \alpha\gamma$. This gives rise to a total of four cases; we computed the probability density of X/Y for just one. The density of X/Y for the other cases can be computed analogously. One may also consider when $\alpha < 0 < \beta$ and $0 < \gamma < \delta$ or when $0 < \alpha < \beta$ and $\gamma < 0 < \delta$ in the statement of Lemma A.0.3 to fully derive the probability density function of X/Y.

Remark A.0.4. To find the pdf of $\left|\frac{X}{Y}\right|$, we simply observe that for any $z \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{X}{Y}\right| < z\right) &= \mathbb{P}\left(-z < \frac{X}{Y} < z\right) \\ &= F_{X/Y}(z) - F_{X/Y}(-z)\end{aligned}$$

Differentiating shows that

$$f_{|X/Y|}(z) = f_{X/Y}(z) - f_{X/Y}(-z).$$