Global and Local Limit Laws for Eigenvalues of the Gaussian Unitary Ensemble and the Wishart Ensemble

Author:
Zhou Fan

Supervisor:
Prof. Horng-Tzer Yau

Submitted to the Harvard University Department of Mathematics in partial fulfillment of the requirements for the degree of A.B. in Mathematics

March 22, 2010

Contents

1 Introduction

Linear algebra deals with the study of vector spaces and linear transformations between them. These transformations can be represented as matrices, and various properties of a linear transformation are reflected in the properties of its corresponding matrix. When the elements of a matrix are replaced by random variables, tools from probability theory can be used to study the resulting properties of these matrices. This paper presents a survey of several results in random matrix theory, the study of spectral properties of matrices with random elements, and analyzes two families of random matrices in a unified way. To motivate our discussion, we begin with three examples:

Example 1.1 (Nuclear Physics [18]). *Consider the nucleus of a large atom (e.g. Uranium-238). We are interested in determining the energy levels of this nucleus. From quantum mechanics, the nuclear energy levels E_n are given as the eigenvalues of the Hamiltonian operator H of the system, $H\psi_n = E_n\psi_n$. Unfortunately, for large atoms, the Hamiltonian H cannot be explicitly computed and so we cannot explicitly determine its spectrum. However, we may instead model the system by approximating the infinite-dimensional Hilbert space of wave functions with a large finite-dimensional space, and approximating H by a matrix operator on this space with random elements. We may then study local statistics of the energy levels, such as their pairwise joint distributions or distributions concerning the spacings between them, using this model. By imposing conditions on the joint distribution of the matrix elements based on the symmetries of the system, we find that the local statistical properties of the eigenvalue distributions of these random matrix models closely match those of observed nuclear energy levels in large atoms.*

Example 1.2 (Wireless Communications [27]). *Consider the transmission of information over a wireless communication channel. The relation between a vector of transmitted data x and received data y can be modeled by a linear channel $y = Hx + n$, where H is a random channel matrix and n is a vector of random Gaussian noise. An information theoretic quantity of interest is the channel capacity, an upper bound on the rate of information transmission over the channel. Suppose that $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^n$, and the matrix HH^* has* *eigenvalues* $\lambda_1, \ldots, \lambda_n$*. Let* $F_H(x) = \frac{1}{n} \cdot \#\{\lambda_i \leq x\}$ *be the cumulative distribution function of the probability mass function of these eigenvalues. Then, under suitable assumptions on the distributions of* x *and* H*, the channel capacity is given by the expected value of*

$$n \int_0^\infty \log\left(1 + \frac{n\mathbb{E}[\|x\|^2]}{k\mathbb{E}[\|n\|^2]} x\right) dF_H(x)$$

over the distribution of the channel matrix H*. In particular, the channel capacity is dependent on the global distribution* F_H *of the eigenvalues of* HH^**.*

Example 1.3 (Financial Portfolio Optimization [16, 21])**.** *Consider a collection of* n *stocks. We are interested in understanding the risk associated to a portfolio of these stocks and in constructing portfolios of low risk. If we model the return of each stock as a random variable* r_i *and consider the covariance matrix* C *with entries* $C_{ij} = \mathrm{Cov}(r_i, r_j)$*, then the risk associated to a portfolio* $p = (p_1, \ldots, p_n)$ *where* p_i *is the amount invested in stock* i *is given by* p^tCp*. Low risk portfolios can be selected to have large components in the directions of eigenvectors of* C *with the lowest eigenvalues. However, by estimating* C *using the sample covariance matrix of observed returns for each stock, the noise in the sample matrix can cause us to misidentify the eigenvectors corresponding to the lowest eigenvalues of* C *and to underestimate the risk associated with the chosen portfolio when* n *is large. We may use properties of the global eigenvalue distribution of the random sample covariance matrix to devise more accurate methods of portfolio selection and adjust for the noise factor in the computation of risk.*

There exists a body of research pertaining to each of these applications of random matrix theory. It is not the goal of this paper to discuss the details of these applications; we refer the interested reader to the listed references. We present these examples as an illustration of the diversity of the range of applications of this theory and as motivation for the specific families of random matrices and the specific properties of their eigenvalue distributions that we will examine. In particular, we will focus on two families of random matrices, defined as follows:

Definition 1.4. *Let* $\{\alpha_{ij}^{(n)}\}_{n\in\mathbb{N}, 1\leq i\leq j\leq n}$ *and* $\{\beta_{ij}^{(n)}\}_{n\in\mathbb{N}, 1\leq i<j\leq n}$ *be i.i.d. random variables, normally distributed with mean 0 and variance 1. Let* W_n *be an* $n \times n$ *matrix for each* n*, with diagonal entries* $(W_n)_{ii} = \alpha_{ii}^{(n)}$ *for* $1 \leq i \leq n$*, above-diagonal entries* $(W_n)_{ij} = \frac{1}{\sqrt{2}}(\alpha_{ij}^{(n)} + \mathbf{i}\beta_{ij}^{(n)})$ *for* $1 \leq i < j \leq n$*, and below-diagonal entries* $(W_n)_{ij} = \frac{1}{\sqrt{2}}(\alpha_{ij}^{(n)} - \mathbf{i}\beta_{ij}^{(n)})$ *for* $1 \leq j < i \leq n$*. We call* $\{W_n\}_{n\in\mathbb{N}}$ *the* **Gaussian Unitary Ensemble (GUE)***.*

Definition 1.5. *Let* $p \geq 1$*, and let* $\{\alpha_{ij}^{(n)}\}_{n\in\mathbb{N}, 1\leq i\leq n, 1\leq j\leq \lfloor pn\rfloor}$ *and* $\{\beta_{ij}^{(n)}\}_{n\in\mathbb{N}, 1\leq i\leq n, 1\leq j\leq \lfloor pn\rfloor}$ *be i.i.d. random variables, normally distributed with mean 0 and variance 1. Let* Y_n *be an* $n \times \lfloor pn \rfloor$ *matrix for each* n*, with entries* $(Y_n)_{ij} = \frac{1}{\sqrt{2}}(\alpha_{ij}^{(n)} + \mathbf{i}\beta_{ij}^{(n)})$*, and let* $M_n = Y_nY_n^*$*. We call* $\{M_n\}_{n\in\mathbb{N}}$ *the* **Wishart Ensemble** with parameter p.

We note that both families of matrices are Hermitian, and that the matrices of the Wishart Ensemble are, in addition, positive semi-definite. The GUE is relevant as a matrix model for nuclear energy levels under specific symmetries in Example 1.1, while the Wishart Ensemble is the model of interest in Examples 1.2 and 1.3.

All three of the examples above deal with random matrices of large dimensionality. In single-variate statistics, large collections of random variables are analyzed using limit theorems such as the Law of Large Numbers and the Central Limit Theorem. The goal of this paper is to develop similar limit theorems for spectral properties of interest for the GUE and Wishart Ensemble, as the matrix size tends to infinity. Our examples motivate the study of the limit theorems for two distinct spectral properties: the global distribution of eigenvalues, as relevant to Example 1.2, and the local statistics of the eigenvalue distribution, as relevant to Example 1.1. We will address these properties separately in the subsequent sections. We will see that the assumption of a normal distribution in Definitions 1.4 and 1.5 is not necessary for our study of the global eigenvalue distribution, but we will rely on this assumption when we turn to the examination of the local statistics.

A theme of this paper is the unified derivation of our results for the GUE and Wishart Ensemble, the studies of which had historical origins in different fields of application. Each section of the paper takes advantage of a similarity between the two families of matrices to derive a general result, which is then specialized to the GUE and Wishart cases. In Section 2, we prove the existence of a limit law for the global empirical distribution of eigenvalues for a class of general band matrices, using a combinatorial and graph theoretic approach. In Section 3, we use a change of variables formula to derive the form of the joint density function for eigenvalues of matrix distributions invariant under conjugation by unitary matrices. Finally, in Section 4, we compute a local correlation function for matrix eigenvalues in terms of orthogonal polynomials and derive a limit law for this function in the cases of the GUE and Wishart Ensemble with $p = 1$.

2 Convergence of the Empirical Distribution of Eigenvalues

We prove in this section the convergence in probability of the empirical distribution of eigenvalues for a class of general band matrices, and we specialize the result to the GUE and Wishart Ensemble by an explicit computation of the limit distribution. The result was first proven for the GUE by Wigner in [28] and [29], and we will follow Wigner's general strategy using the method of moments, with the generalizations provided in [2]. Our presentation draws on [1] and [2]. The class of band matrices we choose to work with is much more general than what is needed for the GUE and Wishart Ensemble—we do not require $\alpha_{ij}^{(n)}$ and $\beta_{ij}^{(n)}$ from Definitions 1.4 and 1.5 to have normal distributions, or even to be identically distributed. Specifically, we will consider a class of matrices X_n according to the following definition:

Definition 2.1. *Let $s : [0,1] \times [0,1] \to [0,\infty)$ be a symmetric bounded measurable function with discontinuity set of measure zero. Let $\{\xi_{ij}^{(n)}\}_{n\in\mathbb{N},1\leq i,j\leq n}$ be a collection of random variables with the following properties:*

1. *$\xi_{ii}^{(n)}$ is real-valued for all i and $\xi_{ij}^{(n)}$ is complex-valued for all $i \neq j$.*
2. *For each n, $\{\xi_{ij}^{(n)}\}_{1\leq i\leq j\leq n}$ is independent.*
3. *For all $i \neq j$, $\xi_{ij}^{(n)} = \overline{\xi_{ji}^{(n)}}$.*
4. *$\mathbb{E}[\xi_{ij}^{(n)}] = 0$ for all i, j.*
5. *$\mathbb{E}\left[\left|\xi_{ij}^{(n)}\right|^2\right] = s\left(\frac{i}{n}, \frac{j}{n}\right)$ for all i, j.*
6. *For each $k \in \mathbb{N}$, $\sup_{n,i,j} \mathbb{E}\left[\left|\xi_{ij}^{(n)}\right|^k\right] < \infty$.*

For each $n \in \mathbb{N}$, let X_n be the matrix with $X_n(i,j) = \frac{1}{\sqrt{n}}\xi_{ij}^{(n)}$.

In particular, the GUE scaled by $\frac{1}{\sqrt{n}}$ satisfies this definition with $s \equiv 1$. The primary concern of this section is the following notion of an empirical eigenvalue distribution:

Definition 2.2. *Suppose that X_n has eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$. The* **empirical distribution of eigenvalues** *of X_n is the counting measure $L_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}$. The* **mean empirical distribution of eigenvalues** *is $\overline{L_n} = \mathbb{E}L_n$.*

With this notation, L_n is a random measure over our probability space such that $L_n(S)$ is the fraction of eigenvalues of X_n contained in the set S. $\overline{L_n}$ is a fixed measure determined by our matrix class $\{X_n\}$, so that $\overline{L_n}(S)$ is the expected fraction of eigenvalues in S for a random matrix from our matrix class. The central result of this section is the following proposition:

Proposition 2.3. *There exists a measure μ of bounded support, symmetric about 0, and uniquely defined by its moments, such that L_n converges weakly, in probability, to μ. That is, for any $f \in C_b(\mathbb{R})$ and any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int f\, dL_n - \int f\, d\mu\right| > \varepsilon\right) = 0.$$

Taking f to be a continuous approximation of an indicator function for an interval S, this proposition tells us that for large n, $L_n(S) \approx \int_S d\mu$ with high probability, i.e. the fraction of eigenvalues of X_n contained in S is approximately the integral of a fixed density function over S. Visually, we may view this result as stating that a histogram of eigenvalues of X_n will, with high probability, converge in shape to the graph of a fixed density function. We will prove Proposition 2.3 in two steps in Sections 2.1 and 2.2 and explicitly compute the limit measure μ for the GUE and Wishart Ensemble in Section 2.3.

2.1 Convergence in moment of the mean empirical distribution $\overline{L_n}$

Our proof of Proposition 2.3 is based on a computation of the moments of $\overline{L_n}$ using a combinatorial analysis. We will use the following definitions:

Definition 2.4. *An* **n-path** *of length k is a $(k+1)$-tuple of indices $\mathbf{i} = (i_1, i_2, \ldots, i_{k+1}) \in \{1, \ldots, n\}^{k+1}$. The path is* **closed** *if $i_{k+1} = i_1$. Let $l(\mathbf{i}) = k$ be the length and $w(\mathbf{i}) = \#\{i_1, \ldots, i_{k+1}\}$ be the number of distinct indices of the path $\mathbf{i}$. Let us consider the path $\mathbf{i} = (i_1)$ of a single index to be a closed path with $l(\mathbf{i}) = 0$ and $w(\mathbf{i}) = 1$. For any ordered pair (i, i') of indices, let $f_{\mathbf{i}}(i, i') = \#\{j \mid i_j = i, i_{j+1} = i'\}$. Let $b_{\mathbf{i}}(i, i') = \#\{j \mid i_j = i', i_{j+1} = i\}$ if $i \neq i'$ and $b_{\mathbf{i}}(i, i') = 0$ if $i = i'$.*

Definition 2.5. *A* **Wigner n-path** *is a closed n-path $\mathbf{i}$ with the following properties:*

1. *For each $j = 1, \ldots, k$, $f_{\mathbf{i}}(i_j, i_{j+1}) = 1$ and $b_{\mathbf{i}}(i_j, i_{j+1}) = 1$.*
2. *$w(\mathbf{i}) = \frac{l(\mathbf{i})}{2} + 1$.*

We may think of an n-path of length k as a walk of k steps along the edges of the complete undirected graph of n vertices. Then $f_{\mathbf{i}}(i, i')$ and $b_{\mathbf{i}}(i, i')$ are the numbers of times the walk traverses the edge $i \to i'$ in the forward and backward directions respectively. (If $i = i'$, we count a traversal of this self-loop as a traversal in the forward direction only.) The first condition of the definition of a Wigner n-path specifies that the path traverses no self-loops and traverses each edge along the path exactly twice, once in each direction. Then the number of undirected edges traversed by the Wigner n-path is $\frac{l(\mathbf{i})}{2}$, and hence the second condition specifies that the subgraph of undirected edges traversed by the Wigner n-path is a tree.

We consider an equivalence relation on paths under permutation of the vertex labels:

Definition 2.6. *Two paths* $\mathbf{i}$ *and* $\mathbf{i}'$ *are* **equivalent** *if there is a bijection* $\varphi : \mathbb{N} \to \mathbb{N}$ *that maps* $\mathbf{i}$ *to* $\mathbf{i}'$ *when applied componentwise. Let* $[\mathbf{i}]$ *denote the equivalence class of* $\mathbf{i}$*. As paths in an equivalence class have the same length and number of distinct indices, let* $l([\mathbf{i}]) = l(\mathbf{i})$ *and* $w([\mathbf{i}]) = w(\mathbf{i})$*. Let* $\mathcal{C}_k$ *be the set of equivalence classes of closed paths of length* k *and* $\mathcal{W}_k \subset \mathcal{C}_k$ *be the set of equivalence classes of Wigner paths of length* k*.*

We note that this notion of equivalence does not make reference to the size n of the index set, so an equivalence class $C \in \mathcal{C}_k$ contains an infinite number of paths, and the n-paths belonging to C will be denoted as $C \cap \{1, \ldots, n\}^{k+1}$.

Lemma 2.7. *The moments of* $\overline{L_n}$ *satisfy*

$$\lim_{n\to\infty} \int x^k d\overline{L_n} = \lim_{n\to\infty} \frac{1}{n^{k/2+1}} \sum_{C\in\mathcal{W}_k} \sum_{\mathbf{i}\in C\cap\{1,\ldots,n\}^{k+1}} \prod_{j=1}^{k} s\left(\tfrac{i_j}{n}, \tfrac{i_{j+1}}{n}\right)^{1/2}, \tag{1}$$

if the limit on the right exists. In particular, $\lim_{n\to\infty} \int x^k d\overline{L_n} = 0$ *if* k *is odd.*

Proof. Letting $\lambda_1 \leq \ldots \leq \lambda_n$ be the eigenvalues of X_n,

$$\begin{aligned}
\int x^k d\overline{L_n} &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \lambda_i^k\right] \\
&= \mathbb{E}\left[\frac{1}{n} \operatorname{tr} X_n^k\right] \\
&= \frac{1}{n} \sum_{i_1,\ldots,i_k=1}^{n} \mathbb{E}[X_n(i_1,i_2)X_n(i_2,i_3)\ldots X_n(i_k,i_1)] \\
&= \frac{1}{n^{k/2+1}} \sum_{C\in\mathcal{C}_k} \sum_{\mathbf{i}\in C\cap\{1,\ldots,n\}^{k+1}} \mathbb{E}\left[\xi^{(n)}_{i_1 i_2} \ldots \xi^{(n)}_{i_k i_1}\right].
\end{aligned}$$

If $E_{\mathbf{i}} = \big\{(i,i') \mid i \leq i', (i,i') = (i_j, i_{j+1}) \text{ or } (i',i) = (i_j,i_{j+1}) \text{ for some } j\big\}$ is the set of edges traversed by any closed path $\mathbf{i}$, then

$$\mathbb{E}\left[\xi^{(n)}_{i_1 i_2} \ldots \xi^{(n)}_{i_k i_1}\right] = \prod_{(i,i')\in E_{\mathbf{i}}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')} \left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')}\right]$$

by the independence condition on $\{\xi^{(n)}_{ij}\}_{i\leq j}$. Therefore, since each $\xi^{(n)}_{ij}$ has mean zero, $\mathbb{E}[\xi^{(n)}_{i_1 i_2} \ldots \xi^{(n)}_{i_k i_1}] = 0$ for any $\mathbf{i}$ that traverses an edge only once. Also, by the bounded moment condition on the $\xi^{(n)}_{ij}$,

$$\left|\prod_{(i,i')\in E_{\mathbf{i}}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')} \left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')}\right]\right| \leq \prod_{(i,i')\in E_{\mathbf{i}}} \mathbb{E}\left[\left|\xi^{(n)}_{ii'}\right|^{f_{\mathbf{i}}(i,i')+b_{\mathbf{i}}(i,i')}\right] < A_{[\mathbf{i}]}$$

for some constant $A_{[\mathbf{i}]}$ depending on the equivalence class of $\mathbf{i}$. If C is an equivalence class of paths for which $l(C) = k$ and $w(C) = m$, then there are $n(n-1)\ldots(n-m+1)$ paths in $C \cap \{1,\ldots,n\}^{k+1}$. Thus, if $m < \frac{k}{2} + 1$, then

$$\lim_{n\to\infty} \left| \frac{1}{n^{k/2+1}} \sum_{\mathbf{i}\in C\cap\{1,\ldots,n\}^{k+1}} \mathbb{E}\left[\xi^{(n)}_{i_1 i_2} \ldots \xi^{(n)}_{i_k i_1}\right] \right| \leq \lim_{n\to\infty} \frac{n(n-1)\ldots(n-m+1)}{n^{k/2+1}} \cdot A_C = 0.$$

On the other hand, if each edge in $E_{\mathbf{i}}$ is traversed at least twice, then $\mathbf{i}$ traverses at most $\frac{k}{2}$ edges, and hence $m \leq \frac{k}{2} + 1$. Equality holds when each edge is traversed exactly twice and

$w(\mathbf{i}) = \frac{l(\mathbf{i})}{2} + 1$, i.e. the subgraph traversed by $\mathbf{i}$ is a tree. A path that traverses each edge of a tree exactly twice must traverse each edge once in either direction, so $\mathbf{i}$ is a Wigner path. Thus

$$\lim_{n\to\infty} \int x^k d\overline{L_n} = \lim_{n\to\infty} \frac{1}{n^{k/2+1}} \sum_{C\in\mathcal{W}_k} \sum_{\mathbf{i}\in C\cap\{1,\dots,n\}^{k+1}} \prod_{(i,i')\in E_{\mathbf{i}}} \mathbb{E}\left[\xi_{ii'}^{(n)}\xi_{ii'}^{(n)*}\right],$$

if the limit on the right exists. As $\mathbb{E}\left[\xi_{ii'}^{(n)}\xi_{ii'}^{(n)*}\right] = \mathbb{E}\left[\left|\xi_{ii'}^{(n)}\right|^2\right] = s\left(\frac{i}{n}, \frac{i'}{n}\right)$, this gives the desired result. If k is odd, then $\mathcal{W}_k$ is empty, so the limit is 0. □

Let us recall the following solution to the classical Hamburger moment problem, whose proof can be found in [23]:

Lemma 2.8. *Given a sequence of real values $\{m_k\}_{k=0}^{\infty}$, there exists a Borel measure μ with moments $\int_{-\infty}^{\infty} x^k d\mu = m_k$ if and only if the $k \times k$ Hankel matrix $\{h_{ij}\}_{0\le i,j\le k-1}$ given by $h_{ij} = m_{i+j}$ is positive semi-definite for each k.*

These lemmas are sufficient to show that the mean empirical eigenvalue distributions $\overline{L_n}$ converge in moment to a fixed measure μ:

Proposition 2.9. *There exists a Borel measure μ of bounded support, symmetric about 0, such that for all $k \ge 0$,*

$$\lim_{n\to\infty} \int x^k d\overline{L_n} = \int x^k d\mu.$$

Proof. Let us fix $C \in \mathcal{W}_k$ with $m = \frac{k}{2} + 1 = w(C)$. Let $\mathbf{i}^C \in C$ be the Wigner path such that the distinct indices of $\mathbf{i}^C$, in the order in which they are visited, are $1, 2, \dots, m$. Define a function $f_C : [0,1]^m \to [0,\infty)$ by $f_C(x_1,\dots,x_m) = \prod_{j=1}^k s(x_{i_j^C}, x_{i_{j+1}^C})^{1/2}$ if $k \ge 1$, or $f_C(x_1) = 1$ if $k = 0$. Then f_C is bounded with discontinuity set of measure zero, because the same properties hold for s and $i_j^C \ne i_{j+1}^C$ for any j if $\mathbf{i}^C$ is a Wigner path. So f_C is Riemann integrable, and we have

$$\begin{aligned}\int_0^1 \cdots \int_0^1 f_C(x_1,\dots,x_m)dx_1\dots dx_m &= \lim_{n\to\infty} \frac{1}{n^m} \sum_{x_1=1}^n \cdots \sum_{x_m=1}^n f_C\left(\tfrac{x_1}{n},\dots,\tfrac{x_m}{n}\right) \\ &= \lim_{n\to\infty} \frac{1}{n^m} \sum_{\substack{x_1,\dots,x_m=1\\ x_1\ne\dots\ne x_m}}^n f_C\left(\tfrac{x_1}{n},\dots,\tfrac{x_m}{n}\right),\end{aligned}$$

where the second equality holds because f_C is bounded. But by the definition of f_C,

$$\sum_{\substack{x_1,\dots,x_m=1\\ x_1\ne\dots\ne x_m}}^n f_C\left(\tfrac{x_1}{n},\dots,\tfrac{x_m}{n}\right) = \sum_{\mathbf{i}\in C\cap\{1,\dots,n\}^{k+1}} \prod_{j=1}^k s\left(\tfrac{i_j}{n}, \tfrac{i_{j+1}}{n}\right)^{1/2},$$

which is the quantity appearing on the right hand side of (1) in Lemma 2.7. Thus the limit in Lemma 2.7 exists, and we have

$$\lim_{n\to\infty} \int x^k d\overline{L_n} = \sum_{C\in\mathcal{W}_k} \mathbb{E}[f_C]$$

for even k, where $\mathbb{E}[f_C]$ denotes the average value of f_C over $[0,1]^m$. For each n, the $k \times k$ Hankel matrix of moments for $\overline{L_n}$ as defined in Lemma 2.8 is positive semi-definite for all k. As all moments of $\overline{L_n}$ converge as $n \to \infty$, the limits of these $k \times k$ Hankel matrices as $n \to \infty$ exist and must also be positive semi-definite for all k. Hence there exists a Borel measure μ whose moments are the limits of those of $\overline{L_n}$, by Lemma 2.8.

To show that μ has bounded support, let us compute the size of $\mathcal{W}_k$. Consider the map $\varphi : \mathcal{W}_k \to \mathbb{Z}^{k+1}$ such that $\varphi([\mathbf{i}])_1 = 0$, and $\varphi([\mathbf{i}])_{j+1} = \varphi([\mathbf{i}])_j + 1$ if (i_j, i_{j+1}) is the first traversal of that edge in $\mathbf{i}$ and $\varphi([\mathbf{i}])_{j+1} = \varphi([\mathbf{i}])_j - 1$ if (i_j, i_{j+1}) is the second traversal of that edge in $\mathbf{i}$, for each $j = 1, \ldots, k$. Then $\varphi([\mathbf{i}])_{k+1} = 0$ and $\varphi([\mathbf{i}])_j \geq 0$ for all j, i.e. $\varphi([\mathbf{i}])$ is a Dyck path. It is straightforward to verify that φ is a bijection between Wigner paths and Dyck paths, the latter enumerated by the Catalan numbers. Hence, for even k, $|\mathcal{W}_k| = \frac{1}{k/2+1}\binom{k}{k/2} \leq 2^k$. Together with the bound $\mathbb{E}[f_C] \leq (\|s\|_\infty^{1/2})^k$, we obtain that $\lim_{n\to\infty} \int x^k d\overline{L_n} \leq \left(2(\|s\|_\infty^{1/2})\right)^k$, so μ is supported on $\left[-2(\|s\|_\infty^{1/2}), 2(\|s\|_\infty^{1/2})\right]$. It is symmetric about 0 because its odd moments are 0 by Lemma 2.7. □

2.2 Convergence of the empirical distribution L_n

To conclude the proof of Proposition 2.3, we extend our notion of path equivalence to an equivalence relation on pairs of paths:

Definition 2.10. *Two ordered pairs of paths $(\mathbf{i}, \mathbf{i}')$ and $(\mathbf{j}, \mathbf{j}')$ are equivalent if there is a bijection $\varphi : \mathbb{N} \to \mathbb{N}$ that maps $\mathbf{i}$ to $\mathbf{j}$ and $\mathbf{i}'$ to $\mathbf{j}'$ when applied componentwise. Let $[(\mathbf{i}, \mathbf{i}')]$ denote the equivalence class of $(\mathbf{i}, \mathbf{i}')$. Let $\mathcal{P}_k$ be the set of equivalence classes of pairs of closed paths of length k.*

The following lemma shows convergence in probability of the moments of L_n to those of $\overline{L}_n$:

Lemma 2.11. *For all $\varepsilon > 0$ and $k \geq 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int x^k dL_n - \int x^k d\overline{L_n}\right| \geq \varepsilon\right) = 0.$$

Proof. As $\mathbb{E}\left[\int x^k dL_n\right] = \int x^k d\overline{L_n}$, Chebyshev's inequality gives

$$\begin{aligned}
\mathbb{P}\left(\left|\int x^k dL_n - \int x^k d\overline{L_n}\right| \geq \varepsilon\right) &\leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\left(\int x^k dL_n - \int x^k d\overline{L_n}\right)^2\right] \\
&= \frac{1}{\varepsilon^2}\left(\mathbb{E}\left[\left(\int x^k dL_n\right)^2\right] - \left(\int x^k d\overline{L_n}\right)^2\right) \\
&= \frac{1}{\varepsilon^2}\left(\mathbb{E}\left[(\tfrac{1}{n}\operatorname{tr} X_n^k)^2\right] - \mathbb{E}\left[\tfrac{1}{n}\operatorname{tr} X_n^k\right]^2\right) \\
&= \frac{1}{\varepsilon^2 n^2}\sum_{\substack{i_1,\ldots,i_k=1\\ i'_1,\ldots,i'_k=1}}^{n}\Big(\mathbb{E}[X_n(i_1,i_2)\ldots X_n(i_k,i_1)X_n(i'_1,i'_2)\ldots X_n(i'_k,i'_1)] \\
&\quad - \mathbb{E}[X_n(i_1,i_2)\ldots X_n(i_k,i_1)]\mathbb{E}[X_n(i'_1,i'_2)\ldots X_n(i'_k,i'_1)]\Big) \\
&= \frac{1}{\varepsilon^2 n^{k+2}}\sum_{C\in\mathcal{P}_k}\sum_{(\mathbf{i},\mathbf{i}')\in C\cap\{1,\ldots,n\}^{k+1}\times\{1,\ldots,n\}^{k+1}}\Big(\mathbb{E}\left[\xi^{(n)}_{i_1i_2}\ldots\xi^{(n)}_{i_ki_1}\xi^{(n)}_{i'_1i'_2}\ldots\xi^{(n)}_{i'_ki'_1}\right] \\
&\quad - \mathbb{E}\left[\xi^{(n)}_{i_1i_2}\ldots\xi^{(n)}_{i_ki_1}\right]\mathbb{E}\left[\xi^{(n)}_{i'_1i'_2}\ldots\xi^{(n)}_{i'_ki'_1}\right]\Big).
\end{aligned} \tag{2}$$

If $E_{\mathbf{i}} = \left\{(i,i') \mid i \leq i', (i,i') = (i_j, i_{j+1}) \text{ or } (i',i) = (i_j, i_{j+1}) \text{ for some } j\right\}$ is the set of edges traversed by any closed path $\mathbf{i}$, then

$$\mathbb{E}\left[\xi^{(n)}_{i_1i_2}\ldots\xi^{(n)}_{i_ki_1}\right] = \prod_{(i,i')\in E_{\mathbf{i}}}\mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')}\left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')}\right]$$

and

$$\mathbb{E}\left[\xi^{(n)}_{i_1 i_2}\ldots\xi^{(n)}_{i_k i_1}\xi^{(n)}_{i'_1 i'_2}\ldots\xi^{(n)}_{i'_k i'_1}\right] = \prod_{(i,i')\in E_{\mathbf{i}}\cup E_{\mathbf{i}'}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')+f_{\mathbf{i}'}(i,i')}\left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')+b_{\mathbf{i}'}(i,i')}\right]$$

by the independence condition on $\{\xi^{(n)}_{ij}\}_{i\leq j}$. Hence for any pair of paths $(\mathbf{i},\mathbf{i}')$ that together traverse an edge only once, the corresponding term in the sum in (2) is 0, and any other pair of paths traverses at most k distinct undirected edges. Also, for any pair of paths $(\mathbf{i},\mathbf{i}')$ for which $E_{\mathbf{i}}$ and $E_{\mathbf{i}'}$ are disjoint, the corresponding term in (2) is 0 since

$$\begin{aligned}&\prod_{(i,i')\in E_{\mathbf{i}}\cup E_{\mathbf{i}'}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')+f_{\mathbf{i}'}(i,i')}\left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')+b_{\mathbf{i}'}(i,i')}\right]\\ &= \prod_{(i,i')\in E_{\mathbf{i}}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}}(i,i')}\left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}}(i,i')}\right]\prod_{(i,i')\in E_{\mathbf{i}'}} \mathbb{E}\left[\left(\xi^{(n)}_{ii'}\right)^{f_{\mathbf{i}'}(i,i')}\left(\xi^{(n)*}_{ii'}\right)^{b_{\mathbf{i}'}(i,i')}\right].\end{aligned}$$

Any pair of paths for which $E_{\mathbf{i}}$ and $E_{\mathbf{i}'}$ are not disjoint is such that the undirected edges of $E_{\mathbf{i}}\cup E_{\mathbf{i}'}$ form a connected subgraph. Hence any equivalence class C of closed path pairs that contributes to the sum in (2) defines a pair of paths with at most $k+1$ distinct indices, and so $\#\{(\mathbf{i},\mathbf{i}')\in C\}\leq n(n-1)\ldots(n-k)$. Together with the bounded moment condition on $\xi^{(n)}_{ij}$, this implies

$$\mathbb{P}\left(\left|\int x^k dL_n - \int x^k d\overline{L_n}\right| \geq \varepsilon\right) \leq \frac{n(n-1)\ldots(n-k)}{\varepsilon^2 n^{k+2}}\cdot \max_{[\mathbf{i},\mathbf{i}']\in\mathcal{P}_k} A_{[\mathbf{i},\mathbf{i}']}$$

for a collection of bounds $A_{[\mathbf{i},\mathbf{i}']}$. Taking the limit as $n\to\infty$ gives the desired result. □

Proof of Proposition 2.3. By Proposition 2.9, $\overline{L_n}$ converges in moment to a measure μ, symmetric about 0, of bounded support. Suppose that μ is supported on $[-C,C]$, and set $B>\max(1,C^2)\geq C$. Given any bounded continuous function f, there exists a polynomial Q such that $g=f-Q$ satisfies $\sup_{|x|\leq B}|g(x)|<\varepsilon/8$ by the Weierstrass approximation theorem. Then we have

$$\begin{aligned}\left|\int f\,dL_n-\int f\,d\mu\right| &\leq \left|\int g\mathbb{1}_{|x|\leq B}\,dL_n-\int g\mathbb{1}_{|x|\leq B}\,d\mu\right| + \left|\int g\mathbb{1}_{|x|>B}\,dL_n\right| + \left|\int Q\,dL_n-\int Q\,d\mu\right|\\ &\leq \frac{\varepsilon}{4} + \left|\int g\mathbb{1}_{|x|>B}\,dL_n\right| + \left|\int Q\,dL_n-\int Q\,d\overline{L_n}\right| + \left|\int Q\,d\overline{L_n}-\int Q\,d\mu\right|,\end{aligned}$$

so

$$\begin{aligned}\mathbb{P}\left(\left|\int f\,dL_n-\int f\,d\mu\right|>\varepsilon\right) &\leq \mathbb{P}\left(\left|\int g\mathbb{1}_{|x|>B}\,dL_n\right|>\frac{\varepsilon}{4}\right)+\mathbb{P}\left(\left|\int Q\,dL_n-\int Q\,d\overline{L_n}\right|>\frac{\varepsilon}{4}\right)\\ &\quad+\mathbb{P}\left(\left|\int Q\,d\overline{L_n}-\int Q\,d\mu\right|>\frac{\varepsilon}{4}\right).\end{aligned}$$

As f is bounded, Q is a polynomial, and $B>1$, there exists some k such that $g(x)\leq|x|^k$ on $|x|>B$. Then

$$\mathbb{P}\left(\left|\int g\mathbb{1}_{|x|>B}\,dL_n\right|>\frac{\varepsilon}{4}\right)\leq\mathbb{P}\left(\left|\int |x|^k\mathbb{1}_{|x|>B}\,dL_n\right|>\frac{\varepsilon}{4}\right)\leq\frac{4}{\varepsilon}\int|x|^k\mathbb{1}_{|x|>B}\,d\overline{L_n}$$

by Markov's inequality. Cauchy-Schwartz's inequality then gives

$$\int|x|^k\mathbb{1}_{|x|>B}\,d\overline{L_n}\leq\sqrt{\int x^{2k}d\overline{L_n}}\sqrt{\int\mathbb{1}_{|x|>B}\,d\overline{L_n}},$$

and Chebyshev's inequality gives

$$\int \mathbb{1}_{|x|>B}\, d\overline{L_n} = \overline{L_n}\left\{|x|^k > B^k\right\} \leq \frac{1}{B^{2k}} \int x^{2k} d\overline{L_n}.$$

Hence, by Proposition 2.9,

$$\limsup_{n\to\infty} \mathbb{P}\left(\left|\int |x|^k \mathbb{1}_{|x|>B}\, dL_n\right| > \frac{\varepsilon}{4}\right) \leq \lim_{n\to\infty} \frac{4\int x^{2k} d\overline{L_n}}{\varepsilon B^k} = \frac{4\int x^{2k} d\mu}{\varepsilon B^k} \leq \frac{4C^{2k}}{\varepsilon B^k}.$$

We note that $\mathbb{P}\left(\left|\int |x|^k \mathbb{1}_{|x|>B}\, dL_n\right| > \frac{\varepsilon}{4}\right)$ is increasing in k as $B > 1$, and $\lim_{k\to\infty} \frac{4C^{2k}}{\varepsilon B^k} = 0$ because $B > C^2$. Thus

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int g \mathbb{1}_{|x|>B}\, dL_n\right| > \frac{\varepsilon}{4}\right) = \lim_{n\to\infty} \mathbb{P}\left(\left|\int |x|^k \mathbb{1}_{|x|>B}\, dL_n\right| > \frac{\varepsilon}{4}\right) = 0.$$

By Lemma 2.11,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int Q\, dL_n - \int Q\, d\overline{L_n}\right| > \frac{\varepsilon}{4}\right) = 0,$$

and by Proposition 2.9,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int Q\, d\overline{L_n} - \int Q\, d\mu\right| > \frac{\varepsilon}{4}\right) = 0.$$

Putting this together gives the desired result,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int f\, dL_n - \int f\, d\mu\right| > \varepsilon\right) = 0.$$

We note that as the weak limit μ of $\overline{L_n}$ must be unique and the only information we used regarding μ was the values of its moments, μ must be uniquely defined by its moments. □

2.3 Computation of the limit distribution of L_n for the GUE and Wishart Ensemble

Recall from the proof of Proposition 2.9 that the limit distribution μ has moments

$$\int x^k d\mu = \sum_{C\in\mathcal{W}_k} \mathbb{E}[f_C]$$

where, for each $C \in \mathcal{W}_k$, $f_C : [0,1]^{k/2+1} \to [0,\infty)$ is the function

$$f_C(x_1,\ldots,x_{k/2+1}) = \prod_{j=1}^{k} s\left(x_{i_j^C}, x_{i_{j+1}^C}\right)^{1/2} \tag{3}$$

for $\mathbf{i}^C = (i_1^C,\ldots,i_{k+1}^C)$ the path in C whose distinct indices are $1,2,\ldots,\frac{k}{2}+1$ in the order visited (or $f_C(x_1) = 1$ if $k = 0$). If we suppose that $s \equiv 1$ as in the case of the GUE, this immediately gives the k^{th} moment of μ as $|\mathcal{W}_k|$, which we computed in the proof of Proposition 2.9 to be the Catalan number $C_{k/2}$. This was the original argument of Wigner in [28] and specifies the limit measure μ. To achieve greater generality, we will follow the steps of [2] and first prove the following proposition, which will be of use in the computation of the limit distribution μ for Wishart matrices:

Proposition 2.12. *For each $x \in [0,1]$, define a formal power series*

$$\Phi(x,t) = \sum_{k=0}^{\infty}\left(\sum_{C\in\mathcal{W}_k} \mathbb{E}[f_C(x_1,\ldots,x_{k/2+1})|x_1 = x]\right) t^{k+1}.$$

Then $\{\Phi(x,t)\}_{x\in[0,1]}$ is the unique collection of formal power series with constant term 0 and linear coefficient 1 for all x that satisfies the identity

$$\Phi(x,t) = t\left(1 - t\int_0^1 s(x,y)\Phi(y,t)dy\right)^{-1}, \tag{4}$$

where the right-hand-side is short-hand for its formal power series expansion in t.

Proof. Let us denote $\tilde{f}_C(x) = \mathbb{E}[f_C(x_1,\ldots,x_{k/2+1})|x_1 = x]$. For any $C \in \mathcal{W}_k$ with $l(C) = k > 0$, let $\mathbf{i} = (i_1, i_2, \ldots, i_k, i_1)$ be a path in C and let $j = \max\{j \le k \mid i_j = i_1\}$. Then $i_{j'} \neq i_1$ for any $j < j' \le k$ and, viewing $\mathbf{i}$ as the traversal of a tree, we see that $i_{j+1} = i_k$. Let $C_1 \in \mathcal{W}_{j-1}$ and $C_2 \in \mathcal{W}_{k-j-1}$ be the equivalence classes of the paths $(i_1, i_2, \ldots, i_j)$ and $(i_{j+1}, i_{j+2}, \ldots, i_k)$ respectively. (We note that we might have $j = 1$ and $l(C_1) = 0$.) Then, by the definition of f in (3), we have the recursive identity

$$\tilde{f}_C(x) = \int_0^1 s(x, x_{j+1})\tilde{f}_{C_1}(x)\tilde{f}_{C_2}(x_{j+1})dx_{j+1}. \tag{5}$$

As the correspondence between $\mathcal{W}_k$ and $\bigcup_{j=1}^{k-1}\mathcal{W}_{j-1}\times\mathcal{W}_{k-j-1}$ sending C to (C_1, C_2) is a bijection, we may sum over all equivalence classes of Wigner paths of length k to obtain

$$\sum_{C\in\mathcal{W}_k}\tilde{f}_C(x) = \sum_{j=1}^{k-1}\sum_{C_1\in\mathcal{W}_{j-1}}\sum_{C_2\in\mathcal{W}_{k-j-1}}\int_0^1 s(x,y)\tilde{f}_{C_1}(x)\tilde{f}_{C_2}(y)dy.$$

Then

$$\begin{aligned}
&\Phi(x,t)\left(1 - t\int_0^1 s(x,y)\Phi(y,t)dy\right)\\
&= \sum_{k=0}^{\infty}\left(\sum_{C\in\mathcal{W}_k}\tilde{f}_C(x)\right)t^{k+1}\cdot\left(1 - t\int_0^1 s(x,y)\sum_{k=0}^{\infty}\left(\sum_{C\in\mathcal{W}_k}\tilde{f}_C(y)\right)t^{k+1}dy\right)\\
&= \sum_{k=0}^{\infty}\left[\sum_{C\in\mathcal{W}_k}\tilde{f}_C(x) - \sum_{j=1}^{k-1}\sum_{C_1\in\mathcal{W}_{j-1}}\sum_{C_2\in\mathcal{W}_{k-j-1}}\int_0^1 s(x,y)\tilde{f}_{C_1}(x)\tilde{f}_{C_2}(y)dy\right]t^{k+1}\\
&= t,
\end{aligned}$$

since for $k = l(C) = 0$, $f_C(x_1) = 1$ for all x_1. Rearranging gives equation (4). To see that this family of power series is unique, we note that if $(1 + a_1t + a_2t^2 + a_3t^3 + \ldots)^{-1} = 1 + b_1t + b_2t^2 + \ldots$, then each $b_i = -a_1b_{i-1} - a_2b_{i-2} - \ldots - a_i$ is defined as a function of $\{a_1,\ldots,a_i,b_1,\ldots,b_{i-1}\}$. Hence, equation (4) specifies the coefficient of t^k in $\Phi(x,t)$ as a function of the coefficients of t^j in $\Phi(y,t)$ for all $y \in [0,1]$ and $j \le k-2$. So fixing the condition that the constant term of $\Phi(x,t)$ is 0 and the linear coefficient is 1 for all $x \in [0,1]$ uniquely specifies $\Phi(x,t)$, and the linear coefficient of $\Phi(x,t)$ is indeed 1 as it is equal to $\tilde{f}_C(x)$ for the zero-length path C. □

2.3.1 The GUE and the semicircle law

Proposition 2.12, along with the fact that the k^{th} moment of the limit distribution μ is given by the coefficient of t^{k+1} in $\int_0^1\Phi(x,t)dx$, allows us to compute μ for specific matrix classes. We will see that for the GUE, μ is the following semicircle law supported on $[-2,2]$:

Lemma 2.13. *Consider the semicircle law $\sigma(x)dx$ given by the density function*

$$\sigma(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbb{1}_{|x|\le 2}.$$

Then the odd moments of $\sigma(x)dx$ are 0 and the even moments are

$$\int x^{2k}\sigma(x)dx = C_k$$

where C_k is the k^{th} Catalan number.

Proof. For all k, integration by parts gives

$$\int_{-\pi/2}^{\pi/2} \sin^{2k}\theta\cos^2\theta\,d\theta = \int_{-\pi/2}^{\pi/2} (\cos\theta)\left(\sin^{2k}\theta\cos\theta\,d\theta\right) = \frac{1}{2k+1}\int_{-\pi/2}^{\pi/2}\sin^{2k+2}\theta\,d\theta.$$

Substituting $x = 2\sin\theta$ gives

$$\begin{aligned}\int x^{2k}\sigma(x)dx &= \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}dx = \frac{2^{2k+1}}{\pi}\int_{-\pi/2}^{\pi/2}\sin^{2k}\theta\cos^2\theta\,d\theta\\ &= \frac{2^{2k+1}}{(2k+1)\pi}\int_{-\pi/2}^{\pi/2}\sin^{2k+2}\theta\,d\theta.\end{aligned}$$

On the other hand,

$$\begin{aligned}\int x^{2k}\sigma(x)dx &= \frac{2^{2k+1}}{\pi}\int_{-\pi/2}^{\pi/2}\sin^{2k}\theta\cos^2\theta\,d\theta\\ &= \frac{2^{2k+1}}{\pi}\left(\int_{-\pi/2}^{\pi/2}\sin^{2k}\theta\,d\theta - \int_{-\pi/2}^{\pi/2}\sin^{2k+2}\theta\,d\theta\right)\\ &= \frac{2^{2k+1}}{\pi}\int_{-\pi/2}^{\pi/2}\sin^{2k}\theta\,d\theta - (2k+1)\int x^{2k}\sigma(x)dx,\end{aligned}$$

so

$$\begin{aligned}\int x^{2k}\sigma(x)dx &= \frac{2^{2k+1}}{(2k+2)\pi}\int_{-\pi/2}^{\pi/2}\sin^{2k}\theta\,d\theta = \frac{4(2k-1)}{2k+2}\int x^{2k-2}\sigma(x)dx\\ &= \frac{2k(2k-1)}{(k+1)k}\int x^{2k-2}\sigma(x)dx.\end{aligned}$$

This recursion with the initial value $\int \sigma(x)dx = 1$ gives the result

$$\int x^{2k}\sigma(x)dx = \frac{(2k)!}{k!(k+1)!} = \frac{\binom{2k}{k}}{k+1} = C_k.$$

□

Then we have the following theorem:

Theorem 2.14. *Suppose $\int_0^1 s(x,y)dy = 1$ for all $x \in [0,1]$. Then L_n converges weakly, in probability, to the semicircle law $\sigma(x)dx$.*

Proof. Given $C \in \mathcal{W}_k$, we define $\tilde{f}_C(x)$, C_1, and C_2 as in the proof of Proposition 2.12. Supposing that $\tilde{f}_{C_1}$ and $\tilde{f}_{C_2}$ are constant over $x \in [0,1]$, equation (5) gives

$$\tilde{f}_C(x) = \int_0^1 s(x,x_{j+1})\tilde{f}_{C_1}(x)\tilde{f}_{C_2}(x_{j+1})dx_{j+1} = \tilde{f}_{C_1}(x)\tilde{f}_{C_2}(x)\int_0^1 s(x,y)dy = \tilde{f}_{C_1}(x)\tilde{f}_{C_2}(x),$$

which is also constant over $x \in [0,1]$. As $\tilde{f}_C(x) = 1$ for $l(C) = 0$, induction on $l(C)$ gives that $\tilde{f}_C$ is constant over $x \in [0,1]$ for all equivalence classes C of Wigner paths. Then $\Phi(x,t) = \sum_{k=0}^{\infty}\sum_{C\in\mathcal{W}_k}\tilde{f}_C(x)t^{k+1}$ is independent of x, so we may write $\Phi(x,t) = \Theta(t)$.

Equation (4) in Proposition 2.12 thus becomes $\Theta(t) = t(1 - t\Theta(t))^{-1}$, which we may solve to obtain

$$\Theta(t) = \frac{1 - \sqrt{1 - 4t^2}}{2t} \tag{6}$$

as the solution whose power series has constant term 0 and linear coefficient 1. We note that, letting a_j be the coefficient of t^j in the power series expansion of $\frac{1-\sqrt{1-4t^2}}{2t}$, $\Theta(t) = t(1 - t\Theta(t))^{-1}$ gives $a_j = 0$ for even j, $a_1 = 1$, and the recurrence $a_{2k+1} = a_{2k-1}a_1 + a_{2k-3}a_3 + \ldots + a_1 a_{2k-1}$. Hence this defines a_{2k+1} as the k^{th} Catalan number. These are precisely the moments of the semicircle law, so the result follows from Proposition 2.12. □

Setting $s \equiv 1$ gives the result for the GUE:

Corollary 2.15 (Wigner [28, 29])**.** *Let $\{W_n\}_{n\in\mathbb{N}}$ be the GUE, and let L_n^W be the empirical eigenvalue distribution of $\frac{1}{\sqrt{n}}W_n$. Then L_n^W converges weakly, in probability, to the semicircle law $\sigma(x)dx$.*

2.3.2 The Wishart Ensemble and the law of Marčenko-Pastur

We conclude the discussion of convergence of the empirical eigenvalue distribution by using Proposition 2.12 to compute the limit distribution μ for the Wishart Ensemble.

Lemma 2.16. *Let $p \geq 1$, let $A = [0, \frac{1}{1+p})$, and let $B = [\frac{1}{1+p}, 1]$. Suppose $s \equiv 0$ on $A \times A$ and $B \times B$, so that for each n, X_n is of the form*

$$X_n = \begin{bmatrix} 0 & Y_n \\ Y_n^* & 0 \end{bmatrix}$$

where Y_n is a $\lfloor \frac{n}{1+p} \rfloor \times \lceil \frac{pn}{1+p} \rceil$ matrix. Suppose furthermore that

$$\int_A s(x, y)dy = \mathbb{1}_B(x), \int_B s(x, y)dy = p\mathbb{1}_A(x).$$

Then L_n converges weakly, in probability, to a measure μ defined by

$$\int f\, d\mu = Kf(0) + \frac{1}{(p+1)\pi} \int_{\left|\frac{y^2-p-1}{\sqrt{p}}\right| \leq 2} \frac{f(y)^{2k}\sqrt{4p - (y^2 - p - 1)^2}}{|y|} dy \tag{7}$$

for some constant K.

Proof. Given $C \in \mathcal{W}_k$, we define $\tilde{f}_C(x)$, C_1, and C_2 as in the proof of Proposition 2.12. Supposing that $\tilde{f}_{C_1}(x) = \mathbb{1}_A(x)\tilde{f}_{C_1}^A + \mathbb{1}_B(x)\tilde{f}_{C_1}^B$ and $\tilde{f}_{C_2}(x) = \mathbb{1}_A(x)\tilde{f}_{C_2}^A + \mathbb{1}_B(x)\tilde{f}_{C_2}^B$ for constants $\tilde{f}_{C_1}^A, \tilde{f}_{C_1}^B, \tilde{f}_{C_2}^A, \tilde{f}_{C_2}^B$, equation (5) gives

$$\begin{aligned}
\tilde{f}_C(x) &= \int_0^1 s(x, x_{j+1})\tilde{f}_{C_1}(x)\tilde{f}_{C_2}(x_{j+1})dx_{j+1} \\
&= \int_0^1 s(x, y)\left(\mathbb{1}_A(x)\tilde{f}_{C_1}^A + \mathbb{1}_B(x)\tilde{f}_{C_1}^B\right)\left(\mathbb{1}_A(y)\tilde{f}_{C_2}^A + \mathbb{1}_B(y)\tilde{f}_{C_2}^B\right)dy \\
&= \left(\mathbb{1}_A(x)\tilde{f}_{C_1}^A + \mathbb{1}_B(x)\tilde{f}_{C_1}^B\right)\left(\tilde{f}_{C_2}^A \int_A s(x, y)dy + \tilde{f}_{C_2}^B \int_B s(x, y)dy\right) \\
&= \mathbb{1}_A(x)\tilde{f}_{C_1}^A\tilde{f}_{C_2}^B p + \mathbb{1}_B(x)\tilde{f}_{C_1}^B\tilde{f}_{C_2}^A.
\end{aligned}$$

As $\tilde{f}_C(x) = 1$ for $l(C) = 0$, induction on $l(C)$ gives that $\tilde{f}_C(x) = \mathbb{1}_A(x)\tilde{f}_C^A + \mathbb{1}_B(x)\tilde{f}_C^B$ for some constants $\tilde{f}_C^A, \tilde{f}_C^B$ for all equivalence classes C of Wigner paths. Then we may write $\Phi(x,t) = \mathbb{1}_A(x)\Phi_A(t) + \mathbb{1}_B(x)\Phi_B(t)$, and Equation (4) in Proposition 2.12 becomes

$$\mathbb{1}_A(x)\Phi_A(t) + \mathbb{1}_B(x)\Phi_B(t) = t\big(1 - \mathbb{1}_B(x)t\Phi_A(t) - p\mathbb{1}_A(x)t\Phi_B(t)\big)^{-1}.$$

Taking $x \in A$ and $x \in B$, we may separate this into a pair of equations $\Phi_A(t) = t(1 - pt\Phi_B(t))^{-1}$ and $\Phi_B(t) = t(1 - t\Phi_A(t))^{-1}$, and solve for $\Phi_A(t)$ and $p\Phi_B(t)$ to obtain

$$\begin{aligned}
\Phi_A(t) &= \frac{1 + t^2 - pt^2 - \sqrt{(1 - t^2 + pt^2)^2 - 4pt^2}}{2t}, \\
p\Phi_B(t) &= \frac{1 - t^2 + pt^2 - \sqrt{(1 - t^2 + pt^2)^2 - 4pt^2}}{2t}, \\
\int \Phi(x,t)dx &= \tfrac{1}{p+1} \left(\Phi_A(t) + p\Phi_B(t)\right) \\
&= \frac{1 - \sqrt{(1 - t^2 + pt^2)^2 - 4pt^2}}{(p+1)t}.
\end{aligned}$$

We may write

$$(1 - t^2 + pt^2)^2 - 4pt^2 = 1 - 2(p+1)t^2 + (p-1)^2t^4 = \left(1 - (p+1)t^2\right)^2 - 4pt^4.$$

Let us set $\Theta(t) = \frac{1-\sqrt{1-4t^2}}{2t} = \sum_{k=0}^{\infty} C_k t^{2k+1}$ from (6) in the proof of Theorem 2.14, whose coefficients C_k are the moments of the semicircle law $\sigma(x)dx$. Then

$$\Theta\left(\frac{\sqrt{p}t^2}{1 - (p+1)t^2}\right) = \frac{1 - (p+1)t^2 - \sqrt{(1 - (p+1)t^2)^2 - 4pt^4}}{2\sqrt{p}t^2}.$$

Noting that $\frac{\sqrt{p}t^2}{1-(p+1)t^2} = \sqrt{p}(t^2 + (p+1)t^4 + (p+1)^2t^6 + \ldots)$ and expanding the composition of power series,

$$\begin{aligned}
\int \Phi(x,t)dx &= t\left(1 + \frac{2\sqrt{p}}{p+1}\,\Theta\left(\frac{\sqrt{p}t^2}{1 - (p+1)t^2}\right)\right) \\
&= t + \frac{2\sqrt{p}}{p+1} \sum_{k=1}^{\infty} \sum_{j=0}^{\lfloor \frac{k-1}{2} \rfloor} C_j p^{j+\frac{1}{2}} (p+1)^{k-1-2j} \binom{k-1}{2j} t^{2k+1},
\end{aligned}$$

where we have used the fact that the number of partitions of k into ordered $(2j+1)$-tuples of positive integers is $\binom{k-1}{2j}$. Then Proposition 2.12 implies that the limit measure μ of the empirical measures L_n of X_n has even moments

$$\begin{aligned}
\int x^{2k} d\mu &= \frac{2p}{p+1} \sum_{j=0}^{\lfloor \frac{k-1}{2} \rfloor} C_j p^j (p+1)^{k-1-2j} \binom{k-1}{2j} \\
&= \frac{p}{(p+1)\pi} \int_{-2}^{2} \sum_{j=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2j} x^{2j} p^j (p+1)^{k-1-2j} \sqrt{4 - x^2}\, dx \\
&= \frac{p}{(p+1)\pi} \int_{-2}^{2} \sum_{j=0}^{k-1} \binom{k-1}{j} x^j \sqrt{p}^j (p+1)^{k-1-j} \sqrt{4 - x^2}\, dx \\
&= \frac{p}{(p+1)\pi} \int_{-2}^{2} (\sqrt{p}x + p + 1)^{k-1} \sqrt{4 - x^2}\, dx
\end{aligned}$$

for $k \geq 1$. Making the substitutions $y = (\sqrt{p}x + p + 1)^{1/2}$ and $y = -(\sqrt{p}x + p + 1)^{1/2}$, we have

$$\begin{aligned}
\int x^{2k} d\mu &= \frac{p}{2(p+1)\pi} \int_{\left|\frac{y^2-p-1}{\sqrt{p}}\right| \leq 2} y^{2k-2} \sqrt{4 - \left(\frac{y^2 - p - 1}{\sqrt{p}}\right)^2} \cdot \frac{2|y|}{\sqrt{p}} dy \\
&= \frac{1}{(p+1)\pi} \int_{\left|\frac{y^2-p-1}{\sqrt{p}}\right| \leq 2} \frac{y^{2k} \sqrt{4p - (y^2 - p - 1)^2}}{|y|} dy.
\end{aligned}$$

If we set K so that

$$\frac{1}{(p+1)\pi}\int_{\left|\frac{y^2-p-1}{\sqrt{p}}\right|\le 2}\frac{\sqrt{4p-(y^2-p-1)^2}}{|y|}dy = 1-K,$$

then we note that the measure defined as in (7) has the same moments as the limit measure μ of L_n, so the result follows from Proposition 2.3. □

This lemma allows us to compute the limit empirical eigenvalue distribution for the Wishart Ensemble:

Theorem 2.17 (Marčenko, Pastur [17]). *Let $\{M_n\}_{n\in\mathbb{N}}$ be the Wishart Ensemble with parameter p, and let L_n^M be the empirical measure of $\frac{1}{n}M_n$. Set $a=(\sqrt{p}-1)^2$ and $b=(\sqrt{p}+1)^2$. Then L_n^M converges weakly, in probability, to the Marčenko-Pastur law $\mu_{MP}(x)dx$ with density*

$$\mu_{MP}(x)=\frac{\sqrt{(x-a)(b-x)}}{2\pi x}\mathbb{1}_{[a,b]}.$$

Proof. Suppose $M_n = Y_nY_n^*$ as in Definition 1.5, and let

$$X_{n+\lfloor pn\rfloor}=\begin{bmatrix} 0 & \frac{1}{\sqrt{n}}Y_n \\ \frac{1}{\sqrt{n}}Y_n^* & 0\end{bmatrix}.$$

Let $A=[0,\frac{1}{p+1})$ and $B=[\frac{1}{p+1},1]$. Then we note that the conditions in Definition 2.1 are satisfied for $X_{n+\lfloor pn\rfloor}$ with $s(x,y)=0$ if $(x,y)\in(A\times A)\cup(B\times B)$ and $s(x,y)=p+1$ if $(x,y)\in(A\times B)\cup(B\times A)$. Thus, by Lemma 2.16, the empirical eigenvalue measures $L_{n+\lfloor pn\rfloor}$ of $X_{n+\lfloor pn\rfloor}$ converge weakly, in probability, to the distribution μ given by (7).

We note that if Y_n has rank r, then M_n has rank r and $X_{n+\lfloor pn\rfloor}$ has rank $2r$. If $X_{n+\lfloor pn\rfloor}\begin{bmatrix}v_1\\ v_2\end{bmatrix}=\begin{bmatrix}\lambda v_1\\ \lambda v_2\end{bmatrix}$, then $X_{n+\lfloor pn\rfloor}\begin{bmatrix}v_1\\ -v_2\end{bmatrix}=\begin{bmatrix}-\lambda v_1\\ \lambda v_2\end{bmatrix}$ and $\frac{1}{n}M_nv_1=\frac{1}{\sqrt{n}}Y_n\cdot\frac{1}{\sqrt{n}}Y_n^*v_1=\lambda^2v_1$. Hence the $2r$ nonzero eigenvalues of $X_{n+\lfloor pn\rfloor}$ come in oppositely signed pairs, and the r nonzero eigenvalues of $\frac{1}{n}M_n$ are the squares of those of $X_{n+\lfloor pn\rfloor}$. This implies that for any $g\in C_b(\mathbb{R})$ with $g(0)=0$,

$$(n+\lfloor pn\rfloor)\cdot\int g(x^2)L_{n+\lfloor pn\rfloor}(dx)=2n\int g(x)L_n^M(dx).$$

Then

$$\begin{aligned}\int g(x)L_n^M(dx) &\xrightarrow{\mathbb{P}}\frac{p+1}{2}\int g(x^2)\mu(dx)\\ &=\frac{1}{2\pi}\int_{\left|\frac{x^2-p-1}{\sqrt{p}}\right|\le 2}\frac{g(x^2)\sqrt{4p-(x^2-p-1)^2}}{|x|}dx\\ &=\frac{1}{\pi}\int_{-2\sqrt{p}+p+1}^{2\sqrt{p}+p+1}\frac{g(y)\sqrt{4p-(y-p-1)^2}}{2y}dy.\end{aligned}$$

Setting $a=(\sqrt{p}-1)^2$ and $b=(\sqrt{p}+1)^2$, we can write this as

$$\int g(x)L_n^M(dx)\xrightarrow{\mathbb{P}}\int_a^b\frac{g(x)\sqrt{(x-a)(b-x)}}{2\pi x}dx.$$

We may check that the integral on the right evaluates to 1 for $g\equiv 1$, so in fact this holds for all $g\in\mathbb{C}_b(\mathbb{R})$. Thus L_n^M converges weakly, in probability, to $\mu_{\mathrm{MP}}(x)dx$. □

Let us remark before continuing that these results can be strengthened in various ways. It was shown in [17] using a Stieltjes-transform method that the convergence in Theorems 2.14 and 2.17 in fact hold almost surely. The condition of all moments finite in Definition 2.1 can be weakened; see [4].

3 Joint Distribution of Eigenvalues

The convergence of the empirical eigenvalue distribution provides us with a global histogram picture of the eigenvalues for the GUE and the Wishart Ensemble, and the combinatorial approach used in the previous section holds more generally for matrices whose entries are not normally distributed. If we make use of the normal distributions in Definitions 1.4 and 1.5, then we can obtain information that is much more precise regarding the eigenvalue distributions. Recall from the spectral theorem that any Hermitian matrix H can be factored as $H = UDU^*$, where U is a unitary matrix whose columns are eigenvectors of H, and D is a diagonal matrix containing the eigenvalues of H. If we parameterize U and D with real variables, we may view this as a change of variables formula from these variables to the entries of H. The distribution of matrix entries for H thus induces a distribution over our parameter space for U and D, and in particular this allows us to derive the joint distribution of eigenvalues of the matrix. We will detail these steps in this section and derive the joint distribution of eigenvalues for matrices invariant under unitary conjugation, including matrices in the GUE and the Wishart Ensemble. Our presentation draws from [1], [5], and [18].

Throughout this section, we will work with the following parametrization of the space of Hermitian matrices:

Definition 3.1. *Let $\mathcal{H}_n$ be the space of $n \times n$ Hermitian matrices, parameterized by the n^2 real variables $\alpha_{ii} = (H_n)_{ii}$ for $1 \le i \le n$ and $\alpha_{ij} = \sqrt{2}\,\mathrm{Re}(H_n)_{ij}$ and $\beta_{ij} = \sqrt{2}\,\mathrm{Im}(H_n)_{ij}$ for $i < j$. Let $\varphi : \mathbb{R}^{n^2} \to \mathcal{H}_n$ be this parametrization map, and endow $\mathcal{H}_n$ with the measure $dH_n = \prod_{i=1}^n d\alpha_{ii} \prod_{i<j} d\alpha_{ij} d\beta_{ij}$ induced by Lebesgue measure on the parameter space.*

Let us observe that

$$\mathrm{tr}\, H_n^2 = \sum_{i,j=1}^{n} H_{ij}H_{ji} = \sum_{i=1}^{n} H_{ii}^2 + 2 \sum_{1 \le i < j \le n} (\mathrm{Re}(H_{ij})^2 + \mathrm{Im}(H_{ij})^2) = \|\varphi^{-1}(H_n)\|_2^2$$

under this parametrization. We note that dH_n satisfies invariance under unitary conjugation in the following sense:

Proposition 3.2. *Let U_n be a unitary matrix and consider the map $H_n \to G_n = U_n H_n U_n^*$. Then $dG_n = dH_n$, i.e., the Jacobian of the transformation $\varphi^{-1}(H_n) \to \varphi^{-1}(G_n)$ has determinant ± 1.*

Proof. Let us denote $u : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ as the transformation $u(x) = \varphi^{-1}(U_n \varphi(x) U_n^*)$ that sends $\varphi^{-1}(H_n)$ to $\varphi^{-1}(G_n)$. As u is linear and $\|\varphi^{-1}(H_n)\|_2^2 = \mathrm{tr}\, H_n^2 = \mathrm{tr}\, G_n^2 = \|\varphi^{-1}(G_n)\|_2^2$, this shows that $\|x\|_2 = \|ux\|_2$ for all $x \in \mathbb{R}^{n^2}$. Hence u is orthogonal and $|\det u| = 1$. □

As a result, any distribution of random Hermitian matrices of the form $f(H)dH$ where f is dependent on only the eigenvalues of H is also invariant under unitary conjugation, and we will see that the GUE and Wishart Ensemble are examples of such distributions. For such matrices, we may use a change of variables idea to compute the joint distribution of eigenvalues. To make this idea rigorous, we will need the following technical lemma, taken from [1], regarding the parametrization of unitary matrices:

Lemma 3.3. *Let $\mathcal{U}_n$ be the group of $n \times n$ unitary matrices and $\mathcal{D}_n$ be the group of $n \times n$ real diagonal matrices. Let $\Lambda = \{\lambda \in \mathbb{R}^n \mid \lambda_1 > \ldots > \lambda_n\}$, and let $d : \Lambda \to \mathcal{D}_n$ map λ to the matrix with λ along the diagonal. Then there exists a set $O \subset \mathbb{R}^{n(n-1)}$ of full measure and a smooth, injective parametrization $p : O \to \mathcal{U}_n$ such that the map $\gamma : \Lambda \times O \to \mathcal{H}_n$ given by $\gamma(\lambda, x) = p(x)d(\lambda)p(x)^*$ is smooth and injective with image of full measure in $\mathcal{H}_n$.*

Proof. Let $\mathcal{V}_n \subset \mathcal{U}_n$ be the set of unitary matrices with nonzero leading principal minors and nonzero diagonal entries, and consider the map $q : \mathcal{V}_n \to \mathbb{R}^{n(n-1)}$ defined as

$$q\left(\begin{bmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \cdots & u_{nn} \end{bmatrix}\right) = \Bigg(\operatorname{Re}\frac{u_{12}}{u_{11}}, \operatorname{Im}\frac{u_{12}}{u_{11}}, \dots, \operatorname{Re}\frac{u_{1n}}{u_{11}}, \operatorname{Im}\frac{u_{1n}}{u_{11}}, \operatorname{Re}\frac{u_{23}}{u_{22}}, \operatorname{Im}\frac{u_{23}}{u_{22}}, \dots, \operatorname{Re}\frac{u_{2n}}{u_{22}}, \operatorname{Im}\frac{u_{2n}}{u_{22}}, \dots, \operatorname{Re}\frac{u_{n-1,n}}{u_{n-1,n-1}}, \operatorname{Im}\frac{u_{n-1,n}}{u_{n-1,n-1}}\Bigg).$$

Let O be the image of q. For any $x \in O$, set $v_{ii} = 1$ for $i = 1$ to n and $v_{ij} = \operatorname{Re}\frac{u_{ij}}{u_{ii}} + \mathbf{i}\operatorname{Im}\frac{u_{ij}}{u_{ii}}$ for $i < j$, and define recursively

$$(v_{i1}, \dots, v_{i,i-1}) = \begin{bmatrix} v_{11} & \cdots & v_{1,i-1} \\ \vdots & \ddots & \vdots \\ v_{i-1,1} & \cdots & v_{i-1,i-1} \end{bmatrix}^{-1} \begin{bmatrix} v_{1i} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{i-1,i} & \cdots & v_{i-1,n} \end{bmatrix} \begin{bmatrix} -v_{ii} \\ \vdots \\ -v_{in} \end{bmatrix}$$

for $i = 2, \dots, n$; this is the unique way to complete the matrix $V = (v_{ij})_{1 \le i,j \le n}$ to have orthogonal rows. Hence the i^{th} row of V is a scalar multiple of the i^{th} row of U, and so the inverse matrices used in the construction exist because U has nonzero leading principal minors. In fact, $\mathbb{R}^{n(n-1)} \setminus O$ is precisely the set such that one of the leading principal minors of V or a diagonal entry of V is 0. As the entries of V are rational functions of the entries of $q(U)$, $\mathbb{R}^{n(n-1)} \setminus O$ is Zariski closed in $\mathbb{R}^{n(n-1)}$ and hence has zero measure. For $x \in O$, let $p(x)_{ij} = v_{ij}/\|(v_{i1}, \dots, v_{in})\|_2$, so that $p(x) \in \mathcal{U}_n$. By the uniqueness of the construction of V, p is injective (and smooth). Indeed, we note that $p(q(U))$ is the matrix whose rows are multiples of those of U with positive real diagonal entries.

We see that γ is smooth since p is smooth, and it is injective since the parameter vector (x, λ) uniquely determines the eigenvalues and eigenvectors (up to scalar multiples) of X if the elements of λ are distinct. Consider the set $\mathcal{G}$ of matrices $X \in \mathcal{H}_n$ that can be factored as $X = UDU^*$, where U is unitary with all minors nonzero and D is diagonal such that $\prod_{i\in I} D_{ii} \neq \prod_{i \in J} D_{ii}$ for any two nonempty subsets $I, J \subset \{1, \dots, n\}$ of the same cardinality. In particular, U has nonzero entries and nonzero leading principal minors and D has distinct diagonal entries, and we may choose D such that the diagonal entries are strictly decreasing and U such that the diagonal entries are all positive real. Hence $\mathcal{G}$ is contained in the image of γ.

To see that $\mathcal{G}$ and hence the image of γ has full measure in $\mathcal{H}_n$, consider for each $X \in \mathcal{H}_n$ and each $r = 1, \dots, n$ the $\binom{n}{r} \times \binom{n}{r}$ matrix $X^{(r)}$ indexed by pairs of r-element subsets of $\{1, \dots, n\}$ such that $X^{(r)}_{IJ} = \det X_{IJ}$, the minor corresponding to the submatrix of rows indexed by I and columns indexed by J. We note that $X^{(r)*} = (X^*)^{(r)}$, so $X^{(r)}$ is Hermitian, and $(XY)^{(r)} = X^{(r)}Y^{(r)}$ as $\det(XY)_{IJ} = \sum_{K \subset \{1,\dots,n\}, |K|=r} \det X_{IK} \det Y_{KJ}$ by the Cauchy-Binet formula. Hence, if $X = UDU^*$, then we may factor $X^{(r)}$ as $U^{(r)}D^{(r)}U^{(r)*}$ where $U^{(r)}$ is unitary and $D^{(r)}$ is diagonal. The set $\mathcal{H}_n \setminus \mathcal{G}$ is characterized by the matrices such that two diagonal entries of $D^{(r)}$ are equal for some r or an element of $U^{(r)}$ is zero for some r.

The condition that $D^{(r)}$ has a repeated diagonal entry is the condition that the discriminant of the characteristic polynomial of $X^{(r)}$ is zero, which is a polynomial condition on the entries of $X^{(r)}$ and hence of X. Supposing that $X^{(r)}$ has distinct eigenvalues, let λ be one such eigenvalue with eigenvector v, let $A = X^{(r)} - \lambda I$, and let $A^{(i,j)}$ be A with row i and column j removed. Then $AA^{\text{adj}} = (\det A)I = 0$ where $A^{\text{adj}}_{ij} = (-1)^{i+j} \det A^{(i,j)}$, and A has kernel spanned by v so each column of A^{adj} is a nonzero multiple of v. Then $v_i = 0$ if and only if $A^{\text{adj}}_{ii} = 0$, which holds if and only if λ is also an eigenvalue of $A^{(i,i)}$. Hence an entry of

$U^{(r)}$ is zero if and only if A and $A^{(i,i)}$ have a common eigenvalue for some i, which is the condition that the resultant of the characteristic polynomials of A and $A^{(i,i)}$ is 0. This is a polynomial condition on the entries of A and hence of $X^{(r)}$ and X. So the set $\mathcal{H}_n \setminus \mathcal{G}$ corresponds to a Zariski closed subset of the parameter space $\{\alpha_{ij}\}_{i\leq j} \cup \{\beta_{ij}\}_{i<j}$ of $\mathcal{H}_n$ and hence has zero measure in $\mathcal{H}_n$. □

Proposition 3.4. *There exists a constant C such that for any function $f : \mathcal{H}_n \to \mathbb{R}$ that can be written as $f(H_n) = g(\lambda_1, \ldots, \lambda_n)$, depending only on the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ of H_n, we have*

$$\int f(H_n) dH_n = C \int_{\lambda_1 \geq \ldots \geq \lambda_n} g(\lambda_1, \ldots, \lambda_n) \prod_{i<j} (\lambda_i - \lambda_j)^2 d\lambda_1 \ldots d\lambda_n.$$

Proof. Extending $g : \mathbb{R}^n \to \mathbb{R}$ to $\tilde{f} : \mathbb{R}^{n^2} \to \mathbb{R}$ given by $\tilde{f}(\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_{n(n-1)}) = g(\lambda_1, \ldots, \lambda_n)$, we have by the preceding lemma that

$$\int_{\{\lambda_1 > \ldots > \lambda_n\} \times O} \tilde{f} |D(\varphi^{-1} \circ \gamma)| d\lambda_1 \ldots d\lambda_n dx_1 \ldots dx_{n(n-1)} = \int f(H_n) dH_n$$

for a full-measure subset $O \subset \mathbb{R}^{n(n-1)}$ and smooth, injective parametrization $\gamma(\lambda, x) = p(x) d(\lambda) p(x)^*$, where $p(x)$ is unitary. Consider the linear map $r_x : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ such that $r_x(v) = \varphi^{-1}(p(x)^* \varphi(v) p(x))$. As $\|v\|_2^2 = \operatorname{tr} \varphi(v)^2$ and $\operatorname{tr}(p(x)^* \varphi(v) p(x))^2 = \operatorname{tr} p(x)^* \varphi(v)^2 p(x) = \operatorname{tr} \varphi(v)^2$, r_x is orthogonal and $|\det r_x| = 1$. Then $|D(\varphi^{-1} \circ \gamma(\lambda, x))| = |r_x \cdot D(\varphi^{-1} \circ \gamma(\lambda, x))|$.

We have, for each k, the matrix equation

$$\frac{\partial \gamma(\lambda, x)}{\partial \lambda_k} = p(x) \frac{\partial d(\lambda)}{\partial \lambda_k} p(x)^*,$$

and hence the column $r_x \cdot \varphi^{-1} \left(\frac{\partial \gamma(\lambda, x)}{\partial \lambda_k} \right)$ of $r_x \cdot D(\varphi^{-1} \circ \gamma)$ is given by $\alpha_{kk} = 1$ and $\alpha_{ij}, \beta_{ij} = 0$ for all other i, j. We also have, for each k, the matrix equation

$$\frac{\partial \gamma(\lambda, x)}{\partial x_k} = \frac{\partial p(x)}{\partial x_k} d(\lambda) p(x)^* + p(x) d(\lambda) \left(\frac{\partial p(x)}{\partial x_k} \right)^* = p(x)(s_k(x) d(\lambda) + d(\lambda) s_k(x)^*) p(x)^*$$

for $s_k(x) = p(x)^* \frac{\partial p(x)}{\partial x_k}$. We note that $p(x)^* p(x) = I$ for all x, so $s_k(x) = -s_k(x)^*$. Then $s_k(x) d(\lambda) + d(\lambda) s_k(x)^* = s_k(x) d(\lambda) - d(\lambda) s_k(x)$, and the column $r_x \cdot \varphi^{-1} \left(\frac{\partial \gamma(\lambda, x)}{\partial x_k} \right)$ is given by $x_{ii} = 0$ for all i, and $x_{ij} = (\lambda_j - \lambda_i) \operatorname{Re}(s_k(x)_{ij})$ and $y_{ij} = (\lambda_j - \lambda_i) \operatorname{Im}(s_k(x)_{ij})$ for all $i < j$. Putting this together, we have that

$$|D(\varphi^{-1} \circ \gamma(\lambda, x))| = |r_x \cdot D(\varphi^{-1} \circ \gamma(\lambda, x))| = c(x) \cdot \prod_{i<j} (\lambda_j - \lambda_i)^2$$

for some function c depending only on x. Then

$$\begin{aligned} \int f(H_n) dH_n &= \int_{\lambda_1 > \ldots > \lambda_n} g(\lambda_1, \ldots, \lambda_n) \prod_{i<j} (\lambda_i - \lambda_j)^2 d\lambda_1 \ldots d\lambda_n \int_O c(x) dx_1 \ldots dx_{n(n-1)} \\ &= C \int_{\lambda_1 \geq \ldots \geq \lambda_n} g(\lambda_1, \ldots, \lambda_n) \prod_{i<j} (\lambda_i - \lambda_j)^2 d\lambda_1 \ldots d\lambda_n \end{aligned}$$

for a constant $C = \int_O c(x) dx_1 \ldots dx_{n(n-1)}$. □

This proposition gives us the joint density function of eigenvalues for any matrix ensemble exhibiting this type of unitary invariance.

3.1 Joint eigenvalue density for the GUE

By Definitions 1.4 and 3.1, we see that W_n from the GUE has the distribution

$$\frac{1}{(2\pi)^{\frac{n^2}{2}}}\exp\left(-\frac{1}{2}\left(\sum_{1\le i\le j\le n}\alpha_{ij}^2+\sum_{1\le i<j\le n}\beta_{ij}^2\right)\right)\prod_{1\le i\le j\ne n}d\alpha_{ij}\prod_{1\le i<j\le n}d\beta_{ij}=\frac{1}{(2\pi)^{\frac{n^2}{2}}}e^{-\frac{1}{2}\operatorname{tr}H_n^2}dH_n.$$

This gives the following result for the GUE as an immediate corollary of Proposition 3.4:

Theorem 3.5 (Ginibre [12]). *Let W_n be the $n\times n$ matrix of the GUE, with eigenvalues $\lambda_1\ge\ldots\ge\lambda_n$. Then the joint distribution of these eigenvalues is given by the density function*

$$\hat{\rho}(\lambda_1,\ldots,\lambda_n)=C\cdot\mathbb{1}_{\lambda_1\ge\ldots\ge\lambda_n}\prod_{1\le i<j\le n}(\lambda_i-\lambda_j)^2\prod_{i=1}^n e^{-\frac{1}{2}\lambda_i^2},$$

where

$$C=\left(\int_{\lambda_1\ge\ldots\ge\lambda_n}\prod_{1\le i<j\le n}(\lambda_i-\lambda_j)^2\prod_{i=1}^n e^{-\frac{1}{2}\lambda_i^2}d\lambda_i\right)^{-1}.$$

3.2 Joint eigenvalue density for the Wishart Ensemble

In the case of the Wishart Ensemble, we must first derive the form of the distribution for the matrix entries. This result was first obtained by Wishart for real matrices in [30] and extended to the complex Wishart distribution by Goodman in [14]; our proof follows the ideas from that of the real Wishart distribution in [3].

Proposition 3.6. *Let $\mathcal{P}_n$ be the space of positive-definite Hermitian matrices. Then M_n, the $n\times n$ matrix of the Wishart Ensemble with parameter p, has the distribution*

$$C\cdot\mathbb{1}_{H_n\in\mathcal{P}_n}(\det H_n)^{\lfloor pn\rfloor-n}e^{-\operatorname{tr}H_n}dH_n$$

for a constant C.

Proof. Let $m=\lfloor pn\rfloor$. Suppose $M_n=Y_nY_n^*$ as in Definition 1.5, and let Y_n have rows $v_1,\ldots,v_n$ with $v_i\in\mathbb{C}^m$ for each i. Taking the standard inner product $\langle v,w\rangle=v^*w$, let $w_1,\ldots,w_n$ be the orthonormal vectors obtained through Gram-Schmidt orthogonalization on $v_1,\ldots,v_n$, i.e., $w_1=v_1/\|v_1\|$ and

$$w_i=\frac{v_i-\sum_{j=1}^{i-1}\langle v_i,w_j\rangle w_j}{\|v_i-\sum_{j=1}^{i-1}\langle v_i,w_j\rangle w_j\|}$$

for $i=2,\ldots,n$. This gives us the factorization $Y_n=T_nU_n$, where T_n is an $n\times n$ lower-triangular matrix and U_n has orthonormal rows $w_1,\ldots,w_n$. Hence $M_n=Y_nY_n^*=T_nU_nU_n^*T_n^*=T_nT_n^*$. The entries of T_n are given by $t_{ij}=\langle v_i,w_j\rangle$ for $i>j$ and $t_{ii}=\|v_i-\sum_{j=1}^{i-1}\langle v_i,w_j\rangle w_j\|\ge 0$.

We note that each v_i is of the form $\frac{1}{\sqrt{2}}(x_{i1}+\mathbf{i}y_{i1},\ldots,x_{im}+\mathbf{i}y_{im})$, where each x_{ij} and y_{ij} has a standard normal distribution. Then the joint density function for $\{x_{ij}\}_{j=1}^m\cup\{y_{ij}\}_{j=1}^m$ is given by

$$\frac{1}{(2\pi)^m}\exp\left(-\frac{1}{2}\sum_{j=1}^m x_{ij}^2+y_{ij}^2\right).$$

This is radially symmetric in the $\mathbb{C}^m$-vector $\frac{1}{\sqrt{2}}(x_{i1}+\mathbf{i}y_{i1},\ldots,x_{im}+\mathbf{i}y_{im})$, and hence the distribution of v_i is invariant under any unitary change of basis. This implies that, if $v_1,\ldots,v_{i-1}$

are given, then $t_{i1}, \ldots, t_{i,i-1}$ are conditionally independent with the same distribution as the entries of Y_n, and t_{ii} is conditionally independent from $t_{i1}, \ldots, t_{i,i-1}$ with the same distribution as $\|(x_{ii} + \mathbf{i}y_{ii}, \ldots, x_{im} + \mathbf{i}y_{im})\|$. The former distribution is given by $\frac{1}{\sqrt{2}}(x + \mathbf{i}y)$ with x and y independent standard normal. As $\|(x_{ii} + \mathbf{i}y_{ii}, \ldots, x_{im} + \mathbf{i}y_{im})\|^2 = \frac{1}{2}\sum_{j=i}^m x_{ij}^2 + y_{ij}^2$, $2t_{ii}^2$ has the χ^2-distribution with $2(m - i + 1)$ degrees of freedom, so the latter distributions are given by the densities $C_i t_{ii}^{2m-2i+1} e^{-t_{ii}^2}$ for constants C_i. As these distributions do not depend on $v_1, \ldots, v_{i-1}$, we in fact obtain that $\{t_{ij}\}_{i\geq j}$ are unconditionally independent with these distributions. Hence, if we parameterize T_n by $\{u_{ii}\}_{i=1}^n = \{t_{ii}\}_{i=1}^n$, $\{u_{ij}\}_{i>j} = \sqrt{2}\,\mathrm{Re}\{t_{ij}\}_{i>j}$, and $\{v_{ij}\}_{i>j} = \sqrt{2}\,\mathrm{Im}\{t_{ij}\}_{i>j}$, then the joint density of these parameters is given by

$$C \prod_{i=1}^n u_{ii}^{2m-2i+1} e^{-u_{ii}^2} \prod_{i>j} e^{-\frac{1}{2}(u_{ij}^2 + v_{ij}^2)} = C \prod_{i=1}^n u_{ii}^{2m-2i+1} e^{-\operatorname{tr} T_n T_n^*}$$

for a constant $C > 0$.

Let us parameterize $\mathcal{H}_n$ by α_{ij} and β_{ij} as in Definition 3.1. Then the map $T_n \to M_n = T_n T_n^*$ is given by

$$\alpha_{ii} = \frac{1}{2}\sum_{k<i}(u_{ik}^2 + v_{ik}^2) + u_{ii}^2$$

for $i = 1$ to n and

$$\frac{1}{\sqrt{2}}(\alpha_{ij} + \mathbf{i}\beta_{ij}) = \frac{1}{2}\sum_{k<i}(u_{ik} + \mathbf{i}v_{ik})(u_{jk} - \mathbf{i}v_{jk}) + \frac{1}{\sqrt{2}}u_{ii}(u_{ji} - \mathbf{i}v_{ji}),$$

or equivalently,

$$\alpha_{ij} = \frac{1}{\sqrt{2}}\sum_{k<i}(u_{ik}u_{jk} + v_{ik}v_{jk}) + u_{ii}u_{ji},$$

$$\beta_{ij} = \frac{1}{\sqrt{2}}\sum_{k<i}(u_{jk}v_{ik} - u_{ik}v_{jk}) - u_{ii}v_{ji}$$

for $i < j$. If we order the parameters for T_n in the order $u_{11}, u_{21}, v_{21}, u_{31}, v_{31}, \ldots, u_{n1}, v_{n1}, u_{22}, u_{32}, v_{32}, \ldots, u_{nn}$ and the parameters for M_n in the order $\alpha_{11}, \alpha_{12}, \beta_{12}, \alpha_{13}, \beta_{13}, \ldots, \alpha_{1n}, \beta_{1n}, \alpha_{22}, \alpha_{23}, \beta_{23}, \ldots, \alpha_{nn}$, then the Jacobian of this map is the determinant of an upper-triangular matrix with diagonal entries $\frac{\partial\alpha_{ii}}{\partial u_{ii}} = 2u_{ii}$, $\frac{\partial\alpha_{ij}}{\partial u_{ji}} = u_{ii}$, and $\frac{\partial\beta_{ij}}{\partial v_{ji}} = -u_{ii}$. Hence the Jacobian has absolute value $2^n \prod_{i=1}^n u_{ii}^{2n-2i+1}$. As $\prod_{i=1}^n u_{ii} = \det T_n = (\det M_n)^{1/2}$, the distribution of M_n is given by $C(\det H_n)^{m-n} e^{-\operatorname{tr} H_n} dH_n$ for a constant C, over the range of the map $T_n \to M_n$, i.e. the positive definite matrices H_n. □

This gives the following result as an immediate corollary of Proposition 3.4; the analogous result for real matrices was first obtained independently in [11], [13], [15], and [22]:

Theorem 3.7. *Let M_n be the $n \times n$ matrix of the Wishart Ensemble with parameter p, with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$. Then the joint distribution of these eigenvalues is given by the density function*

$$\hat{\rho}(\lambda_1, \ldots, \lambda_n) = C \cdot \mathbb{1}_{\lambda_1 \geq \ldots \geq \lambda_n \geq 0} \prod_{1\leq i<j\leq n}(\lambda_i - \lambda_j)^2 \prod_{i=1}^n \lambda_i^{\lfloor pn\rfloor - n} e^{-\lambda_i},$$

where

$$C = \left(\int_{\lambda_1 \geq \ldots \geq \lambda_n \geq 0} \prod_{1\leq i<j\leq n}(\lambda_i - \lambda_j)^2 \prod_{i=1}^n \lambda_i^{\lfloor pn\rfloor - n} e^{-\lambda_i} d\lambda_i\right)^{-1}.$$

4 Convergence of the Local Eigenvalue Correlation Function

The formulas for the joint density of eigenvalues from the previous section allow us to study the local statistics of eigenvalues in the GUE and Wishart Ensemble. We note that the density functions are not symmetric because they assume that the inputs are ordered, and we may symmetrize as follows:

Definition 4.1. *Suppose that an $n \times n$ random matrix has joint density of (ordered) eigenvalues given by $\hat{\rho}(x_1, \ldots, x_n) = \mathbb{1}_{x_1 \geq \ldots \geq x_n} f(x_1, \ldots, x_n)$ for a symmetric function f. Its* **joint density of unordered eigenvalues** *is $\rho(x_1, \ldots, x_n) = \frac{1}{n!} f(x_1, \ldots, x_n)$.*

This joint density of unordered eigenvalues satisfies the properties that, for any function f and permutation σ of $\{1, \ldots, n\}$, we have

$$\int_{\mathbb{R}^n} f(x_1, \ldots, x_n) \rho(x_1, \ldots, x_n) dx_1 \ldots dx_n = \int_{\mathbb{R}^n} f(\sigma(x_1), \ldots, \sigma(x_n)) \rho(x_1, \ldots, x_n) dx_1 \ldots dx_n, \tag{8}$$

and for any symmetric function f,

$$\int_{\mathbb{R}^n} f(x_1, \ldots, x_n) \rho(x_1, \ldots, x_n) dx_1 \ldots dx_n = \int_{x_1 \geq \ldots \geq x_n} f(x_1, \ldots, x_n) \hat{\rho}(x_1, \ldots, x_n) dx_1 \ldots dx_n. \tag{9}$$

The focus of our study of local eigenvalue statistics of the GUE and Wishart Ensemble will be the following k-point eigenvalue correlation functions, introduced by Dyson in [6] and [7]:

Definition 4.2. *The k-point eigenvalue correlation function of an $n \times n$ random matrix X, for $1 \leq k \leq n$, is*

$$R_k(x_1, \ldots, x_k) = \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} \rho(x_1, \ldots, x_n)\, dx_{k+1} \ldots dx_n.$$

We note that these are simply scaled marginal densities of k unordered eigenvalues, and we may interpret $R_k(x_1, \ldots, x_k)$ as an "expectation density" of finding eigenvalues close to $x_1, \ldots, x_k$ in the following sense:

Proposition 4.3. *Given sets $A_1, \ldots, A_k \subset \mathbb{R}$, $\int_{A_1} \cdots \int_{A_k} R_k(x_1, \ldots, x_k)\, dx_1 \ldots dx_k$ is the expected number of ordered k-tuples of distinct eigenvalues $(\lambda_{i_1}, \ldots, \lambda_{i_k})$ of X such that $\lambda_{i_j} \in A_j$ for each j from 1 to k.*

Proof. The proof is straightforward from equations (8) and (9). Indeed,

$$\begin{aligned}
&\int_{A_1} \cdots \int_{A_k} R_k(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \\
&= \int_{\mathbb{R}^k} \mathbb{1}_{A_1}(x_1) \ldots \mathbb{1}_{A_k}(x_k) R_k(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \\
&= \frac{n!}{(n-k)!} \int_{\mathbb{R}^n} \mathbb{1}_{A_1}(x_1) \ldots \mathbb{1}_{A_k}(x_k) \rho(x_1, \ldots, x_n)\, dx_1 \ldots dx_n \\
&= \int_{\mathbb{R}^n} \sum_{\sigma \in \Sigma_{k,n}} \mathbb{1}_{A_1}(x_{\sigma(1)}) \ldots \mathbb{1}_{A_k}(x_{\sigma(k)}) \rho(x_1, \ldots, x_n)\, dx_1 \ldots dx_n \\
&= \int_{x_1 \geq \ldots \geq x_n} \sum_{\sigma \in \Sigma_{k,n}} \mathbb{1}_{A_1}(x_{\sigma(1)}) \ldots \mathbb{1}_{A_k}(x_{\sigma(k)}) \hat{\rho}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n,
\end{aligned}$$

where $\Sigma_{k,n}$ is the set of all ordered k-tuples of distinct indices from 1 to n, and the last quantity is our desired expectation. ☐

The formula for $\rho(x_1, \ldots, x_n)$ can, in the case of matrices exhibiting unitary invariance, be expressed in terms of polynomials orthogonal with respect to certain weight functions. This allows us to write the k-point eigenvalue correlation functions R_k in a relatively simple determinantal form, which we will derive in Section 4.1. This will then translate the problem of determining a limit law for R_k into one of determining the asymptotic growth for these orthogonal polynomials, and we will use the asymptotic properties of the specific polynomials associated to the GUE and Wishart Ensemble to derive limit laws for the bulk of the spectrum in Section 4.2. Our presentation draws on [1], [5], [18], [20], and [24].

4.1 Properties of the k-point eigenvalue correlation functions

In this section, let us consider an $n \times n$ random matrix X with joint density of unordered eigenvalues

$$\rho(x_1, \ldots, x_n) = C \prod_{1 \le i < j \le n} |x_i - x_j|^2 \prod_{i=1}^{n} w(x_i), \tag{10}$$

for a weight function $w : \mathbb{R} \to \mathbb{R}_{\ge 0}$ such that $\int |x^k| w(x) dx < \infty$ for all k. For polynomials p, let us denote $\int p(x)^2 w(x) dx$ as the norm-squared of p with respect to w, and for two polynomials p and q, let us say that they are orthogonal with respect to w if $\int p(x) q(x) w(x)\, dx = 0$. We will use the following notation:

Definition 4.4. *Given a function $K : \mathbb{R}^2 \to \mathbb{R}$, denote*

$$\det_{i,j=1}^{k} K(x_i, x_j) = \begin{vmatrix} K(x_1, x_1) & \cdots & K(x_1, x_k) \\ \vdots & \ddots & \vdots \\ K(x_k, x_1) & \cdots & K(x_k, x_k) \end{vmatrix}$$

for $k \ge 1$. Denote

$$\det(I - K)_{[a,b]} = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{[a,b]^k} \det_{i,j=1}^{k} K(x_i, x_j)\, dx_1 \ldots dx_k.$$

The central result of this section is that we may express the k-point eigenvalue correlation function as a determinant of a matrix of values related to orthogonal polynomials:

Proposition 4.5 (Mehta, Gaudin [19])**.** *Suppose the joint density of unordered eigenvalues of X is given by equation (10), and let R_k be its k-point eigenvalue correlation function. Let $\{p_k\}_{k=0}^{\infty}$ be monic polynomials orthogonal with respect to w, with p_k of degree k and norm-squared c_k with respect to w. Then for each k from 1 to n,*

$$R_k(x_1, \ldots, x_k) = \det_{i,j=1}^{k} K(x_i, x_j) \tag{11}$$

where

$$K(x, y) = \sum_{j=0}^{n-1} \frac{\sqrt{w(x) w(y)}\, p_j(x) p_j(y)}{c_j}.$$

Proof. We show by induction that $R_k(x_1, \ldots, x_k) = n! C \cdot c_0 \ldots c_{n-1} \det_{i,j=1}^{k} K(x_i, x_j)$ (where C is the constant in the joint eigenvalue density of X). For the base case $k = n$,

$$R_n(x_1, \ldots, x_n) = n! \rho(x_1, \ldots, x_n) = n! C \prod_{1 \le i < j \le n} |x_i - x_j|^2 \prod_{i=1}^{n} w(x_i).$$

We note that $\prod_{1\le i<j\le n} |x_i - x_j| = \left|\det_{i,j=1}^{n} x_i^{j-1}\right|$ is the van der Monde determinant, and $\det_{i,j=1}^{n} x_i^{j-1} = \det_{i,j=1}^{n} p_{j-1}(x_i)$ because our polynomials p_j are monic and adding a multiple of one row of a matrix to another does not change the determinant. Hence

$$R_n(x_1,\dots,x_n) = n!C\cdot c_0\dots c_{n-1}\left(\det_{i,j=1}^{n}\frac{\sqrt{w(x_i)}p_{j-1}(x_i)}{\sqrt{c_{j-1}}}\right)^2.$$

We then note that

$$\left(\det_{i,j=1}^{n} f(x_i,x_j)\right)^2 = \left(\det_{i,j=1}^{n} f(x_i,x_j)\right)\left(\det_{i,j=1}^{n} f(x_j,x_i)\right) = \left(\det_{i,j=1}^{n}\sum_{k=1}^{n} f(x_i,x_k)f(x_j,x_k)\right),$$

where the first equality uses $\det A = \det A^T$ and the second uses $(\det A)(\det B) = \det AB$. Thus

$$R_n(x_1,\dots,x_n) = n!C\cdot c_0\dots c_{n-1}\det_{i,j=1}^{n} K(x_i,x_j)$$

as desired. For the inductive step, suppose

$$\begin{aligned}R_{k-1}(x_1,\dots,x_{k-1}) &= \frac{1}{n-k+1}\int R_k(x_1,\dots,x_k)dx_k\\ &= \frac{n!C\cdot c_0\dots c_{n-1}}{n-k+1}\int \det_{i,j=1}^{k} K(x_i,x_j)\,dx_k.\end{aligned}$$

We may expand the determinant as

$$\det_{i,j=1}^{k} K(x_i,x_j) = \sum_{\sigma\in\Sigma_{k,k}} \operatorname{sgn}(\sigma)K(x_1,x_{\sigma(1)})\dots K(x_k,x_{\sigma(k)}),$$

where $\Sigma_{k,k}$ is the set of permutations on $\{1,\dots,k\}$. For each j from 1 to k, there is a bijection $f_j : \{\sigma\in\Sigma_{k,k} \mid \sigma(j)=k\}\to\Sigma_{k-1,k-1}$ such that $f_j(\sigma)(j)=\sigma(k)$ and $f_j(\sigma)(i)=\sigma(i)$ for all $i\ne j$. We note that $\operatorname{sgn} f_j(\sigma) = -\operatorname{sgn}\sigma$ if $j\ne k$ and $\operatorname{sgn} f_k(\sigma)=\operatorname{sgn}\sigma$. As

$$\begin{aligned}\int K(x,y)K(y,z)dy &= \int\sqrt{w(x)w(z)}w(y)\left(\sum_{j=0}^{n-1}\frac{p_j(x)p_j(y)}{c_j}\right)\left(\sum_{j=0}^{n-1}\frac{p_j(y)p_j(z)}{c_j}\right)dy\\ &= \sqrt{w(x)w(z)}\sum_{i,j=0}^{n-1}\frac{p_i(x)p_j(z)}{c_ic_j}\int p_i(y)p_j(y)w(y)dy\\ &= K(x,z),\end{aligned}$$

where the last equality uses that $\int p_i(y)p_j(y)w(y)dy = 0$ if $i\ne j$ and $\int p_i(y)p_j(y)w(y)dy = c_i$ if $i=j$. Thus we have for $j\ne k$,

$$\begin{aligned}&\int\sum_{\sigma(j)=k}\operatorname{sgn}(\sigma)K(x_1,x_{\sigma(1)})\dots K(x_j,x_k)\dots K(x_k,x_{\sigma(k)})dx_k\\ &= \sum_{\sigma(j)=k}\operatorname{sgn}(\sigma)K(x_1,x_{\sigma(1)})\dots K(x_j,x_{\sigma(k)})\dots K(x_{k-1},x_{\sigma(k-1)})\\ &= -\sum_{\sigma\in\Sigma_{k-1,k-1}}\operatorname{sgn}(\sigma)K(x_1,x_{\sigma(1)})\dots K(x_{k-1},x_{\sigma(k-1)})\\ &= -\det_{i,j=1}^{k-1} K(x_i,x_j).\end{aligned}$$

On the other hand, as

$$\int K(x,x)dx = \int \sum_{j=0}^{n-1} \frac{w(x)p_j(x)^2}{c_j}\, dx = n,$$

we have for $j = k$,

$$\begin{aligned}
&\int \sum_{\sigma(k)=k} \operatorname{sgn}(\sigma) K(x_1, x_{\sigma(1)}) \dots K(x_{k-1}, x_{\sigma(k-1)}) \dots K(x_k, x_k) dx_k \\
=&\, n \sum_{\sigma \in \Sigma_{k-1,k-1}} \operatorname{sgn}(\sigma) K(x_1, x_{\sigma(1)}) \dots K(x_{k-1}, x_{\sigma(k-1)}) \\
=&\, n \det_{i,j=1}^{k-1} K(x_i, x_j).
\end{aligned}$$

Summing over j from 1 to k gives

$$R_{k-1}(x_1, \dots, x_{k-1}) = n! C \cdot c_0 \dots c_{n-1} \det_{i,j=1}^{k-1} K(x_i, x_j)$$

as desired, completing the induction. In particular,

$$\begin{aligned}
1 &= \int_{\mathbb{R}^n} \rho(x_1, \dots, x_n)\, dx_1 \dots dx_n = \frac{1}{n} \int R_1(x_1)\, dx_1 \\
&= \frac{n! C \cdot c_0 \dots c_{n-1}}{n} \int K(x_1, x_1)\, dx_1 = n! C \cdot c_0 \dots c_{n-1},
\end{aligned}$$

giving the desired result. □

Proposition 4.5 will allow us to obtain a limit theorem for the k-point eigenvalue correlation functions of the GUE and Wishart Ensemble as $n \to \infty$. Before doing so, though, let us comment that the k-point eigenvalue correlation functions of X allow us to compute various other local statistics of the eigenvalue distribution of X. As an example, we may obtain a formula for the probability of finding exactly m eigenvalues in a closed interval:

Proposition 4.6. *Suppose the joint density of unordered eigenvalues of X is given by equation (10). Suppose $\{p_k\}_{k=0}^{\infty}$ are monic polynomials orthogonal with respect to w, with p_k of degree k and norm-squared c_k with respect to w. For $m \geq 0$, let $S_m([a,b])$ be the probability that X has exactly m eigenvalues in $[a,b]$. Then*

$$S_m([a,b]) = \frac{1}{m!} \left(-\frac{d}{d\gamma} \right)^m \det(I - \gamma K)_{[a,b]} \Big|_{\gamma=1}$$

where

$$K(x,y) = \sum_{j=0}^{n-1} \frac{\sqrt{w(x)w(y)}\, p_j(x) p_j(y)}{c_j}.$$

Proof. Let us denote $\Sigma'_{m,n}$ as the collection of (unordered) subsets of $\{1, \dots, n\}$ of size m. We may use (8) and (9) to obtain

$$\begin{aligned}
S_m([a,b]) &= \int_{x_1 \geq \dots \geq x_n} \sum_{\sigma \in \Sigma'_{m,n}} \prod_{i \in \sigma} \mathbb{1}_{[a,b]}(x_i) \prod_{i \notin \sigma} \left(1 - \mathbb{1}_{[a,b]}(x_i)\right) \hat{\rho}(x_1, \dots, x_n) dx_1 \dots dx_n \\
&= \int_{\mathbb{R}^n} \sum_{\sigma \in \Sigma'_{m,n}} \prod_{i \in \sigma} \mathbb{1}_{[a,b]}(x_i) \prod_{i \notin \sigma} \left(1 - \mathbb{1}_{[a,b]}(x_i)\right) \rho(x_1, \dots, x_n) dx_1 \dots dx_n \\
&= \binom{n}{m} \int_{\mathbb{R}^n} \prod_{k=1}^{m} \mathbb{1}_{[a,b]}(x_k) \prod_{k=m+1}^{n} \left(1 - \mathbb{1}_{[a,b]}(x_k)\right) \rho(x_1, \dots, x_n) dx_1 \dots dx_n.
\end{aligned}$$

We note that

$$-\frac{d}{d\gamma}\int_{\mathbb{R}^n}\prod_{k=1}^{m}\mathbb{1}_{[a,b]}(x_k)\prod_{k=m+1}^{n}\left(1-\gamma\mathbb{1}_{[a,b]}(x_k)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n$$

$$=\int_{\mathbb{R}^n}\sum_{k=m+1}^{n}\mathbb{1}_{[a,b]}(x_k)\prod_{j=1}^{m}\mathbb{1}_{[a,b]}(x_j)\prod_{\substack{j=m+1\\ j\neq k}}^{n}\left(1-\gamma\mathbb{1}_{[a,b]}(x_j)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n$$

$$=(n-m)\int_{\mathbb{R}^n}\prod_{k=1}^{m+1}\mathbb{1}_{[a,b]}(x_k)\prod_{k=m+2}^{n}\left(1-\gamma\mathbb{1}_{[a,b]}(x_k)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n.$$

Hence

$$S_m([a,b])=\frac{1}{m!}\left(-\frac{d}{d\gamma}\right)^m\left[\int_{\mathbb{R}^n}\prod_{k=1}^{n}\left(1-\gamma\mathbb{1}_{[a,b]}(x_k)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n\right]_{\gamma=1}.$$

We then have that

$$\int_{\mathbb{R}^n}\prod_{k=1}^{n}\left(1-\gamma\mathbb{1}_{[a,b]}(x_k)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n$$

$$=\int_{\mathbb{R}^n}\left(1+\sum_{k=1}^{n}(-\gamma)^k\sum_{\sigma\in\Sigma'_{k,n}}\prod_{i\in\sigma}\mathbb{1}_{[a,b]}(x_i)\right)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n$$

$$=1+\sum_{k=1}^{n}(-\gamma)^k\cdot\binom{n}{k}\int_{\mathbb{R}^n}\prod_{i=1}^{k}\mathbb{1}_{[a,b]}(x_i)\rho(x_1,\ldots,x_n)dx_1\ldots dx_n$$

$$=1+\sum_{k=1}^{n}\frac{(-\gamma)^k}{k!}\int_{[a,b]^k}R(x_1,\ldots,x_k)dx_1\ldots dx_k$$

$$=1+\sum_{k=1}^{n}\frac{(-\gamma)^k}{k!}\int_{[a,b]^k}\det_{i,j=1}^{k}K(x_i,x_j)dx_1\ldots dx_k$$

by Proposition 4.5. We note that

$$(K(x_i,x_j))_{i,j=1}^k=\sum_{j=0}^{n-1}\frac{1}{c_j}\begin{pmatrix}\sqrt{w(x_1)}p_j(x_1)\\ \vdots\\ \sqrt{w(x_k)}p_j(x_k)\end{pmatrix}\begin{pmatrix}\sqrt{w(x_1)}p_j(x_1) & \ldots & \sqrt{w(x_k)}p_j(x_k)\end{pmatrix},$$

which has rank at most n. Hence $\det_{i,j=1}^k K(x_i,x_j)=0$ for all $k>n$, so

$$S_m([a,b])=\frac{1}{m!}\left(-\frac{d}{d\gamma}\right)^m\det(I-\gamma K)_{[a,b]}\Big|_{\gamma=1}.$$

□

4.2 Limiting behavior of the correlation function in the bulk of the spectrum

Proposition 4.5 expresses the k-point eigenvalue correlation functions of fixed random matrices X in terms of polynomials orthogonal with respect to corresponding weight functions. By using certain Plancherel-Rotach asymptotics concerning the polynomials corresponding to the GUE and Wishart Ensemble, we may obtain asymptotics for their k-point eigenvalue correlation functions. Specifically, we will prove the following two results:

Theorem 4.7. *Let $\{W_n\}_{n\in\mathbb{N}}$ be the GUE, and let $R_k^{(n)}(x_1,\ldots,x_k)$ be the k-point eigenvalue correlation function of the scaled matrix $\frac{1}{\sqrt{n}}W_n$ for each n. Then for any $c\in(-2,2)$ and any distinct values $\xi_1,\ldots,\xi_k\in\mathbb{R}$,*

$$\lim_{n\to\infty}\frac{1}{(n\sigma(c))^k}R_k^{(n)}\left(c+\frac{\xi_1}{n\sigma(c)},\ldots,c+\frac{\xi_k}{n\sigma(c)}\right)=\det_{i,j=1}^{k}K(\xi_i,\xi_j),$$

where σ is the semicircle density from Theorem 2.14 and

$$K(\xi,\eta)=\frac{\sin\pi(\xi-\eta)}{\pi(\xi-\eta)}.$$

Theorem 4.8. *Let $\{M_n\}_{n\in\mathbb{N}}$ be the Wishart Ensemble with parameter $p=1$, and let $R_k^{(n)}(x_1,\ldots,x_k)$ be the k-point eigenvalue correlation function of the scaled matrix $\frac{1}{n}M_n$ for each n. Then for any $c\in(0,4)$ and any distinct values $\xi_1,\ldots,\xi_k\in\mathbb{R}$,*

$$\lim_{n\to\infty}\frac{1}{(n\mu_{MP}(c))^k}R_k^{(n)}\left(c+\frac{\xi_1}{n\mu_{MP}(c)},\ldots,c+\frac{\xi_k}{n\mu_{MP}(c)}\right)=\det_{i,j=1}^{k}K(\xi_i,\xi_j),$$

where μ_{MP} is the Marčenko-Pastur density from Theorem 2.17 and

$$K(\xi,\eta)=\frac{\sin\pi(\xi-\eta)}{\pi(\xi-\eta)}.$$

Intuitively, these results state that if we pick a point c in the interior of the support of the spectrum for the scaled GUE or Wishart Ensemble and then rescale the matrices so that the eigenvalue density is 1 at c, then the local k-point correlation functions around c have a pointwise limit. The method of proof for these results was introduced by Dyson, who established the $c=0$ case of Theorem 4.7 in [6].

To prove these limit results, we will need to compute the asymptotics of the kernel function

$$K(x,y)=\sum_{j=0}^{n-1}\frac{\sqrt{w(x)w(y)}p_j(x)p_j(y)}{c_j}$$

from Proposition 4.5. As a first step, let us express this function in closed form.

Lemma 4.9. *Suppose $\{p_k\}_{k=0}^{\infty}$ are monic polynomials orthogonal with respect to a weight function w, with p_k of degree k and norm-squared c_k with respect to w. Then for $x\neq y$,*

$$\sum_{j=0}^{n-1}\frac{p_j(x)p_j(y)}{c_j}=\frac{p_n(x)p_{n-1}(y)-p_{n-1}(y)p_n(x)}{c_{n-1}(x-y)}.$$

Proof. Let $q_j(x)=\frac{1}{\sqrt{c_j}}p_j(x)$ for all j. We note that for any j, $\{q_i\}_{i=0}^{j-1}$ forms a basis for the space of polynomials of degree at most $j-1$, so q_j is orthogonal, with respect to w, to any polynomial of degree at most $j-1$. For $j\geq 2$, $q_j(x)-\sqrt{\frac{c_{j-1}}{c_j}}xq_{j-1}(x)$ is a polynomial of degree $j-1$ (because p_j is monic), and hence it is a linear combination of $\{q_i\}_{i=0}^{j-1}$. We note that

$$\begin{aligned}&\int\left(q_j(x)-\sqrt{\frac{c_{j-1}}{c_j}}xq_{j-1}(x)\right)q_i(x)w(x)dx\\&=\int q_j(x)q_i(x)w(x)dx-\sqrt{\frac{c_{j-1}}{c_j}}\int xq_{j-1}(x)q_i(x)w(x)dx\\&=0\end{aligned}$$

for all $i < j-2$. Hence $q_j(x) - \sqrt{\frac{c_{j-1}}{c_j}} x q_{j-1}(x) = C_j q_{j-1}(x) + D_j q_{j-2}(x)$ for some constants C_j and D_j. Multiplying by $q_{j-2}(x)w(x)$ and integrating, this gives

$$D_j = -\sqrt{\frac{c_{j-1}}{c_j}} \int q_{j-1}(x)(xq_{j-2}(x))w(x)dx = -\sqrt{\frac{c_{j-1}^2}{c_j c_{j-2}}},$$

as $xq_{j-2}(x) = \sqrt{\frac{c_{j-1}}{c_{j-2}}} q_{j-1}(x) + \ldots$. So for all $j \geq 2$ and some constants C_j,

$$q_j(x) = \left(\sqrt{\frac{c_{j-1}}{c_j}} x + C_j\right) q_{j-1}(x) - \sqrt{\frac{c_{j-1}^2}{c_j c_{j-2}}} q_{j-2}(x).$$

Then

$$\begin{aligned} q_j(x)q_{j-1}(y) - q_{j-1}(x)q_j(y) &= \left[\left(\sqrt{\frac{c_{j-1}}{c_j}} x + C_j\right) q_{j-1}(x) - \sqrt{\frac{c_{j-1}^2}{c_j c_{j-2}}} q_{j-2}(x)\right] q_{j-1}(y) \\ &\quad - q_{j-1}(x) \left[\left(\sqrt{\frac{c_{j-1}}{c_j}} y + C_j\right) q_{j-1}(y) - \sqrt{\frac{c_{j-1}^2}{c_j c_{j-2}}} q_{j-2}(y)\right] \\ &= \sqrt{\frac{c_{j-1}}{c_j}} (x-y) q_{j-1}(x) q_{j-1}(y) \\ &\quad + \sqrt{\frac{c_{j-1}^2}{c_j c_{j-2}}} \left(q_{j-1}(x)q_{j-2}(y) - q_{j-2}(x)q_j(y)\right), \end{aligned}$$

so

$$\sqrt{\frac{c_j}{c_{j-1}}} \frac{q_j(x)q_{j-1}(y) - q_{j-1}(x)q_j(y)}{x-y} = q_{j-1}(x)q_{j-1}(y) + \sqrt{\frac{c_{j-1}}{c_{j-2}}} \frac{q_{j-1}(x)q_{j-2}(y) - q_{j-2}(x)q_j(y)}{x-y}.$$

Summing over j from 2 to n gives

$$\sqrt{\frac{c_n}{c_{n-1}}} \frac{q_n(x)q_{n-1}(y) - q_{n-1}(x)q_n(y)}{x-y} = \sum_{j=1}^{n-1} q_j(x)q_j(y) + \sqrt{\frac{c_1}{c_0}} \frac{q_1(x)q_0(y) - q_0(x)q_1(y)}{x-y}.$$

As $q_1(x) = \frac{1}{\sqrt{c_1}} x = \sqrt{\frac{c_0}{c_1}} x q_0(x)$, we have

$$\sqrt{\frac{c_1}{c_0}} \frac{q_1(x)q_0(y) - q_0(x)q_1(y)}{x-y} = q_0(x)q_0(y),$$

and substituting $q_j(x) = \frac{1}{\sqrt{c_j}} p_j(x)$ gives the desired result. □

4.2.1 Hermite polynomials and the GUE

The orthogonal polynomials corresponding to the eigenvalue density functions of the scaled GUE are based on the classical Hermite polynomials, which are orthogonal with respect to the weight $w(x) = e^{-x^2}$:

Definition 4.10. *The Hermite polynomial of degree n is $h_n(x) = e^{x^2}(-1)^n \frac{d^n}{dx^n} e^{-x^2}$.*

Lemma 4.11. *The degree n Hermite polynomial h_n has leading coefficient 2^n, and*

$$\int_{-\infty}^{\infty} h_n(x)h_m(x)e^{-x^2} dx = \sqrt{\pi} 2^n n! \delta_{mn}.$$

Proof. From the definition of h_n,

$$h_n(x) = -e^{x^2}(-1)^{n-1}\frac{d}{dx}\left[\frac{d^{n-1}}{dx^{n-1}}e^{-x^2}\right] = -e^{x^2}\frac{d}{dx}[h_{n-1}(x)e^{-x^2}] = 2xh_{n-1}(x) - h'_{n-1}(x).$$

As $h_0 = 1$, h_n has leading coefficient 2^n by induction. Supposing without loss of generality that $m \geq n$,

$$\int_{-\infty}^{\infty} h_n(x)h_m(x)e^{-x^2}dx = (-1)^m\int_{-\infty}^{\infty} h_n(x)\frac{d^m}{dx^m}\left(e^{-x^2}\right)dx = \int_{-\infty}^{\infty}\frac{d^m}{dx^m}\left(h_n(x)\right)e^{-x^2}dx,$$

where the first equality follows from the definition of h_m and the second from integration by parts m times. If $m > n$, then this is 0, while if $m = n$, then $\frac{d^n}{dx^n}h_n(x) = 2^n n!$ and so this gives $\sqrt{\pi}2^n n!$. □

To establish the limit law in Theorem 4.7, we will assume the following Plancherel-Rotach asymptotic of the Hermite polynomials h_n, whose proof may be found in [24]:

Proposition 4.12. *Let $\varphi \in [\varepsilon, \pi - \varepsilon]$ for some $\varepsilon > 0$, and let $x_n = (2n+1)^{1/2}\cos\varphi$. Then*

$$e^{-\frac{x_n^2}{2}}h_n(x_n) = 2^{\frac{n}{2}+\frac{1}{4}}(n!)^{\frac{1}{2}}(\pi n)^{-\frac{1}{4}}(\sin\varphi)^{-\frac{1}{2}}\left[\sin\left[\left(\frac{n}{2}+\frac{1}{4}\right)(\sin 2\varphi - 2\varphi) + \frac{3\pi}{4}\right] + O\left(\frac{1}{n}\right)\right],$$

where the error term is $O(1/n)$ uniformly over $\varphi \in [\varepsilon, \pi - \varepsilon]$.

These tools allow us to complete the proof of the asymptotic limit of the k-point eigenvalue correlation function for the scaled GUE.

Proof of Theorem 4.7. Rescaling the result from Theorem 3.5, the joint density of unordered eigenvalues of $\frac{1}{\sqrt{n}}W_n$, where W_n is the $n \times n$ matrix of the GUE, is given by

$$\rho(x_1, \ldots, x_n) = C_n\prod_{i<j}|x_i - x_j|^2\prod_{i=1}^{n}e^{-\frac{nx_i^2}{2}}$$

for some constant C_n. From Lemma 4.11, we have

$$\int_{-\infty}^{\infty} h_j\left(\sqrt{\frac{n}{2}}x\right)h_k\left(\sqrt{\frac{n}{2}}x\right)e^{-\frac{nx^2}{2}}dx = \sqrt{\frac{2\pi}{n}}2^j j!\delta_{jk},$$

and $h_j\left(\sqrt{\frac{n}{2}}x\right)$ has leading coefficient $(2n)^{j/2}$. Thus the monic polynomials orthogonal with respect to weight $e^{-\frac{nx^2}{2}}$ are given by $p_j(x) = (2n)^{-j/2}h_j\left(\sqrt{\frac{n}{2}}x\right)$, and p_j has norm-squared

$$\int_{-\infty}^{\infty} p_j(x)^2 e^{-nx^2/2}dx = n^{-j-\frac{1}{2}}\sqrt{2\pi}j!.$$

Hence, by Lemma 4.9, the kernel function of Proposition 4.5 is given for $x \neq y$ by

$$\begin{aligned}K_n(x,y) &= \frac{e^{-\frac{n}{4}(x^2+y^2)}n^{n-\frac{1}{2}}(p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y))}{\sqrt{2\pi}(n-1)!(x-y)}\\ &= \frac{e^{-\frac{n}{4}(x^2+y^2)}\left[h_n\left(\sqrt{\frac{n}{2}}x\right)h_{n-1}\left(\sqrt{\frac{n}{2}}y\right) - h_{n-1}\left(\sqrt{\frac{n}{2}}x\right)h_n\left(\sqrt{\frac{n}{2}}y\right)\right]}{2^n\sqrt{\pi}(n-1)!(x-y)}.\end{aligned}$$

Let us set $x = c + \frac{\xi}{n\sigma(c)}$ and $y = c + \frac{\eta}{n\sigma(c)}$, where $c \in (-2, 2)$. Then for all sufficiently large n, we may set $\theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2} \in (0, \pi)$ such that

$$\cos\theta_{n,1} = \sqrt{\frac{n}{2(2n+1)}}x, \quad \cos\theta_{n,2} = \sqrt{\frac{n}{2(2n-1)}}x,$$
$$\cos\varphi_{n,1} = \sqrt{\frac{n}{2(2n+1)}}y, \quad \cos\varphi_{n,2} = \sqrt{\frac{n}{2(2n-1)}}y,$$

so that $\theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2} \to \cos^{-1}\left(\frac{c}{2}\right)$ as $n \to \infty$. Then, using Proposition 4.12, this gives

$$K_n(x,y) = \frac{n\sigma(c)}{\pi(\xi-\eta)}\Bigg(\frac{\sin\left[\left(\frac{n}{2}+\frac{1}{4}\right)(\sin 2\theta_{n,1}-2\theta_{n,1})+\frac{3\pi}{4}\right]\sin\left[\left(\frac{n}{2}-\frac{1}{4}\right)(\sin 2\varphi_{n,2}-2\varphi_{n,2})+\frac{3\pi}{4}\right]}{(\sin\theta_{n,1}\sin\varphi_{n,2})^{\frac{1}{2}}}$$
$$-\frac{\sin\left[\left(\frac{n}{2}-\frac{1}{4}\right)(\sin 2\theta_{n,2}-2\theta_{n,2})+\frac{3\pi}{4}\right]\sin\left[\left(\frac{n}{2}+\frac{1}{4}\right)(\sin 2\varphi_{n,1}-2\varphi_{n,1})+\frac{3\pi}{4}\right]}{(\sin\theta_{n,2}\sin\varphi_{n,1})^{\frac{1}{2}}}+O\left(\tfrac{1}{n}\right)\Bigg).$$

We may compute the Taylor expansions of $\sin 2\theta = 2\sin\theta\cos\theta = 2\sqrt{1-\cos^2\theta}\cos\theta$ and $\theta = \cos^{-1}(\cos\theta)$ in terms of $\frac{1}{n}$ for $\theta = \theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2}$ to obtain the following asymptotic identities:

$$\begin{aligned}
\left(\tfrac{n}{2}+\tfrac{1}{4}\right)(\sin 2\theta_{n,1}-2\theta_{n,1}) &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\xi-\frac{1}{2}\cos^{-1}\left(\frac{c}{2}\right)+O\left(\tfrac{1}{n}\right)\\
\left(\tfrac{n}{2}-\tfrac{1}{4}\right)(\sin 2\theta_{n,2}-2\theta_{n,2}) &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\xi+\frac{1}{2}\cos^{-1}\left(\frac{c}{2}\right)+O\left(\tfrac{1}{n}\right)\\
\left(\tfrac{n}{2}+\tfrac{1}{4}\right)(\sin 2\varphi_{n,1}-2\varphi_{n,1}) &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\eta-\frac{1}{2}\cos^{-1}\left(\frac{c}{2}\right)+O\left(\tfrac{1}{n}\right)\\
\left(\tfrac{n}{2}-\tfrac{1}{4}\right)(\sin 2\varphi_{n,2}-2\varphi_{n,2}) &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\eta+\frac{1}{2}\cos^{-1}\left(\frac{c}{2}\right)+O\left(\tfrac{1}{n}\right)
\end{aligned}$$

Using the trigonometric identity

$$\sin(A-C)\sin(B+C)-\sin(A+C)\sin(B-C) = \sin(A-B)\sin(2C)$$

with

$$\begin{aligned}
A &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\xi+\frac{3\pi}{4},\\
B &= n\left(\frac{c}{4}\sqrt{4-c^2}-\cos^{-1}\left(\frac{c}{2}\right)\right)+\pi\eta+\frac{3\pi}{4},\\
C &= \frac{1}{2}\cos^{-1}\left(\frac{c}{2}\right),
\end{aligned}$$

this gives

$$K_n\left(c+\frac{\xi}{n\sigma(c)}, c+\frac{\eta}{n\sigma(c)}\right) = \frac{n\sigma(c)}{\pi(\xi-\eta)}\left(\frac{\sin\pi(\xi-\eta)\sin\left(\cos^{-1}\left(\frac{c}{2}\right)\right)}{\sin\left(\cos^{-1}\left(\frac{c}{2}\right)\right)}+o(1)\right),$$

and hence

$$\frac{1}{n\sigma(c)}K_n\left(c+\frac{\xi}{n\sigma(c)}, c+\frac{\eta}{n\sigma(c)}\right) \to \frac{\sin\pi(\xi-\eta)}{\pi(\xi-\eta)}.$$

Together with Proposition 4.5, this establishes the desired result. □

4.2.2 Laguerre polynomials and the Wishart Ensemble for $p = 1$

We may carry out the same argument for the Wishart Ensemble with parameter $p = 1$. The orthogonal polynomials corresponding to the eigenvalue density function of the scaled Wishart Ensemble for $p = 1$ are based on the classical Laguerre polynomials, which are orthogonal with respect to the weight $w(x) = e^{-x}\mathbb{1}_{x\geq 0}$:

Definition 4.13. *The Laguerre polynomial of degree n is $l_n(x) = \frac{1}{n!}e^x\frac{d^n}{dx^n}(e^{-x}x^n)$.*

Lemma 4.14. *The degree n Laguerre polynomial l_n has leading coefficient $\frac{(-1)^n}{n!}$, and*

$$\int_0^\infty l_n(x)l_m(x)e^{-x}dx = \delta_{mn}.$$

Proof. The highest order term of $l_n(x)$ is the x^n term with coefficient $\frac{1}{n!}e^x \frac{d^n}{dx^n}(e^{-x}) = \frac{(-1)^n}{n!}$. Supposing without loss of generality that $m \geq n$,

$$\int_0^\infty l_n(x)l_m(x)e^{-x}dx = \frac{1}{m!}\int_0^\infty l_n(x)\frac{d^m}{dx^m}\left(e^{-x}x^m\right)dx = \frac{(-1)^m}{m!}\int_0^\infty \frac{d^m}{dx^m}\left(l_n(x)\right)e^{-x}x^m dx,$$

where the first equality follows from the definition of l_m and the second from integration by parts m times (and we note that $\frac{d^k}{dx^k}(e^{-x}x^m) = 0$ at $x = 0$ and $x \to \infty$ for all $k < m$). If $m > n$, then this is 0, while if $m = n$, then $\frac{d^n}{dx^n}l_n(x) = (-1)^n$ and so this gives 1 as the integral of the density for the Gamma distribution. □

Analogous to the GUE case, we will assume the following Plancherel-Rotach asymptotic for the Laguerre polynomials l_n, whose proof may be found in [24]:

Proposition 4.15. *Let $\varphi_n \in [\varepsilon, \frac{\pi}{2} - \varepsilon n^{-\frac{1}{2}}]$ for some $\varepsilon > 0$, and let $x_n = (4n+2)\cos^2\varphi_n$. Then*

$$e^{-\frac{x_n}{2}}l_n(x_n) = (-1)^n(\pi\sin\varphi_n)^{-\frac{1}{2}}(x_n n)^{-\frac{1}{4}}\left[\sin\left[\left(n+\frac{1}{2}\right)(\sin 2\varphi_n - 2\varphi_n) + \frac{3\pi}{4}\right] + O\left(\frac{1}{(nx_n)^{\frac{1}{2}}}\right)\right],$$

where the error term $O\left(\frac{1}{(nx_n)^{\frac{1}{2}}}\right)$ holds uniformly for $\varphi_n \in [\varepsilon, \frac{\pi}{2} - \varepsilon n^{-\frac{1}{2}}]$.

Using this, we may derive the asymptotic limit of the k-point eigenvalue correlation function for the scaled Wishart Ensemble.

Proof of Theorem 4.8. Rescaling the result from Theorem 3.7, the joint density of unordered eigenvalues of $\frac{1}{n}M_n$, where M_n is the $n \times n$ matrix of the Wishart Ensemble with parameter $p = 1$, is given by

$$\rho(x_1, \ldots, x_n) = C_n \prod_{i<j}|x_i - x_j|^2 \prod_{i=1}^n e^{-nx_i}\mathbb{1}_{x_i \geq 0}$$

for some constant C_n. From Lemma 4.14, we have

$$\int_0^\infty l_j(nx)l_k(nx)e^{-nx}dx = \frac{1}{n}\delta_{jk},$$

and $l_j(nx)$ has leading coefficient $\frac{(-n)^j}{j!}$. Thus the monic polynomials orthogonal with respect to weight $e^{-nx}\mathbb{1}_{x\geq 0}$ are given by $p_j(x) = (-n)^{-j}j!l_j(nx)$, and p_j has norm-squared

$$\int_0^\infty p_j(x)^2 e^{-nx}dx = n^{-2j-1}(j!)^2.$$

Hence, by Lemma 4.9, the kernel function of Proposition 4.5 is given for $x \neq y$ by

$$\begin{aligned} K_n(x,y) &= \frac{e^{-\frac{n}{2}(x+y)}\mathbb{1}_{x\geq 0, y\geq 0}n^{2n-1}(p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y))}{(n-1)!^2(x-y)} \\ &= \frac{e^{-\frac{n}{2}(x+y)}\mathbb{1}_{x\geq 0, y\geq 0}n\left[l_{n-1}(nx)l_n(ny) - l_n(nx)l_{n-1}(ny)\right]}{x-y}. \end{aligned}$$

Let us set $x = c + \frac{\xi}{n\mu_{\mathrm{MP}}(c)}$ and $y = c + \frac{\eta}{n\mu_{\mathrm{MP}}(c)}$, where $c \in (0,4)$. Then for all sufficiently large n, we may set $\theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2} \in (0, \frac{\pi}{2})$ such that

$$\begin{aligned} \cos^2\theta_{n,1} &= \frac{n}{4n+2}x, \quad \cos^2\theta_{n,2} = \frac{n}{4n-2}x, \\ \cos^2\varphi_{n,1} &= \frac{n}{4n+2}y, \quad \cos^2\varphi_{n,2} = \frac{n}{4n-2}y, \end{aligned}$$

so that $\theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2} \to \cos^{-1}\left(\frac{\sqrt{c}}{2}\right)$ as $n \to \infty$. Then, using Proposition 4.15, this gives

$$K_n(x,y) = \frac{n\mu_{\mathrm{MP}}(c)}{\pi(\xi-\eta)(xy)^{\frac14}}\Bigg(\frac{\sin\left[(n+\frac12)(\sin 2\theta_{n,1}-2\theta_{n,1})+\frac{3\pi}{4}\right]\sin\left[(n-\frac12)(\sin 2\varphi_{n,2}-2\varphi_{n,2})+\frac{3\pi}{4}\right]}{(\sin\theta_{n,1}\sin\varphi_{n,2})^{\frac12}} - \frac{\sin\left[(n-\frac12)(\sin 2\theta_{n,2}-2\theta_{n,2})+\frac{3\pi}{4}\right]\sin\left[(n+\frac12)(\sin 2\varphi_{n,1}-2\varphi_{n,1})+\frac{3\pi}{4}\right]}{(\sin\theta_{n,2}\sin\varphi_{n,1})^{\frac12}} + O\left(\tfrac1n\right)\Bigg).$$

Using the Taylor expansions of $\sin 2\theta = 2\sin\theta\cos\theta = 2\sqrt{(1-\cos^2\theta)(\cos^2\theta)}$ and $\theta = \cos^{-1}\left(\sqrt{\cos^2\theta}\right)$ in terms of $\frac1n$ for $\theta = \theta_{n,1}, \theta_{n,2}, \varphi_{n,1}, \varphi_{n,2}$, we obtain:

$$\begin{aligned}
\left(n+\tfrac12\right)(\sin 2\theta_{n,1}-2\theta_{n,1}) &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\xi - \cos^{-1}\left(\frac{\sqrt{c}}{2}\right) + O\left(\tfrac1n\right)\\
\left(n-\tfrac12\right)(\sin 2\theta_{n,2}-2\theta_{n,2}) &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\xi + \cos^{-1}\left(\frac{\sqrt{c}}{2}\right) + O\left(\tfrac1n\right)\\
\left(n+\tfrac12\right)(\sin 2\varphi_{n,1}-2\varphi_{n,1}) &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\eta - \cos^{-1}\left(\frac{\sqrt{c}}{2}\right) + O\left(\tfrac1n\right)\\
\left(n-\tfrac12\right)(\sin 2\varphi_{n,2}-2\varphi_{n,2}) &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\eta + \cos^{-1}\left(\frac{\sqrt{c}}{2}\right) + O\left(\tfrac1n\right)
\end{aligned}$$

Using the trigonometric identity

$$\sin(A-C)\sin(B+C) - \sin(A+C)\sin(B-C) = \sin(A-B)\sin(2C)$$

with

$$\begin{aligned}
A &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\xi + \frac{3\pi}{4},\\
B &= \frac{n}{2}\left(\sqrt{c(4-c)} - 4\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right) + \pi\eta + \frac{3\pi}{4},\\
C &= \cos^{-1}\left(\frac{\sqrt{c}}{2}\right),
\end{aligned}$$

this gives

$$K_n\left(c + \frac{\xi}{n\mu_{\mathrm{MP}}(c)}, c + \frac{\eta}{n\mu_{\mathrm{MP}}(c)}\right) = \frac{n\mu_{\mathrm{MP}}(c)}{\pi(\xi-\eta)\sqrt{c}}\left(\frac{\sin\pi(\xi-\eta)\sin\left(2\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right)}{\sin\left(\cos^{-1}\left(\frac{\sqrt{c}}{2}\right)\right)} + o(1)\right),$$

and hence

$$\frac{1}{n\mu_{\mathrm{MP}}(c)}K_n\left(c + \frac{\xi}{n\mu_{\mathrm{MP}}(c)}, c + \frac{\eta}{n\mu_{\mathrm{MP}}(c)}\right) \to \frac{\sin\pi(\xi-\eta)}{\pi(\xi-\eta)}.$$

Together with Proposition 4.5, this establishes the desired result. □

Some concluding remarks are in order regarding generalizations of these results. It is not a coincidence that the limit forms of the kernel functions in Theorems 4.7 and 4.8 are identical. Asymptotics of the Plancherel-Rotach type, such as in Propositions 4.12 and 4.15, can be derived in greater generality for polynomials orthogonal to a general weight function $w(x)$, and these asymptotics can be used to prove the convergence of the kernel function to the sine kernel given in Theorem 4.7 for general classes of matrices corresponding to these orthogonal polynomials. We refer the reader to [5] for this approach. We assumed in Theorem 4.8 a parameter value of $p = 1$ so that the corresponding orthogonal polynomials are rescaled Laguerre polynomials. The same result is true for $p > 1$ ([9, 26]), but there is no classical result for the asymptotics of the corresponding polynomials (which are generalized Laguerre polynomials $l_n^{(\alpha)}$ for parameter α increasing with n).

As Theorems 4.7 and 4.8 describe the k-point eigenvalue correlation function around a point c in the interior of the support of the limiting spectrum, these results are referred to as convergence results for the “bulk” of the spectrum. The same method of proof yields a limit theorem for the “edge” of the spectrum, where c is on the boundary of the support of the limiting spectrum, using the corresponding Plancherel-Rotach asymptotics in [24]. The limiting kernel function for the edge of the spectrum is different from the sine kernel in Theorems 4.7 and 4.8, and is instead given by $K(x, y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$ where $\mathrm{Ai}(x)$ is the Airy function. We refer the reader to [1] and [5] for details.

Finally, the approach of orthogonal polynomials developed in this section depends on the ability to factor the joint eigenvalue densities from Theorems 3.5 and 3.7 as a product of weight functions of the eigenvalues, which relies on the property of invariance under unitary conjugation for the matrix ensembles, a consequence of the normal distribution of the matrix entries. That Theorems 4.7 and 4.8 hold for matrices with more general distributions of matrix entries was a long-standing conjecture, and results in this direction were proven recently in [8], [9], [10], [25], and [26].

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, Cambridge, to be published.

[2] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probability Theory and Related Fields*, 134:283–338, 2005.

[3] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, Hoboken, third edition, 2003.

[4] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statistica Sinica*, 9:611–677, 1999.

[5] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*. Courant Institute of Mathematical Sciences, New York, 1999.

[6] F. J. Dyson. Statistical theory of the energy levels of complex systems. I, II, III. *Journal of Mathematical Physics*, 3:140–175, 1962.

[7] F. J. Dyson. Correlations between eigenvalues of a random matrix. *Communications in Mathematical Physics*, 19:235–250, 1970.

[8] L. Erdös, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau. Bulk universality for Wigner matrices. arXiv:0905.4176.

[9] L. Erdös, B. Schlein, H.-T. Yau, and J. Yin. The local relaxation flow approach to universality of the local statistics for random matrices. arXiv:0911.3687.

[10] L. Erdös, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. arXiv:1001.3453.

[11] R. A. Fisher. The sampling distribution of some statistics obtained from non-linear equations. *Annals of Eugenics*, 9:238–249, 1939.

[12] J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6:440–449, 1965.

[13] M. A. Girshick. On the sampling theory of roots of determinantal equations. *Annals of Mathematical Statistics*, 10:203.

[14] N. R. Goodman. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 34:152–177, 1963.

[15] P. L. Hsu. On the distribution of roots of certain determinantal equations. *Annals of Eugenics*, 9:250–258, 1939.

[16] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3:391–397, 2000.

[17] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mathematics of the USSR: Sbornik*, 1:457–483, 1967.

[18] M. L. Mehta. *Random Matrices*. Academic Press, San Diego, third edition, 2004.

[19] M. L. Mehta and M. Gaudin. On the density of eigenvalues of a random matrix. *Nuclear Physics*, 18:420–427, 1960.

[20] T. Nagao and M. Wadati. Correlation functions of random matrix ensembles related to classical orthogonal polynomials. *Journal of The Physical Society of Japan*, 60:3298–3322, 1991.

[21] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley. Random matrix approach to cross-correlations in financial data. *Physical Review E*, 65:066126, 2002.

[22] S. N. Roy. p-statistics or some generalizations in the analysis of variance appropriate to multivariate problems. *Sankhya: Indian Journal of Statistics*, 4:381–396, 1939.

[23] J. A. Shohat and J. D. Tamarkin. *The Problem of Moments*. American Mathematical Society, New York, 1943.

[24] G. Szegö. *Orthogonal Polynomials*. American Mathematical Society, New York, 1939.

[25] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics. arXiv:0906.0510.

[26] T. Tao and V. Vu. Random covariance matrices: Universality of local statistics of eigenvalues. arXiv:0912.0966.

[27] A. M. Tulin and S. Verdu. *Random Matrix Theory and Wireless Communications*. Now Publishers, Hanover, 2004.

[28] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *The Annals of Mathematics*, 62:548–564, 1955.

[29] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *The Annals of Mathematics*, 67:325–327, 1958.

[30] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20A:32–52, 1928.