SMALLEST GAPS BETWEEN EIGENVALUES OF REAL GAUSSIAN MATRICES

PATRICK LOPATTO AND MATTHEW MEEKER

1. Introduction

1.1. **Background.** Random matrices have fascinated mathematicians and physicists for decades due to their connections to quantum chaos, number theory, statistics, and numerous other fields. A primary focus is the distribution of gaps between consecutive eigenvalues in random Hermitian matrices, which underpins many of these links. Substantial empirical evidence indicates that this distribution also arises in the spacings between energy levels of disordered quantum systems and the zeros of the Riemann zeta function, to give just two examples [6, 30].

While the distribution of a single eigenvalue gap is now mathematically understood for a wide variety of matrix models, less is known about the smallest and largest gaps. As motivation for their study, we note that the average-case performance of the Toda flow algorithm for diagonalizing a symmetric matrix can be analyzed in terms of the smallest eigenvalue gap of a Gaussian matrix [3, 12]. Additionally, the correspondence between eigenvalue gaps and spacings of zeta function zeros mentioned previously extends to the largest and smallest spacings in a given interval [3].

The rigorous study of extremal eigenvalue gaps was initiated by Vinson in his 2001 Ph.D. dissertation [38]. Using the method of moments, Vinson determined the asymptotic distribution of the smallest eigenvalue gap for the circular unitary ensemble (CUE), the Gaussian unitary ensemble (GUE), and unitarily-invariant β-ensembles. For the CUE and GUE, these results were extended by Ben Arous and Bourgade in [3], where they obtained the joint limiting distributions for the k smallest eigenvalue gaps. Instead of the method of moments, they drew on ideas developed by Soshnikov to study the smallest gaps of determinantal point processes [36]. Further, they also determined the asymptotic distribution of largest gaps for the CUE and GUE in the spectral bulk. The smallest gaps distribution for the Gaussian orthogonal ensemble (GOE) was later established by Feng, Tian, and Wei in [13]. We remark that determining the asymptotic distribution of the largest gaps for the GOE remains an open problem.

All of the matrix models mentioned in the previous paragraph are exactly solvable, in the sense that their eigenvalue correlation functions are given by explicit formulas. There has also been significant interest in studying extremal eigenvalue gaps for matrices outside of this class. In [9], Bourgade studied the distributions of smallest and largest gaps for a quite general set of random matrices, the generalized Wigner matrices, under a smoothness assumption on the entry distributions. He showed that these extremal gaps match those of the GOE/GUE (depending on whether the matrix is real symmetric or complex Hermitian). Landon, Lopatto, and Marcinek provided an alternative comparison argument for the largest gaps in [26], which does not require a smoothness hypothesis. As a consequence of these universality results and the works on the GOE/GUE mentioned previously, the largest gaps distribution is known for all Hermitian generalized Wigner matrices, and the smallest gaps distribution is known for generalized Wigner matrices of both symmetry classes with sufficiently smooth entries.

Further, lower bounds on the smallest gaps were established by Nguyen, Tao, and Vu in [29] for Wigner matrices with arbitrary mean (without a smoothness hypothesis), and for adjacency matrices of random graphs. Lopatto and Luh obtained similar lower bounds in [27] for sparse matrices, including adjacency matrices for sparse random graphs. Feng and Wei studied the smallest gaps for the circular β-ensemble, a generalization of the CUE, for all positive integer β in [14]. In [15], Figalli and Guionnet extended the results of Ben Arous and Bourgade from [3] to a several-matrix model.

So far, we have discussed only models with one-dimensional spectra, with eigenvalues lying on either the real line or the unit circle. In [31], Shi and Jiang studied the distribution of the smallest gaps for the complex Ginibre ensemble, a matrix of independent complex Gaussian random variables, whose spectrum is asymptotically supported in the unit disk in the complex plane. To the best of our knowledge, this is the only previous work identifying the asymptotic distribution of extremal gaps for an ensemble possessing a two-dimensional spectrum. (Shi and Jiang also consider Wishart matrices and unitarily-invariant β-ensembles, which have one-dimensional spectra.)

High-probability lower bounds for the smallest particle gap in a two-dimensional Coulomb gas were proved by Ameur in [1], and by Ameur and Romero in [2]. In [37], Thoma studied the smallest gap for Coulomb gases in two and three dimensions, providing upper and lower bounds and proving asymptotic tightness. We remark that Thoma's lower bound in two dimensions improves on those in [1, 2], and that he proves many other results on the separation of particles that hold in arbitrary dimension.

In this work, we consider the smallest gaps distribution for the real Ginibre ensemble, a matrix of independent real Gaussian random variables, whose spectrum is also asymptotically supported on the unit disk (but with a non-zero probability of real eigenvalues, unlike its complex counterpart). Due to the strong correlations between eigenvalues, we find that the smallest gaps between complex eigenvalues are of order $n^{-3/4}$. This contrasts with the smallest gap between n independent points in the disk, which is of order n^{-1}, and confirms the general principle that random matrix eigenvalues act as mutually repelling particles whose interactions suppress their fluctuations (relative to those of independent particles).

The methodology of Shi and Jiang in [31] was based on the fact that the eigenvalues of a complex Ginibre matrix form a determinantal point process, which makes the study of their smallest gaps amenable to the techniques developed in [3, 36]. However, the eigenvalues of real Ginibre matrices instead have a Pfaffian structure, which makes understanding their smallest gaps substantially more complicated. This work is therefore the first to determine the smallest gaps distribution for a two-dimensional ensemble without determinantal correlation functions.

1.2. **Main Result.** We begin by introducing some concepts necessary to precisely state our main result. When possible, our notation and definitions are chosen to match the previous works [21, 31], for consistency with the existing literature.

Definition 1.1. For all $n \in \mathbb{N}$, let $G_n = (g_{ij})_{1\le i,j\le n}$ denote a $n \times n$ random matrix whose entries are independent Gaussian random variables such that $\mathbb{E}[g_{ij}] = 0$ and $\mathbb{E}[g_{ij}^2] = 1$ for all i, j. The matrix G_n is called the *real Ginibre matrix* (GinOE) of dimension n. We also define $W_n = n^{-1/2} G_n$.

We refer the reader to [10], [11], and [16] for more information about the GinOE and the related unitary and symplectic Ginibre ensembles.

Next, we recall two well-known facts about the spectrum of W_n. First, in the limit as n tends to infinity, the empirical spectral distribution becomes uniformly distributed on the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ [7]. Second, the non-real eigenvalues of W_n come in conjugate pairs, since W_n has real entries. This means that if λ is an eigenvalue of W_n with non-zero imaginary part, then $\bar{\lambda}$ is also an eigenvalue of W_n, and the eigenvalues in the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ completely determine those in the lower half-plane.

Given these considerations, we restrict our attention to the eigenvalues of W_n lying in a given domain contained in the upper half disk

$$\mathbb{D}^+ = \{z \in \mathbb{C} : |z| < 1, \operatorname{Im} z > 0\},$$

and study the asymptotic distribution of the smallest gaps among these eigenvalues.[1] Further, to avoid boundary effects, we will consider only domains at a positive distance from the boundary of $\mathbb{D}^+$. This restriction leads to the following definition.

Definition 1.2. A domain Ω is called *admissible* if $\overline{\Omega} \subset \mathbb{D}^+$.

We will derive the distribution of the smallest gaps from a more general result about the convergence of a certain point process built from these gaps to a Poisson limit. To define this process, we begin with the definition of an order on points of $\mathbb{C}$.

Definition 1.3. For $z_1, z_2 \in \mathbb{C}$, we say that $z_1 \prec z_2$ if $\operatorname{Im}(z_1) < \operatorname{Im}(z_2)$, or if $\operatorname{Im}(z_1) = \operatorname{Im}(z_2)$ and $\operatorname{Re}(z_1) < \operatorname{Re}(z_2)$.

Let $\{\lambda_i\}_{i=1}^n$ denote the eigenvalues of W_n, indexed so that $\lambda_1 \prec \cdots \prec \lambda_n$ when all eigenvalues are distinct. On the measure zero set where the eigenvalues are not distinct, we label the eigenvalues in the same way, except we break ties between equal eigenvalues arbitrarily.

[1] We recall that a domain is is a non-empty connected open subset of $\mathbb{C}$.

Definition 1.4. Let Ω denote an admissible domain and set $\mathbb{R}^+ = [0,\infty)$. We define a point process $\chi_\Omega^{(n)}$ on $\mathbb{R}^+$ as follows. First, for all $i \in \mathbb{N}$ such that $1 \le i \le n$, we define

$$i^* = \arg\min_{j \neq i} \big\{ |\lambda_j - \lambda_i| : \lambda_i \prec \lambda_j, \lambda_j \in \Omega \big\} \tag{1.1}$$

if the set of indices j such that $j \neq i$ and $\lambda_j \in \Omega$ is nonempty. Otherwise, we set $i^* = 0$. We then set

$$\chi_\Omega^{(n)} = \sum_{i:\lambda_i \in \Omega} \delta_{n^{3/4}|\lambda_{i^*} - \lambda_i|} \mathbb{1}_{\{i^* \neq 0\}}. \tag{1.2}$$

Remark 1.5. As noted in [31, Remark 1.1], the point of introducing the order $\prec$ in the definition of i^* is to prevent the duplication of gaps in $\chi_\Omega^{(n)}$, to ensure good limiting behavior. For example, if i^* were defined as the index minimizing $|\lambda_i - \lambda_{i^*}|$ (without the order condition), then the smallest gap would appear twice in the set $\{|\lambda_i - \lambda_{i^*}| : 1 \le i \le n\}$, and $\chi_\Omega^{(n)}$ could not converge to a Poisson process with an absolutely continuous intensity measure.

We now state our main theorem on the Poisson convergence of $\chi_\Omega^{(n)}$. For every set $S \subset \mathbb{C}$, let $|S|$ denote the Lebesgue measure of C.

Theorem 1.6. Let Ω be an admissible domain. As $n \to \infty$, the processes $\chi_\Omega^{(n)}$ converge weakly to a Poisson point process χ_Ω on $\mathbb{R}^+$ with intensity

$$\mathbb{E}\big[\chi_\Omega(A)\big] = \frac{|\Omega|}{\pi} \int_A r^3 \, dr \tag{1.3}$$

for any bounded Borel set $A \subset \mathbb{R}^+$.

For the next corollary, we let

$$t_1^{(n)} = \min\{|\lambda_i - \lambda_j| : \lambda_i \in \Omega \wedge \lambda_j \in \Omega \wedge i < j\} \tag{1.4}$$

denote the smallest gap between eigenvalues in Ω, and more generally we let $t_\ell^{(n)}$ denote the ℓ-th smallest value in this set. We then define the rescaled gaps

$$\omega_\ell^{(n)} = n^{3/4} \left(\frac{\pi}{4|\Omega|} \right)^{1/4} t_\ell^{(n)}. \tag{1.5}$$

Corollary 1.7. For any real numbers $0 < x_1 < y_1 < \cdots < x_k < y_k$, we have

$$\lim_{n\to\infty} \mathbb{P}(x_\ell < \omega_\ell^{(n)} < y_\ell \text{ for all } 1 \le \ell \le k) = \left(e^{-x_k^4} - e^{-y_k^4} \right) \prod_{\ell=1}^{k-1} (y_\ell^4 - x_\ell^4).$$

In particular, the previous result implies that for any fixed $k \in \mathbb{N}$, the rescaled gap $\omega_k^{(n)}$ converges in distribution to a random variable with density $p(x) \propto x^{4k-1} e^{-x^4} \, dx$.

Remark 1.8. The limiting intensity measure in (1.3) is identical to the one for the smallest gaps process of the complex Ginibre ensemble (GinUE); see [31, Theorem 1.1].[2]

1.3. **Proof Ideas.** As mentioned in Section 1.1, the previous analysis of smallest gaps for the GinUE in [31] is based on the determinantal structure of the eigenvalue correlation functions, following ideas of [36] and [3]. Because the GinOE has Pfaffian correlation functions, this methodology does not immediately apply. However, it was observed in [21] and [24] that in the bulk of the spectrum, these Pfaffian correlation functions can be approximated by determinantal ones up to exponentially small additive errors. Using this idea, we are able to place ourselves back in the determinantal framework of [3, 31, 36].

After we make the reduction to a determinantal process, we come to the primary obstacle in adapting the strategy used for the smallest gaps of the GinUE in [31]. Their proof relies heavily on the positive-definiteness of the determinantal kernel for the GinUE, while the kernel in our determinantal approximation is not known to be positive-definite. To overcome this difficulty, we prove a technical lemma showing that the approximating kernel is positive-definite (or exponentially small) everywhere in the spectral bulk, expect possibly for a set of asymptotically vanishing measure (see Lemma 3.2). Using the determinantal approximation in tandem with this lemma, we are able to complete the proof of our main result by carefully tracking the contribution of the exceptional sets on which the kernel is not positive definite.

It is natural to wonder whether Theorem 1.6 holds for more general domains. Indeed, the analysis of the smallest gaps for the GinUE in [31] holds for any region of the complex plane. The methods here should extend in a relatively straightforward way to any region Ω such that $\overline{\Omega} \subset \mathbb{H}$; that is, we require a positive distance from the real axis but permit the region to extend beyond the unit circle. Specifically, one can augment the asymptotic analysis of the Pfaffian kernel in Lemma 2.4, which holds in the interior of the unit disk, with the asymptotics from [25, Remark 3.4] and [5, Theorem B.1] to access to the entire interior of the upper half-plane. We omit this extension for brevity, since it requires lengthy computations. Generalizing our result to regions that intersect the real line would require a precise accounting of the contributions from the real eigenvalues, and seems more difficult.

1.4. **Outline.** In Section 2, we begin by recalling previous results on the correlation functions for the GinOE, the Pfaffian, and the determinantal approximation, and state some straightforward consequences of these results. We also recall some facts about convergence to Poisson distributions and processes. In Section 3, we state a three key lemmas and show how they imply our main theorem and its corollary. In Section 4, we prove each of these key lemmas Section 4. Finally, in Section 5, we prove several auxiliary technical lemmas that we require at various points throughout this work.

1.5. **Acknowledgments.** P. L. is supported by NSF postdoctoral fellowship DMS-2202891.

[2]Consequently, the distributions of the rescaled smallest gaps also match. We note for the convenience of the reader that there is a misprint in the definition of the rescaled gaps $\tau_\ell^{(n)}$ in [31, Corollary 1.1].

2. Preliminary Results

2.1. **Correlation Functions for the GinOE.** We begin by recalling how to compute (symmetrized) expectations of functions of the eigenvalues $(w_i)_{i=1}^n$ of G_n. Let $\mathbb{C}^* = \mathbb{C}\backslash\mathbb{R}$ denote the complex plane with the real line removed, and let $\mathcal{I}_k \subset \{1, \dots, n\}^n$ be the set of pairwise distinct k-tuples of indices. By [8, (5.1)], for all $k, n \in \mathbb{N}$, there exists a function $\rho_k^{(n)} : (\mathbb{C}^*)^k \to \mathbb{R}^+$ such that

$$\int_{(\mathbb{C}^*)^k} f(z_1, \dots, z_k)\rho_k^{(n)}(z_1, \dots, z_k)\, dz_1 \dots dz_k = \mathbb{E}\left[\sum_{(i_1,\dots,i_k)\in\mathcal{I}_k} f(w_{i_1}, \dots, w_{i_k})\right] \tag{2.1}$$

for every compactly supported, bounded, and Borel-measurable function $f : (\mathbb{C}^*)^k \to \mathbb{R}$. We will use the shorthand $\rho_k = \rho_k^{(n)}$, suppressing the n-dependence in the notation. For more on correlation functions, including their definition for general random point fields, we refer the reader to [35].

The next lemma identifies the correlation function ρ_k as the Pfaffian of certain $2k \times 2k$ matrix. For the reader's convenience, we begin by recalling the definition of a Pfaffian.

Definition 2.1. The Pfaffian of a $2n \times 2n$ skew-symmetric matrix $M = (M_{ij})_{i,j=1}^{2n}$ is defined by

$$\mathrm{Pf}(M) = \frac{1}{2^n n!}\sum_{\sigma\in S_{2n}} \mathrm{sgn}(\sigma)\prod_{i=1}^{n} M_{\sigma(2i-1),\sigma(2i)},$$

where S_{2n} is the symmetric group of degree $2n$.

The statement below quotes [21, Theorem 1.1] exactly, which collected certain results from [28, Appendix B.3]. However, we emphasize that these correlation functions were originally identified explicitly in [18], and the Pfaffian form below was first derived in the works [8, 17, 32, 33].

Lemma 2.2. The k-point complex–complex correlation functions of the n-dimensional real Ginibre ensemble G_n are given by

$$\rho_k(z_1, \dots, z_k) = \mathrm{Pf}(K(z_i, z_j))_{1\le i,j\le k}, \tag{2.2}$$

where $(K(z_i, z_j))_{1\le i,j\le k}$ is a $2k \times 2k$ matrix composed of the 2×2 blocks

$$K(z_i, z_j) = \begin{pmatrix} D_n(z_i, z_j) & S_n(z_i, z_j) \\ -S_n(z_j, z_i) & I_n(z_i, z_j) \end{pmatrix},$$

and D_n, I_n, and S_n are defined by

$$S_n(z,w) = \frac{\mathrm{i}e^{-(1/2)(z-\bar{w})^2}}{\sqrt{2\pi}}(\bar{w}-z)G(z,w)s_n(z\bar{w}),$$
$$D_n(z,w) = \frac{e^{-(1/2)(z-w)^2}}{\sqrt{2\pi}}(w-z)G(z,w)s_n(zw),$$
$$I_n(z,w) = \frac{e^{-(1/2)(\bar{z}-\bar{w})^2}}{\sqrt{2\pi}}(\bar{z}-\bar{w})G(z,w)s_n(\bar{z}\bar{w}),$$

where $z, w \in \mathbb{C}^*$ and

$$G(z,w) = \sqrt{\operatorname{erfc}(\sqrt{2}\operatorname{Im}(z))\operatorname{erfc}(\sqrt{2}\operatorname{Im}(w))}, \qquad \operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty \exp(-t^2)\,dt,$$
$$s_n(z) = e^{-z}\sum_{j=0}^{n-1}\frac{z^j}{j!}.$$

Remark 2.3. By a change of variables in the definition (2.1), it follows that the k-th correlation function for the complex eigenvalues of W_n is $n^k\rho_k(\sqrt{n}z_1,\dots\sqrt{n}z_k)$.

The following lemma is useful for controlling the correlation functions of W_n. Its statement is taken from [21, Lemma 2.3].

Lemma 2.4. Let Ω be an admissible domain, and let $d_\Omega = \inf\{|z-w| : z \in \Omega, w \in \partial\mathbb{D}^+\}$ denote the distance between Ω and the boundary of $\mathbb{D}^+$. There exist constants $C(d_\Omega), c(d_\Omega) > 0$ such that

$$\sup_{z,w\in\Omega}\left|D_n(\sqrt{n}z,\sqrt{n}w)\right| \le Ce^{-cn}, \qquad \sup_{z,w\in\Omega}\left|I_n(\sqrt{n}z,\sqrt{n}w)\right| \le Ce^{-cn},$$
$$\sup_{z,w\in\Omega}\left|S_n(\sqrt{n}z,\sqrt{n}w)\right| \le C.$$

The statement of the next lemma is from [21, Lemma 2.4]. It was proved in [20, Appendix B] (and appeared earlier in [19]).

Lemma 2.5. Let $M = (M_{ij})_{i,j=1}^{2n}$ be a skew-symmetric $2n\times 2n$ matrix such that $M_{ij} = 0$ for every pair of indices (i,j) such that $i \equiv j \bmod 2$. Let $\widetilde{M} = (\widetilde{M})_{i,j=1}^n$ be the $n\times n$ matrix defined by $\widetilde{M}_{ij} = M_{2i-1,2j}$. Then $\operatorname{Pf}(M) = \det(\widetilde{M})$.

For all $k \in \mathbb{N}$, define the $k\times k$ matrix $Q^{(k)}(z_1,\dots,z_k) = (S_n(\sqrt{n}z_i,\sqrt{n}z_j))_{1\le i,j\le k}$. The following lemma provides a useful approximation of the correlation functions of W_n by a determinant (see Remark 2.3). It follows from combining Lemma 2.2, Lemma 2.4, and Lemma 2.5; we provide a detailed proof in Section 5.

Lemma 2.6. Fix $k \in \mathbb{N}$. For every admissible domain Ω, there exists a constant $c(k, d_\Omega) > 0$ such that for all $n \in \mathbb{N}$,

$$\sup_{z_1,\dots,z_k\in\Omega}\left|n^k\rho_k(\sqrt{n}z_1,\dots,\sqrt{n}z_k) - n^k\det Q^{(k)}(z_1,\dots,z_k)\right| \le c^{-1}e^{-cn}.$$

The next lemma is a consequence of Lemma 2.6 and provides a useful estimate on $S_n(\sqrt{n}z, \sqrt{n}w)$ when z and w are close together. Its proof also appears in Section 5. We define

$$U(z,w) = \frac{\mathrm{i}e^{(-n/2)(z-\bar{w})^2 - n(\operatorname{Im}(z)^2 + \operatorname{Im}(w)^2)}}{2\pi\sqrt{\operatorname{Im}(z)\operatorname{Im}(w)}}(\bar{w} - z).$$

Lemma 2.7. Fix an admissible domain Ω and $r \in \mathbb{R}^+$. Then there exists a constant $C(r, d_\Omega) > 0$ such that

$$\sup_{z,w\in\Omega:|z-w|<rn^{-3/4}} \left|S_n(\sqrt{n}z, \sqrt{n}w) - U(z,w)\right| \le \frac{C}{n}. \tag{2.3}$$

2.2. **Poisson Convergence.** We require two basic convergence results, one for Poisson random variables and one for Poisson point processes, which are proved in Section 5.

Proposition 2.8. Let $\{X_1\}_{n=1}^{\infty}$ be a sequence of random variables taking values in the non-negative integers, and let X be a Poisson random variable with rate $\lambda > 0$. Suppose that for all $k > 0$,

$$\lim_{n\to\infty} \mathbb{E}\big[X(X-1)\cdots(X-k)\big] = \lambda^k. \tag{2.4}$$

Then the sequence X_n converges to X in distribution.

Proposition 2.9. Let $\{\chi^{(n)}\}_{n=1}^{\infty}$ be a sequence of point processes on $\mathbb{R}$, and let χ be a Poisson point process on $\mathbb{R}$ with a intensity measure μ, which we suppose has no atoms. If $\chi^{(n)}(J)$ converges in distribution to $\chi(J)$ for all bounded Borel sets $J \subset \mathbb{R}$, then the sequence of point processes $\chi^{(n)}$ converges in distribution to χ.

3. Proof of Main Result

We now state a series a lemmas and show how they imply the main result. Their proofs follow in the coming sections.

3.1. **Modified Point Process.** We begin by introducing a auxiliary point process, the *s-modified* point process corresponding to the eigenvalues of W_n. This technique was originally introduced in the works [34, 36].

Given a bounded Borel set $A \subset \mathbb{R}^+$, we define the corresponding set

$$B = \{z \in \mathbb{C} : |z| \in A, 0 \prec z\}, \tag{3.1}$$

omitting the dependence on A in the notation (since the choice of A will always be clear from context). For all $n \in \mathbb{N}$, we set

$$A_n = \{n^{-3/4}a : a \in A\}, \qquad B_n = \{n^{-3/4}b : b \in B\}, \tag{3.2}$$

and define

$$\xi^{(n)} = \sum_{i:\lambda_i\in\Omega} \delta_{\lambda_i}, \quad \widetilde{\xi}^{(n)} = \sum_{i:\lambda_i\in\Omega} \delta_{\lambda_i} \mathbb{1}_{\{\xi^{(n)}(\lambda_i+B_n)=1\}}, \tag{3.3}$$

where the set $\lambda_i + B_n$ has the usual definition as $\lambda_i + B_n = \{\lambda_i + z : z \in B_n\}$.

The following lemma implies that it suffices to study $\widetilde{\xi}^{(n)}$ in order to prove our main result. It is proved in Section 4.

Lemma 3.1. Let Ω be an admissible domain. For all bounded Borel sets $A \subset \mathbb{R}^+$, we have

$$\lim_{n\to\infty} \chi^{(n)}(A) - \widetilde{\xi}^{(n)}(\Omega) = 0, \tag{3.4}$$

where the convergence is in distribution.

We let τ_k denote the k-th correlation function for $\xi^{(n)}$, suppressing the dependence on n in the notation. From the definition of a correlation function (see e.g. [35, Definition 2]) and Remark 2.3, it is straightforward to see that τ_k is given by

$$\tau_k(z_1, \dots, z_k) = n^k \rho_k(\sqrt{n} z_1, \dots, \sqrt{n} z_k) \mathbb{1}_\Omega(z_1) \dots \mathbb{1}_\Omega(z_k). \tag{3.5}$$

Further, we let $\widetilde{\tau}_k$ denote the k-th correlation function for the modified process $\widetilde{\xi}_n$. The function $\widetilde{\tau}_k$ can be written explicitly in terms of the functions $\{\tau_i\}_{i=1}^n$ using the inclusion–exclusion principle; see [35, (4.5)] for details.

For every $k \in \mathbb{N}$, we define the set

$$\Psi_k = \big\{(z_1, \dots, z_k) \in \Omega^k : (z_i + B_n) \cap (z_j + B_n) = \varnothing \text{ for all } i \neq j, 1 \leq i, j \leq k\big\}. \tag{3.6}$$

The following lemma is proved in Section 4.

Lemma 3.2. Fix an admissible domain Ω, a bounded Borel set $A \subset \mathbb{R}^+$, and $k \in \mathbb{N}$.

(1) There exists a set $\mathcal{Z} \subset \Psi_k$ such that $\mu(\mathcal{Z}) = 0$ and the following holds. For all $(z_1, \dots, z_k) \in \Omega^k \setminus \mathcal{Z}$ with pairwise distinct entries,

$$\lim_{n\to\infty} \widetilde{\tau}_k(z_1, \dots, z_k) = \left(\frac{1}{\pi^2} \int_B |z|^2 \, dz\right)^k. \tag{3.7}$$

(2) There exist constants $C, c > 0$, depending only on A, k, and Ω, such that the following statements hold for all $n \in \mathbb{N}$. There exists a set $\mathcal{W}_n \subset \Omega^k$ such that

$$\mathbb{P}(\mathcal{W}_n) \leq C e^{-cn}, \tag{3.8}$$

and for all $(z_1, \dots, z_k) \in \Psi_k \setminus \mathcal{W}_n$, we have

$$\widetilde{\tau}_k(z_1, \dots, z_k) \leq C, \tag{3.9}$$

and for all $(z_1, \dots, z_k) \in \mathcal{W}_n$,

$$\widetilde{\tau}_k(z_1, \dots, z_k) \leq C n^{8k}. \tag{3.10}$$

(3) Set $\overline{\Psi}_k = \Omega^k \setminus \Psi_k$. Then

$$\lim_{n\to\infty} \int_{\overline{\Psi}_k} \widetilde{\tau}_k(z_1, \dots, z_k) \, dz_1 \dots dz_k = 0. \tag{3.11}$$

Proof of Theorem 1.6. By the definition of a correlation function [35, Definition 2], we have

$$\int_{\Omega^k} \widetilde{\tau}_k(z_1, \dots, z_k) \, dz_1 \dots dz_k = \mathbb{E}\left[\frac{\widetilde{\xi}^{(n)}(\Omega)!}{\big(\widetilde{\xi}^{(n)}(\Omega) - k\big)!}\right]. \tag{3.12}$$

Abbreviating $\mathcal{W} = \mathcal{W}_n$, we write

$$\int_{\Omega^k} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k = \int_{\Omega^k} \mathbb{1}_{\Psi_k \setminus \mathcal{W}} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k \tag{3.13}$$

$$+ \int_{\Omega^k} \mathbb{1}_{\mathcal{W}} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k \tag{3.14}$$

$$+ \int_{\Omega^k} \mathbb{1}_{\overline{\Psi}_k} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k. \tag{3.15}$$

Observe that by the Borel–Cantelli lemma and (3.8), we have

$$\lim_{n\to\infty} \mathbb{1}_{\Psi_k \setminus \mathcal{W}} = \mathbb{1}_{\Omega}. \tag{3.16}$$

The integrand in the integral on the right-hand side of (3.13) is uniformly bounded, by Lemma 3.2(1), so by the dominated convergence theorem, Lemma 3.2(2), and (3.16), we conclude that

$$\lim_{n\to\infty} \int_{\Omega^k} \mathbb{1}_{\Psi_k} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k = \left(\frac{1}{\pi^2} \int_B |z|^2\, dz \right)^k \left(\int_\Omega dz \right)^k. \tag{3.17}$$

Using (3.8) and (3.10), we see that

$$\lim_{n\to\infty} \int_{\Omega^k} \mathbb{1}_{\mathcal{W}} \widetilde{\tau}_k(z_1, \dots, z_k)\, dz_1 \dots z_k = 0. \tag{3.18}$$

Further, by (3.15), the limit as n tends to infinity of (3.15) also vanishes. Then combining (3.12), (3.13), (3.18), and (3.17), we find

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{\widetilde{\xi}^{(n)}(\Omega)!}{\left(\widetilde{\xi}^{(n)}(\Omega) - k\right)!} \right] = \left(\int_B |z|^2\, dz \right)^k \left(\frac{1}{\pi^2} \int_\Omega dz \right)^k. \tag{3.19}$$

It follows from Proposition 2.8 that $\widetilde{\xi}^{(n)}(\Omega)$ converges in distribution to a Poisson random variable with intensity equal to

$$\left(\int_B |z|^2\, dz \right) \left(\int_\Omega \frac{dz}{\pi^2} \right) = \frac{|\Omega|}{\pi} \int_A r^3\, dr. \tag{3.20}$$

Then Lemma 3.1 implies that $\chi_\Omega^{(n)}(A)$ converges in distribution to a Poisson random variable with rate given by (3.20). Since this convergence holds for all bounded Borel sets $A \subset \mathbb{R}^+$, Proposition 2.9 implies that $\chi_\Omega^{(n)}$ converges to χ_Ω, as desired. □

For the proof of Corollary 1.7, we additionally require the next lemma, which says that it is unlikely that three eigenvalues all bunch together on the scale of the smallest eigenvalue gap. It is also proved in Section 4. We define the random measure

$$\Xi^{(3)} = \sum_{\lambda_{i_1}, \lambda_{i_2}, \lambda_{i_3} \text{ pairwise distinct}} \delta_{(\lambda_{i_1}, \lambda_{i_2}, \lambda_{i_3})}, \tag{3.21}$$

on $\mathbb{C}^3$, and for all admissible domains Ω and all $M > 0$, we define

$$\mathcal{B}_M = \{(z, x_1, x_2) : z \in \Omega, |x_1 - z| \le M n^{-3/4}, |x_2 - z| \le M n^{-3/4}\}. \tag{3.22}$$

Lemma 3.3. Fix $M > 0$ and an admissible domain Ω. Then

$$\lim_{n\to\infty} \mathbb{E}\big[\Xi^{(3)}(\mathcal{B}_M)\big] = 0. \tag{3.23}$$

Proof of Corollary 1.7. Let $\widetilde{t}_1 \le \cdots \le \widetilde{t}_k$ denote the k smallest values in the set

$$\big\{|\lambda_i - \lambda_{i^*}| : \lambda_i \in \Omega \wedge i^* \neq 0\big\}, \tag{3.24}$$

and define the rescaled values

$$\widetilde{\omega}_\ell = n^{3/4}\left(\frac{\pi}{4|\Omega|}\right)^{1/4} \widetilde{t}_\ell. \tag{3.25}$$

Observe that we have suppressed the n-dependence in the notation for the $\widetilde{t}_k$ and $\widetilde{\omega}_k$. We also use the shorthand $t_k = t_k^{(n)}$. We will first show that the $\widetilde{\omega}_\ell$ have the desired joint limiting distribution, then transfer this result to the ω_ℓ.

We begin by observing that the event

$$\{x_\ell < \widetilde{\omega}_\ell < y_\ell \text{ for all } 1 \le \ell \le k\} \tag{3.26}$$

can be written as the intersection of the events

$$\left\{\chi_\Omega^{(n)}\left(\left(\frac{4|\Omega|}{\pi}\right)^{1/4}(x_\ell, y_\ell)\right) \ge 1\right\}, \tag{3.27}$$

$$\left\{\chi_\Omega\left(\left(\frac{4|\Omega|}{\pi}\right)^{1/4}(x_k, y_k)\right) = 1 \text{ for all } 1 \le \ell \le k-1\right\}, \tag{3.28}$$

and

$$\left\{\chi_\Omega\left(\left(\frac{4|\Omega|}{\pi}\right)^{1/4}(y_{\ell-1}, x_\ell)\right) = 0 \text{ for all } 1 \le \ell \le k\right\}, \tag{3.29}$$

with the convention that $y_0 = 0$. Using this representation, and the fact that $\chi_\Omega(S_1)$ and $\chi_\Omega(S_2)$ are independent when the sets S_1 and S_2 are disjoint (since χ_Ω is a Poisson process), we have

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{P}(x_\ell < \widetilde{\omega}_\ell < y_\ell \text{ for all } 1 \le \ell \le k) = \\ &\Big(1 - e^{-(y_k^4 - x_k^4)}\Big)\prod_{\ell=1}^{k-1}(y_\ell^4 - x_\ell^4)e^{-(y_\ell^4 - x_\ell^4)}\prod_{\ell=1}^{k} e^{(y_\ell^4 - x_\ell^4)} = \Big(e^{-x_k^4} - e^{-y_k^4}\Big)\prod_{\ell=1}^{k-1}(y_\ell^4 - x_\ell^4).\end{aligned} \tag{3.30}$$

Next, we prove that

$$\lim_{n\to\infty} \mathbb{P}(\exists\, \ell \le k \text{ such that } \widetilde{t}_\ell \neq t_\ell) = 0. \tag{3.31}$$

Let $p_\ell < q_\ell$ denote the indices such that $|\lambda_{p_\ell} - \lambda_{q_\ell}| = t_\ell$. To show (3.31), it suffices to show that the probability that the $2k$ eigenvalues $(\lambda_{p_\ell}, \lambda_{q_\ell})_{\ell=1}^k$ are distinct tends to 1 as $n \to \infty$.

Using the notation defined before Lemma 3.3, we find that for every $M > 0$,

$$\mathbb{P}(\exists\, \ell \le k \text{ such that } \widetilde{t}_\ell \neq t_\ell) \le \mathbb{P}\big(\Xi^{(3)}(\mathcal{B}_M) \neq 0\big) + \mathbb{P}(\widetilde{t}_k > Mn^{-3/4}/2) \tag{3.32}$$

$$\le \mathbb{E}\big(\Xi^{(3)}(\mathcal{B}_M)\big) + \mathbb{P}(\widetilde{t}_k > Mn^{-3/4}/2). \tag{3.33}$$

By Lemma 3.3,

$$\lim_{n\to\infty} \mathbb{E}\big[\Xi^{(3)}(\mathcal{B}_M)\big] = 0. \tag{3.34}$$

Therefore

$$\limsup_{n\to\infty} \mathbb{P}(\exists\, \ell \le k \text{ such that } \widetilde{t}_\ell \neq t_\ell) \le \limsup_{n\to\infty} \mathbb{P}(\widetilde{t}_k > Mn^{-3/4}/2), \tag{3.35}$$

and we conclude the left-hand side of (3.35) is zero by taking M to infinity and using (3.30). □

4. Proofs of Main Lemmas

We begin with the proof of Lemma 3.3, since it is used in the proof of Lemma 3.1. We then prove Lemma 3.1 and Lemma 3.2. We use the convention that the letters C and c denote positive constants that may change line to line.

4.1. **Proof of Lemma 3.3.**

Proof of Lemma 3.3. For all $z \in \mathbb{C}$ and $r \in \mathbb{R}^+$, let

$$D(z,r) = \{w \in \mathbb{C} : |z - w| < r\}.$$

Set $M_n = Mn^{-3/4}$. We first note that, by the identification of the correlation functions for W_n in Remark 2.3, and Lemma 2.6, we have

$$\mathbb{E}\big[\Xi^{(3)}(\mathcal{B}_M)\big] = n^3 \int_\Omega \int_{D(\lambda,M_n)^2} \rho_3(\sqrt{n}\lambda, \sqrt{n}z_1, \sqrt{n}z_1)\, dz_1\, dz_2\, d\lambda \tag{4.1}$$

$$= n^3 \int_\Omega \int_{D(\lambda,M_n)^2} \det Q(\lambda, z_1, z_2)\, dz_1\, dz_2\, d\lambda + O(e^{-cn}), \tag{4.2}$$

where we abbreviate $Q(w_1,w_2,w_3) = Q^{(3)}(w_1,w_2,w_3)$. Here, and for the rest of this proof, the implicit constants in the asymptotic O notation depend only on A and Ω. We note that

$$\det Q(\lambda, z_1, z_2) = Q_{11}(Q_{22}Q_{33} - Q_{23}Q_{32}) - Q_{12}(Q_{21}Q_{33} - Q_{23}Q_{31}) \tag{4.3}$$

$$+ Q_{13}(Q_{21}Q_{32} - Q_{31}Q_{22}). \tag{4.4}$$

We begin by bounding the first term in (4.3). By Lemma 2.4, we have $Q_{11} = O(1)$, so we focus on the difference $Q_{22}Q_{33} - Q_{23}Q_{32}$.

Using Lemma 2.7 with $r = \tau$, we have

$$Q_{22} = \frac{\mathrm{i} e^{(-n/2)(z_1 - \bar{z}_1)^2 - n(\mathrm{Im}(z_1)^2 + \mathrm{Im}(z_1)^2)}}{2\pi\sqrt{\mathrm{Im}(z_1)\,\mathrm{Im}(z_1)}}(\bar{z}_1 - z_1) + O(n^{-1}) = \frac{1}{\pi} + O(n^{-1}), \tag{4.5}$$

and similarly

$$Q_{33} = \frac{1}{\pi} + O(n^{-1}). \tag{4.6}$$

Further,

$$Q_{23}Q_{32} = \frac{e^{(-n/2)(z_1-\bar{z}_2)^2+(-n/2)(z_2-\bar{z}_1)^2-2n(\operatorname{Im}(z_1)^2+\operatorname{Im}(z_2)^2)}}{4\pi\operatorname{Im}(z_1)\operatorname{Im}(z_2)}|\bar{z}_1 - z_2|^2 + O(n^{-1}). \tag{4.7}$$

Define $u_1, u_2 \in \mathbb{C}$ by the relations

$$z_1 = \lambda + n^{-3/4}u_1, \qquad z_2 = \lambda + n^{-3/4}u_2. \tag{4.8}$$

Note that

$$z_1 - \bar{z}_2 = z_1 - \bar{z}_1 + n^{-3/4}(\bar{u}_1 - \bar{u}_2) = 2\mathrm{i}\operatorname{Im}(z_1) + n^{-3/4}(\bar{u}_1 - \bar{u}_2) \tag{4.9}$$

and similarly

$$z_2 - \bar{z}_1 = 2\mathrm{i}\operatorname{Im}(z_2) - n^{-3/4}(\bar{u}_1 - \bar{u}_2). \tag{4.10}$$

Using (4.9), (4.10), and $\operatorname{Im}(z_1) - \operatorname{Im}(z_2) = O(n^{-3/4})$, we obtain

$$-\frac{n}{2}(z_1-\bar{z}_2)^2 - \frac{n}{2}(z_2-\bar{z}_1)^2 - 2n\big(\operatorname{Im}(z_1)^2+\operatorname{Im}(z_2)^2\big) \tag{4.11}$$

$$= 2\mathrm{i}n^{1/4}\big(\operatorname{Im}(z_2)-\operatorname{Im}(z_1)\big)(\bar{u}_1-\bar{u}_2) + O(n^{-3/2}) = O(n^{-1/2}). \tag{4.12}$$

Additionally, we have

$$|\bar{z}_1 - z_2|^2 = 4\operatorname{Im}(z_1)\operatorname{Im}(z_2) + O(n^{-3/4}). \tag{4.13}$$

Inserting (4.12) and (4.13) into (4.7), we obtain

$$Q_{23}Q_{32} = \frac{\exp\big(O(n^{-1/2})\big)}{\pi} + O(n^{-3/4}) = \frac{1}{\pi} + O(n^{-1/2}). \tag{4.14}$$

To control the error terms in (4.14), we used the fact that

$$\sup_{z,w\in\Omega:|z-w|<\tau n^{-3/4}} \big|U(z,w)\big| \le C, \tag{4.15}$$

which follows from Lemma 2.4 and Lemma 2.7.

Combining (4.5), (4.6), and (4.14), and recalling that we have $Q_{11} = O(1)$ from Lemma 2.4, we conclude that

$$Q_{11}(Q_{22}Q_{33} - Q_{23}Q_{32}) = O(n^{-1/2}). \tag{4.16}$$

Parallel reasoning (which we omit) yields

$$Q_{12}(Q_{21}Q_{33} - Q_{23}Q_{31}) = O(n^{-1/2}), \qquad Q_{13}(Q_{21}Q_{32} - Q_{31}Q_{22}) = O(n^{-1/2}). \tag{4.17}$$

Inserting (4.16) and (4.17) into (4.3), and using (4.2), we have

$$\mathbb{E}\big[\Xi^{(3)}(\mathcal{B}_M)\big] = n^3\int_\Omega\int_{D(\lambda,M_n)^2} O(n^{-1/2})\,dz_1\,dz_2\,d\lambda = O(n^{-1/2}). \tag{4.18}$$

In the last equation, we use that the area of $D(\lambda, M_n)$ is $O(n^{-3/2})$. Finally, (4.18) implies (4.21), which finishes the proof. $\square$

4.2. **Proof of Lemma 3.1.**

Proof of Lemma 3.1. Fix a constant $\tau > 0$ such that $A \subset (0, \tau)$ and set $\tau_n = n^{-3/4}\tau$ for all $n \in \mathbb{N}$. For all $z \in \mathbb{C}$ and $r \in \mathbb{R}^+$, let

$$D^+(z, r) = \{w \in \mathbb{C} : |z - w| < r, z \prec w\}.$$

We begin by showing that for all indices i such that $\lambda_i \in \Omega$ and $i^* \neq 0$, if

$$\mathbb{1}_{\{\lambda_{i^*} - \lambda_i \in B_n\}} \neq \mathbb{1}_{\{\xi^{(n)}(\lambda_i + B_n) = 1\}}, \tag{4.19}$$

then $\xi^{(n)}(D^+(\lambda_i, \tau_n)) \geq 2$. First, suppose that the left-hand side of (4.19) is 1, while the right-hand side is 0. In this case, $\xi^{(n)}(\lambda_i + B_n) \geq 2$, which implies $\xi^{(n)}(D^+(\lambda_i, \tau_n)) \geq 2$ by the definition of B_n. Next, suppose that the left-hand side of (4.19) is 0, while the right-hand side is 1. Then there exists some $j \neq i^*$ such that $\lambda_j \in \Omega$ and $\lambda_j \in \lambda_i + B_n \subset D^+(\lambda_i, \tau_n)$. Since $|\lambda_i - \lambda_{i^*}| \leq |\lambda_i - \lambda_j|$ by the definition of i^*, we have $\xi^{(n)}(D^+(\lambda_i, \tau_n)) \geq 2$. We conclude that (4.19) implies $\xi^{(n)}(D^+(\lambda_i, \tau_n)) \geq 2$. We also note that if $i^* = 0$, then the term corresponding to λ_i vanishes in both of the sums $\chi_\Omega^{(n)}(A)$ and $\widetilde{\xi}^{(n)}(\Omega)$.

From the previous paragraph, we see that

$$\left|\chi_\Omega^{(n)}(A) - \widetilde{\xi}^{(n)}(\Omega)\right| \leq \sum_{i=1}^{n-1} \mathbb{1}_{\{\xi^{(n)}(D^+(\lambda_i, \tau_n)) \geq 2\}} \leq \Xi^{(3)}(\mathcal{E}), \tag{4.20}$$

where $\Xi^{(3)}$ denotes the 3-point measure defined in (3.21) and we define

$$\mathcal{E} = \{(\lambda, z_1, z_2) : \lambda \in \Omega, (z_1, z_2) \in D^+(\lambda, \tau_n)^2\}.$$

By Lemma 3.3 and the inclusion $\mathcal{E} \subset \mathcal{B}_\tau$, we have

$$\limsup_{n\to\infty} \mathbb{E}\left[\Xi^{(3)}(\mathcal{E})\right] \leq \limsup_{n\to\infty} \mathbb{E}\left[\Xi^{(3)}(\mathcal{B}_\tau)\right] = 0. \tag{4.21}$$

By Markov's inequality and (4.20), this implies (3.4) and completes the proof. $\square$

4.3. **Proof of Lemma 3.2.** We begin by recalling the following lemma from [22, Theorem 7.8.5].

Lemma 4.1. Let M be an $n \times n$ positive-definite Hermitian matrix. For any $\mathcal{I} \subset \{1, 2, \ldots, n\}$, let $M_{\mathcal{I}}$ be the submatrix of M formed by the rows and columns with indices in $\mathcal{I}$. We have $\det(M) \leq \det(M_{\mathcal{I}}) \det(M_{\mathcal{I}^c})$.

The next lemma shows that $Q^{(k)}(z_1, \ldots, z_k)$ is positive-definite, or exponentially small, for all but an asymptotically vanishing set of $(z_1, \ldots, z_k)$. We will use it in conjunction with (4.1) to bound the correlation functions τ_k. It is proved in Section 5.

Lemma 4.2. Fix an admissible domain Ω. For all $k \in \mathbb{N}$, there exist constants $C_k(\Omega), c_k(\Omega) > 0$ such that the following holds. Define

$$\mathcal{C}_k = \{\boldsymbol{z} \in \Omega^k : Q^{(k)}(\boldsymbol{z}) \text{ is not positive definite} \wedge |n^k Q_k(\boldsymbol{z})| \geq C_k e^{-c_k n}\}, \tag{4.22}$$

and let μ denote Lebesgue measure on $\mathbb{C}^k$. Then

$$\mu(\mathcal{C}_k) < C_k e^{-c_k n}. \tag{4.23}$$

We require the following consequence of Lemma 4.2.

Lemma 4.3. Fix an admissible domain Ω. For all $k, m \in \mathbb{N}$ with $m > k$, there exist constants $C_{k,m}(\Omega), c_{k,m}(\Omega) > 0$ such that the following holds. For all $z_1, \ldots, z_k \in \Omega$ and $m > k$, define

$$\mathcal{G}_{k,m}(z_1, \ldots, z_k) = \big\{(z_{k+1}, \ldots, z_m) \in \Omega^{m-k} : (z_1, \ldots, , z_m) \in \mathcal{C}_m\big\}, \tag{4.24}$$

and

$$\mathcal{C}_{k,m} = \Big\{(z_1, \ldots, z_k) \in \Omega^k : \mu\big(\mathcal{G}_{k,m}(z_1, \ldots, z_k)\big) > C_{k,m}e^{-c_{k,m}n}\Big\}. \tag{4.25}$$

Then

$$\mu(\mathcal{C}_{k,m}) < C_{k,m}e^{-c_{k,m}n}. \tag{4.26}$$

Proof. We consider $C_{k,m}, c_{k,m}$ as parameters that will be fixed at the end of the proof. By Fubini's theorem,

$$\mu(\mathcal{C}_k) = \int_{\mathcal{C}_k} d\boldsymbol{z} \geq \int_{\mathcal{C}_{k,m}} \int_{\mathcal{G}_{k,m}} d\boldsymbol{z} \geq \int_{\mathcal{C}_{k,m}} C_{k,m}e^{-c_{k,m}N} = \mu(\mathcal{C}_{k,m})C_{k,m}e^{-c_{k,m}N}. \tag{4.27}$$

We conclude from Lemma 4.2 that

$$\frac{C_k}{C_{k,m}}e^{-(c_k - c_{k,m})n} \geq \mu(\mathcal{C}_{k,m}). \tag{4.28}$$

The conclusion follows after choosing $c_{k,m}$ such that $c_k > 2c_{k,m}$ and $C_{k,m}$ such that $C_{k,m}^2 > C_k$. □

We also recall the well-known error function asymptotic

$$\operatorname{erfc}(x) = \frac{e^{-x^2}}{\sqrt{\pi}x}\left(1 + O(x^{-2})\right), \tag{4.29}$$

which holds as $x \to \infty$.

Proof of Lemma 3.2(1). By a standard computation using the inclusion–exclusion principle (see [35, (4.5)]), we have

$$\widetilde{\tau}_k(z_1, \ldots, z_k) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \int_{z_1+B_n} dx_1 \ldots \int_{z_k+B_n} dx_k$$
$$\int_{((z_1+B_n)\cup\ldots\cup(z_k+B_n))^m} \tau_{2k+m}(z_1, x_1, \ldots, z_k, x_k, y_1, \ldots, y_m)\, dy_1 \ldots dy_m. \tag{4.30}$$

We note that [35, (4.5)] requires that $z_i \notin z_j + B_n$ for all $i \neq j$, which is true for sufficiently large n by the assumption that $(z_1, \ldots, z_k)$ has pairwise distinct entries.

We begin by analyzing the $m = 0$ term, which is

$$\int_{z_1+B_n} \cdots \int_{z_k+B_n} \tau_{2k}(z_1, x_1, \ldots, z_k, x_k)\, dx_1 \ldots dx_k. \tag{4.31}$$

By (3.5) and Lemma 2.6, we have uniformly for all $z_1, x_1, \ldots, z_k, x_k \in \Omega$ that

$$\tau_{2k}(z_1, x_1, \ldots, z_k, x_k, y_1, \ldots, y_m) = n^{2k} \det Q^{(2k)}(z_1, x_1, \ldots, z_k, x_k) + O(e^{-cn}). \tag{4.32}$$

Observe that $Q^{(2k)}$ is $2k \times 2k$ matrix that can be written as a $k \times k$ matrix of 2×2 blocks $(\widetilde{Q}_{ij})_{1 \le i,j \le k}$, where the $\widetilde{Q}_{ij}$ have the form

$$\widetilde{Q}_{ij} = \begin{pmatrix} S_n\left(\sqrt{n} z_i, \sqrt{n} z_j\right) & S_n\left(\sqrt{n} x_i, \sqrt{n} z_j\right) \\ S_n\left(\sqrt{n} z_i, \sqrt{n} x_j\right) & S_n\left(\sqrt{n} x_i, \sqrt{n} x_j\right) \end{pmatrix}.$$

For $1 \le a, b \le 2$, we will will use $\widetilde{Q}_{ij}(a,b)$ to denote the (a,b)-th entry in the (i,j)-th block of $Q^{(2k)}$.

We will analyze the diagonal and off-diagonal blocks separately. Beginning with the off-diagonal blocks where $i \neq j$, we claim that there exist constants $C, c > 0$, depending only on Ω and the values $z_1, \dots, z_k$, such that

$$\max_{i \neq j} \max_{a,b \le 2} \widetilde{Q}_{ij}(a,b) \le C e^{-cn}. \tag{4.33}$$

We will consider only the case $a = b = 1$ in detail, since the others are nearly identical.

By (4.29), uniformly for all $z_i, z_j \in \Omega$, we have the asymptotic expansion

$$S_n(\sqrt{n} z_i, \sqrt{n} z_j) = U(z_i, z_j) s_n(n z_i \bar{z}_j) \left(1 + O(n^{-1})\right). \tag{4.34}$$

By [21, Lemma 4.1], there exists a constant $C(\widetilde{d}_\Omega) > 0$ such that for all $z_i, z_j \in \Omega$,

$$s_n(n z_i \bar{z}_j) = 1 - \frac{1}{\sqrt{2\pi n}} \frac{(z_i \bar{z}_j e^{1 - z_i \bar{z}_j})^n}{1 - z_i \bar{z}_j} \big(1 + R(z_i \bar{z}_j; n)\big), \qquad \big|R(z_i \bar{z}_j; n)\big| \le C n^{-1}. \tag{4.35}$$

Inserting this estimate into (4.34), we find

$$|S_n(\sqrt{n} z_i, \sqrt{n} z_j)| \le C e^{-(n/2) \operatorname{Re}(z_i - \bar{z}_j)^2} e^{-n(\operatorname{Im}(z_i)^2 + \operatorname{Im}(z_j)^2)} \tag{4.36}$$

$$\times \left(1 + C \left| e^{-2(1 - z_i \bar{z}_j)} (z_i \bar{z}_j)^n \right| \right). \tag{4.37}$$

Next, we note that there exists a constant $c(z_1, \dots, z_n) > 0$ such that

$$e^{-(n/2) \operatorname{Re}(z_i - \bar{z}_j)^2} e^{-n(\operatorname{Im}(z_i)^2 + \operatorname{Im}(z_j)^2)} = e^{-(n/2)|z_i - z_j|^2} \le e^{-cn}, \tag{4.38}$$

and

$$\left| e^{-(n/2) \operatorname{Re}(\lambda - \bar{z}_j)^2} e^{-n(\operatorname{Im}(z_i)^2 + \operatorname{Im}(z_j)^2)} \right| \le e^{(n/2)(-|z_i - \bar{z}_j|^2 + 2 + 2 \operatorname{Re}(z_i \bar{z}_j) - 2 \ln |z_i \bar{z}_j|} \tag{4.39}$$

$$= e^{-(n/2)(|z_i|^2 + |z_j|^2 - 2 - \ln |z_i|^2 - \ln |z_j|^2))} \tag{4.40}$$

$$\le e^{-cn}, \tag{4.41}$$

where the last inequality follows from $|z|^2 - 1 - \ln |z|^2 \ge c > 0$ for $z \in \Omega$ (which holds by the convexity of $x \mapsto \ln x$). This completes the proof of (4.33).

We now consider the diagonal blocks, corresponding to the case $i = j$. By Lemma 2.7, these take the form

$$\begin{pmatrix} S_n(z_i, z_i) & S_n(z_i, x_i) \\ S_n(x_i, z_i) & S_n(x_i, x_i) \end{pmatrix} = \begin{pmatrix} U(z_i, z_i) & U(z_i, x_i) \\ U(x_i, z_i) & U(x_i, x_i) \end{pmatrix} + O(n^{-1}), \tag{4.42}$$

where the additive error term denotes a matrix such that each entry is $O(n^{-1})$. Since $|U(z,w)| \le C$ uniformly for $z, w \in \Omega$, by Lemma 2.7 and Lemma 2.4, we have

$$\det \begin{pmatrix} S_n(z_i,z_i) & S_n(z_i,x_i) \\ S_n(x_i,z_i) & S_n(x_i,x_i) \end{pmatrix} = \det \begin{pmatrix} U(z_i,z_i) & U(z_i,x_i) \\ U(x_i,z_i) & U(x_i,x_i) \end{pmatrix} + O(n^{-1}). \tag{4.43}$$

We now compute the determinant on the right-hand side of (4.43). Beginning with the term coming from the diagonal of this block, by direct substitution, we observe that

$$U(z_i,z_i) = U(x_i,x_i) = \frac{1}{\pi}. \tag{4.44}$$

For the term coming from the cross-diagonal, observe first that

$$\begin{aligned} U(z_i,x_i)U(x_i,z_i) &= \frac{-1}{4\pi^2 \operatorname{Im}(x_i)\operatorname{Im}(z_i)} \cdot (\bar{x}_i - z_i)(\bar{z}_i - x_i) \\ &\quad \times \exp\left(-\frac{n}{2}\left[(z_i-\bar{x}_i)^2 + (x_i - \bar{z}_i)^2\right] - 2n\left[\operatorname{Im}(z_i)^2 + \operatorname{Im}(x_i)^2\right]\right) \end{aligned} \tag{4.45}$$

Following the same strategy as in (4.8), since $x_i \in z_i + B_n$, we will define $u_i \in \mathbb{C}$ by the equality

$$x_i = z_i + n^{-3/4}u_i. \tag{4.46}$$

We now return to (4.45) and analyze each factor separately. First, we have

$$\begin{aligned} \frac{-1}{4\pi^2 \operatorname{Im}(x_i)\operatorname{Im}(z_i)} &= \frac{-1}{4\pi^2 \operatorname{Im}(z_i)\left(\operatorname{Im}(z_i) + n^{-3/4}\operatorname{Im}(u_i)\right)} \\ &= \frac{-1}{4\pi^2 \operatorname{Im}(z_i)^2} + O(n^{-3/4}). \end{aligned} \tag{4.47}$$

Secondly, we have

$$\begin{aligned} (\bar{x}_i - z_i)(\bar{z}_i - x_i) &= (-2i\operatorname{Im}(z_i) + n^{-3/4}\bar{u}_i)(-2i\operatorname{Im}(z_i) - n^{-3/4}u_i) \\ &= -4\operatorname{Im}(z_i)^2 + O(n^{-3/4}). \end{aligned} \tag{4.48}$$

Next, in the exponent, we have

$$\begin{aligned} -\frac{n}{2}&\left((z_i-\bar{x}_i)^2 + (x_i-\bar{z}_i)^2\right) \\ &= -\frac{n}{2}\left((2i\operatorname{Im}(z_i) - n^{-3/4}\bar{u}_i)^2 + (2i\operatorname{Im}(z_i) + n^{-3/4}u_i)^2\right) \\ &= -\frac{n}{2}\left(-8\operatorname{Im}(z_i)^2 - 8n^{-3/4}\operatorname{Im}(z_i)\operatorname{Im}(u_i) + n^{-3/2}(u_i^2 + \bar{u}_i^2)\right) \\ &= 4n\operatorname{Im}(z_i)^2 + 4n^{1/4}\operatorname{Im}(z_i)\operatorname{Im}(u_i) - n^{-1/2}(\operatorname{Re}(u_i)^2 - \operatorname{Im}(u_i)^2), \end{aligned} \tag{4.49}$$

and, for the other term in the exponent,

$$\begin{aligned} -2n\left(\operatorname{Im}(z_i)^2 + \operatorname{Im}(x_i)^2\right) &= -2n\left(2\operatorname{Im}(z_i)^2 + 2n^{-3/4}\operatorname{Im}(z_i)\operatorname{Im}(u_i) + n^{-3/2}\operatorname{Im}(u_i)^2\right) \\ &= -4n\operatorname{Im}(z_i)^2 - 4n^{-1/4}\operatorname{Im}(z_i)\operatorname{Im}(u_i)^2 - 2n^{-1/2}\operatorname{Im}(u_i)^2. \end{aligned} \tag{4.50}$$

Summing (4.49) and (4.50), we see that the exponent is

$$-n^{-1/2}(\mathrm{Re}(u_i)^2 - \mathrm{Im}(u_i)^2) - 2n^{-1/2}\,\mathrm{Im}(u_i)^2 = -n^{-1/2}|u_i|^2 \tag{4.51}$$

$$= -n|z_i - x_i|^2, \tag{4.52}$$

noting that $u_i = n^{3/4}(x_i - z_i)$. So, (4.45) simplifies to

$$U(z_i, x_i)U(x_i, z_i) = \frac{1}{\pi^2}\exp\left(-n|z_i - x_i|^2\right) + O(n^{-3/4}). \tag{4.53}$$

Using (4.44) and (4.53), we compute the determinant on the right hand side of (4.43) and find

$$\det\begin{pmatrix} U(z_i, z_i) & U(z_i, x_i) \\ U(x_i, z_i) & U(x_i, x_i)\end{pmatrix} + O(n^{-1}) = \pi^{-2}\left(1 - \exp(-n|z_i - x_i|^2)\right) + O(n^{-3/4}). \tag{4.54}$$

Returning to (4.32), and using (4.33), we see that the only non-negligible contributions to the determinant of $Q^{(2k)}$ are from the 2×2 blocks along the diagonal. Up to an additive error of $O(n^{-3/4})$, this is

$$\prod_{i=1}^{k}\left(n^2\pi^{-2}\int_{z_i+B_n}(1 - \exp(-n|z_i - x_i|^2)\,dx_i\right). \tag{4.55}$$

It suffices to compute each integral from this product individually. Making the change of variable $u = n^{3/4}(x_i - z_i)$, we compute

$$n^2\int_{z_i+B_n}\left(1 - \exp(-n|z_i - x_i|^2\right)dx_i = n^{1/2}\int_B\left(1 - \exp(-n^{-1/2}|u|^2\right)du \tag{4.56}$$

$$= \int_B |u|^2\,du + O(n^{-1/2}). \tag{4.57}$$

We conclude that (4.55) equals

$$\left(\frac{1}{\pi^2}\int_B |u|^2\,du\right)^k + O(n^{-1/2}). \tag{4.58}$$

For future reference, we also note that (4.54) implies that

$$\sup_{z\in\Omega}\left|\int_{z+B_n} n^2\det Q^{(2)}(z,x)\right| \le C. \tag{4.59}$$

It remains to show that the contributions from the terms in (4.30) with $m \ge 1$ are negligible. We begin with the indices $m \ge k+5$. We can bound such a term by

$$\left|\frac{(-1)^m}{m!}\int_{z_1+B_n}\cdots\int_{z_k+B_n}\int_{((z_1+B_n)\cup\ldots\cup(z_k+B_n))^m}\tau_{2k+m}(z_1, x_1, \ldots, z_k, x_k, \boldsymbol{y})\,d\boldsymbol{y}\,d\boldsymbol{x}\right|$$

$$\le \frac{1}{m!}\cdot C^m k^m n^{-3m/2}\cdot C^k n^{-3k/2}\cdot C^{2k+m}n^{2k+m} = \frac{C^{3k+2m}}{m!}k^m n^{(k-m)/2}, \tag{4.60}$$

where we bound the area of integration of $\boldsymbol{y}$ by $(Ckn^{-3/2})^m$, we bound the area of integration of $\boldsymbol{x}$ by $C^k n^{-3k/2}$, and we use the bound $\tau_{2k+m} \leq C^{2k+m} n^{2k+m}$ coming from (2.2), Remark 2.3, and Lemma 2.4. (Observe that the C in the last inequality is independent of k and m.) We conclude that the sum of all terms in (4.30) with $m > k+5$ is $O(n^{-1/2})$.

Next, we consider the terms with $m < k+5$. Fix $m < k+5$ and set $\ell = 2k+m$. We use the notation $C_{k,\ell}^{(n)}$ to make the dependence of the set $C_{k,\ell}$ on n explicit. Using Lemma 4.3, we compute

$$\sum_{n=1}^{\infty} \mu\big(\mathcal{C}_{k,\ell}^{(n)}\big) \leq C_{k,\ell} \sum_{n=1}^{\infty} e^{-c_{k,\ell} n} < \infty. \tag{4.61}$$

Therefore, by the Borel–Cantelli lemma, for every $\boldsymbol{z} \in \Omega^k$ (except a set of measure zero), there exists $n_0(\boldsymbol{z})$ such that $\boldsymbol{z} \in C_{k,\ell}^c$ for every $n \geq n_0$. Since we are proving an asymptotic statement, we suppose $n \geq n_0$ for the rest of the proof. Then, since $\mathcal{G}_{k,\ell}(z_1, \ldots, z_k)$ is exponentially small for such $\boldsymbol{z} \in \Omega^k$ (by the definition of $\mathcal{C}_{k,\ell}$), it remains to bound the integral over $\mathcal{G}_{k,\ell}^c(z_1, \ldots, z_k)$ in the m-th term of (4.30).

Define

$$\mathcal{A}_\ell = \{(z_{k+1}, \ldots, z_\ell \in \Omega^{\ell-k} : Q^{(\ell)}(\mathbf{z}) \text{ is positive definite}\}, \tag{4.62}$$

and set $\mathcal{B}_\ell = \mathcal{G}_{k,\ell}^c(z_1, \ldots, z_k) \setminus \mathcal{A}_\ell$. By the definition of $\mathcal{G}_{k,\ell}$, for $(z_{k+1}, \ldots, z_\ell) \in \mathcal{B}_\ell$ we have

$$n^\ell \big| \det Q^{(\ell)}(z_1, \ldots, z_\ell) \big| \leq Ce^{-cn} \tag{4.63}$$

for some constants $C, c > 0$ depending only on k and Ω. In particular, the constants are uniform in ℓ (and hence m) by the assumed upper bound on m. Using (4.63), (4.32), and that τ_{2k+m} is symmetric in its arguments, we can bound the portion of the integral in the m-th term over the region $\mathcal{B}_\ell$ by

$$\int_{\mathcal{B}_\ell} \det Q^{(2k+m)}(z_1, \ldots, z_k, x_1, \ldots, x_k, y_1, \ldots, y_m) \, d\boldsymbol{x} \, d\boldsymbol{y} \leq Ce^{-cn}, \tag{4.64}$$

after increasing the value of C.

Finally, we consider the integral over $\mathcal{A}_\ell$ where $Q^{(2k+m)}$ is positive-definite. We observe that the integrand here can be estimated by

$$\mathbb{1}_{\mathcal{A}_\ell} \det Q^{(2k+m)}(z_1, \ldots, z_k, \boldsymbol{x}, \boldsymbol{y}) \leq \mathbb{1}_{\mathcal{A}_\ell} \det Q^{(2k)}(z_1, \ldots, z_k, \boldsymbol{x}) \left(\prod_{i=1}^m \det Q^{(1)}(y_i) \right) \tag{4.65}$$

by iteratively applying Lemma 4.1 since the matrices $M_{\mathcal{I}}$ and $M_{\mathcal{I}^c}$ as defined in Lemma 4.1 are positive-definite whenever M is (since all principal minors of a positive-definite matrix are positive-definite). By Lemma 2.4, we have

$$\left(\prod_{i=1}^m \det Q^{(1)}(y_i) \right) \leq C^m. \tag{4.66}$$

Bounding (4.65) using the estimates (4.65) and (4.66), we find

$$\left| \int_{((z_1+B_n)\cup\ldots\cup(z_k+B_n))^m} \int_{z_1+B_n} \cdots \int_{z_k+B_n} \mathbb{1}_{\mathcal{A}_\ell} n^{2k+m} \det Q^{(2k+m)}(z_1,\ldots,z_k,\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{x}\,d\boldsymbol{y} \right|$$
$$\leq Cn^m \cdot n^{-3m/2} \int_{z_1+B_n} \cdots \int_{z_k+B_n} \mathbb{1}_{\mathcal{A}_\ell} n^{2k} \det Q^{(2k)}(z_1,\ldots,z_k,\boldsymbol{x})$$
$$\leq Cn^m \cdot n^{-3m/2} = Cn^{-m/2}, \quad (4.67)$$

where the constant C depends on m and Ω, and changes at each appearance. The bound in (4.67) is $o(1)$, which completes the proof. □

Proof of Lemma 3.2(2). We retain the notations from the previous proof. Our goal is to prove a uniform upper bound on (4.30). The same argument that gave (4.60) shows that the that the sum of all terms in (4.30) with $m \geq k+5$ is $O(n^{-1/2})$. For the terms with $m < k+5$, the previous Borel–Cantelli argument no long suffices, as it is not uniform in n. Therefore, we reason as follows. Define $\mathcal{X}_k = \mathcal{X}_k^{(n)} \subset \Omega^k$ by

$$\mathcal{X}_k = \mathcal{C}_k^c \cap \left(\bigcap_{m=1}^{k+4} \mathcal{C}_{k,m}^c \right). \quad (4.68)$$

For $\boldsymbol{z} \in \mathcal{X}_k$, the same argument that gave (4.67) shows the the terms with $0 \leq m \leq k+4$ in (4.30) are uniformly bounded by a constant $C > 0$ (which depends on k and Ω). By (4.2) and (4.3), there exist constants $C, c > 0$ (depending on k) such that

$$\mathbb{P}(\mathcal{X}_k^c) \leq Ce^{-cn}. \quad (4.69)$$

Then with the choice $\mathcal{W}_n = \mathcal{X}_k^{(n)}$, we have established (3.8) and (3.9). For (3.10), we note that (2.2) and (2.4) together show that that there exists a constant $\widetilde{C} > 0$ such that for every $k \in \mathbb{N}$, we have

$$\sup_{z_1,\ldots,z_k \in \Omega} \tau_k(z_1,\ldots,z_k) \leq \widetilde{C}^k n^k. \quad (4.70)$$

We emphasize that $\widetilde{C}$ does not depend on k. Using (4.70) in the terms with $0 \leq m \leq k+4$ in (4.30) shows that their sum is bounded by

$$Cn^{3k+4} \leq Cn^{8k} \quad (4.71)$$

for some $C(k) > 0$, where we use $k \geq 1$. This shows that (3.10) holds and completes the proof. □

Proof of Lemma 3.2(3). Consider an arbitrary point $(z_1,\ldots,z_k) \in \overline{\Psi}_k$. We begin by defining an equivalence relation $\sim$ on the set $\{1,\ldots,k\}$ as follows. This equivalence relation is determined by two conditions. First, we have $i \sim j$ if $|z_i - z_j| \in B_n$. Second, we also have $i \sim j$ if there exists a sequence $i_1,\ldots,i_s$ of indices such that $i_1 = i$, $i_s = j$, and $z_{i_m} \sim z_{m+1}$ for all $1 \leq m < s$. Let p denote the number of equivalence classes under $\sim$. For each $r \in \mathbb{N}$ with $1 \leq r \leq p$, let i_r denote the index in the r-th equivalence class such that z_{i_r} is the maximal element in this equivalence class under $\prec$. (The set of $\boldsymbol{z} \in \overline{\Psi}_k$ where there is not a unique maximal element in each class has measure zero, and therefore can be neglected.)

By the definition of $\widetilde{\xi}^{(n)}$ and Lemma 2.6, there exist $C(\Omega), c(\Omega) > 0$ such that

$$\widetilde{\tau}_k(z_1, \dots, z_k) \tag{4.72}$$
$$\leq \int_{z_{i_1}+B_n} \cdots \int_{z_{i_p}+B_n} \tau_{k+p}(z_1, \dots, z_k, x_1, \dots, x_p)\, dx_1 \dots dx_p \tag{4.73}$$
$$\leq n^{k+p} \int_{z_{i_1}+B_n} \cdots \int_{z_{i_p}+B_n} \det Q^{(k+p)}(z_1, \dots, z_k, x_1, \dots, x_p)\, dx_1 \dots dx_p + Ce^{-cn}. \tag{4.74}$$

We define $\mathcal{Y}_k \subset \Omega^k$ by

$$\mathcal{Y}_k = \mathcal{C}_k^c \cap \left(\bigcap_{m=1}^{2k} \mathcal{C}_{k,m}^c \right). \tag{4.75}$$

From (4.2) and (4.3), there exist constants $C, c > 0$ such that

$$\mathbb{P}(\mathcal{Y}_k) > 1 - Ce^{-cn}. \tag{4.76}$$

Note that by Lemma 2.4 and Lemma 2.6, there exists a constant $C(\Omega) > 0$ such that

$$\left| \det Q^{(k+p)}(z_1, \dots, z_k, x_1, \dots, x_p) \right| \leq C^{k+p} \leq C^{2k}, \tag{4.77}$$

since $p < k$ by the definitions of p and $\overline{\Psi}_k$. Then (4.76) and (4.77) together imply that

$$\lim_{n \to \infty} \int_{\mathcal{Y}_k^c} \widetilde{\tau}_k(z_1, \dots, z_k) = 0. \tag{4.78}$$

It remains to consider the analogous integral over the region $\mathcal{Y}_k \cap \overline{\Psi}_k$.

Returning to (4.74), we let $\mathcal{A}_{k,p} \subset \mathcal{Y}_k$ denote the set of points $(z_1, \dots, z_{k+p})$ such that $Q^{(k+p)}(z_1, \dots, z_{k+p})$ is positive-definite, and set $\mathcal{B}_{k,p} = \mathcal{Y}_k \setminus \mathcal{A}_{k,p}$. On $\mathcal{B}_{k,p}$, we have $|\tau_{k+p}(\boldsymbol{z})| \leq Ce^{-cn}$ uniformly for all $\boldsymbol{z}$ and all $p < k$, so the contribution from $\mathcal{B}_{k,p}$ is negligible by the same argument that gave (4.78).

On $\mathcal{A}_{k,p}$, we have after using Lemma 4.1 multiple times that

$$n^{k+p} \int_{z_{i_1}+B_n} \cdots \int_{z_{i_p}+B_n} \mathbb{1}_{\mathcal{A}_{k,p}} \det Q^{(k+p)}(z_1, \dots, z_k, x_1, \dots, x_p)\, dx_1 \dots dx_p$$
$$\leq \prod_{j \neq i_1, \dots i_p} n \det Q^{(1)}(z_j) \cdot \prod_{j=1}^{p} \int_{z_{i_j}+B_n} n^2 \det Q^{(2)}(z_{i_j}, x_j)\, dx_j. \tag{4.79}$$

It follows from the estimate (4.59) above that

$$\sup_{z_1, \dots, z_p \in \Omega} \left| \prod_{j=1}^{p} \int_{z_j+B_n} n^2 \det Q^{(2)}(z_j, x_j)\, dx_j \right| \leq C^p. \tag{4.80}$$

We also have, using (2.2), Remark 2.3, and Lemma 2.5, that

$$\sup_{z_1,\dots,z_{k-p}} \left| \prod_{j=1}^{k-p} n \det Q^{(1)}(z_j) \right| \le (Cn)^{k-p}. \tag{4.81}$$

For a given p, the measure of the region of integration in (4.79) is bounded by $C_p n^{-3(k-p)/2}$. Then inserting (4.80) and (4.81) into the right-hand side of (4.79), we find that the integral is bounded is $O(n^{-k/2+p/2})$. Since $k > p$ by the definition of $\overline{\Psi}_k$, this finishes the proof. □

5. Proofs of Auxiliary Lemmas

The following proof is essentially the same as that of [21, Lemma 3.3]. We repeat it here for completeness.

Proof of Lemma 2.6. We recall from (2.2) that $\rho_k(\sqrt{n}z_1, \dots, \sqrt{n}z_k)$ can be written explicitly as a Pfaffian. Using Lemma 2.4, all terms in this expansion involving a factor D_n or I_n are exponentially small. We conclude that

$$\sup_{z_1,\dots,z_k\in\Omega} \left| n^k \rho_k(\sqrt{n}z_1, \dots, \sqrt{n}z_k) - n^k \operatorname{Pf}\left(\widetilde{K}(\sqrt{n}z_i, \sqrt{n}z_j)\right)_{1\le i,j\le k}, \right| \le c^{-1}e^{-cn},$$

where $(\widetilde{K}(z_i, z_j))_{1\le i,j\le k}$ is a $2k \times 2k$ matrix composed of the 2×2 blocks

$$\widetilde{K}(z_i, z_j) = \begin{pmatrix} 0 & S_n(z_i, z_j) \\ -S_n(z_j, z_i) & 0 \end{pmatrix}.$$

The conclusion follows after noting that Lemma 2.5 implies

$$\operatorname{Pf}\left(\widetilde{K}(\sqrt{n}z_i, \sqrt{n}z_j)\right)_{1\le i,j\le k} = \det Q^{(k)}(z_1, \dots, z_k).$$

□

Proof of Lemma 2.7. By (4.29), for all $z, w \in \Omega$, we have the asymptotic expansion

$$S_n(\sqrt{n}z, \sqrt{n}w) = U(z,w)s_n(nz\bar{w})\left(1 + O(n^{-1})\right). \tag{5.1}$$

By [21, Lemma 4.1], there exists a constant $C(\widetilde{d}_\Omega) > 0$ such that for all $z, w \in \Omega$,

$$s_n(nz\bar{w}) = 1 - \frac{1}{\sqrt{2\pi n}} \frac{(z\bar{w}e^{1-z\bar{w}})^n}{1 - z\bar{w}}\big(1 + R(z\bar{w}; n)\big), \qquad \big|R(z\bar{w}; n)\big| \le Cn^{-1}. \tag{5.2}$$

Considering the case $z = w$, the exponential factor becomes

$$|z|^2 e^{1-|z|^2} = \exp\left(1 - |z|^2 + \log(|z|^2)\right).$$

There exists a constant $\varepsilon(\Omega) > 0$ such that for all $z \in \Omega$, we have $\varepsilon < |z|^2 < 1 - \varepsilon$. Then, using concavity of the logarithm, we find that there exists $\delta(\Omega) > 0$ such that for all $z \in \Omega$,

$$1 - |z|^2 + \log(|z|^2) < -\delta.$$

By continuity, it follows that there exists $n_0(r, \Omega) > 0$ such that for all $z, w \in \Omega$ such that $|z - w| < rn^{-3/4}$, and all $n > n_0$,

$$\operatorname{Re}\left(1 - z\bar{w} - \log(z\bar{w})\right) < -\delta/2, \qquad |z\bar{w}e^{1-z\bar{w}}|^n \le e^{-\delta n/2}. \tag{5.3}$$

Inserting (5.3) into (5.2) shows that

$$s_n(nz\bar{w}) = 1 + O(n^{-1}), \tag{5.4}$$

after noting that $(1 - z\bar{w})^{-1}$ is uniformly bounded above for all $z, w \in \Omega$. Then inserting (5.4) into (5.1) completes the proof (after recalling that the left-hand side of (5.1) is $O(1)$, by Lemma 2.4). □

Proof of Proposition 2.8. Note that $X_n \to \chi(J)$ in distribution is implied by the statement that $\mathbb{E}[X^r] \to \mathbb{E}[X^r]$ for all $r \in \mathbb{N}$, by [4, Theorem 30.1] and [4, Theorem 30.2].[3] Convergence of all moments $\mathbb{E}[\chi^{(n)}(J)^r]$ is implied by the convergence of the factorial moments on the left-hand side of (2.4), since the factorial moments of order at most r determine the moment $\mathbb{E}[\chi^{(n)}(J)^r]$, and this completes the proof after recalling the standard fact that the r-th factorial moment of a Poisson random variable with rate λ is λ^r (which follows from a straightforward computation using the probability density function). □

Proof of Proposition 2.9. This follows directly from [23, Theorem 4.15]; see the comment immediately before its statement. Note that our hypothesis on μ implies that χ is simple, as required by theorem. We are also using that the collection of bounded Borel subset of $\mathbb{R}$ forms a *dissecting ring* according to the definition of [23, p. 24]. □

For the proof of Lemma 4.2, we need the following lemma.

Lemma 5.1. Fix an admissible domain Ω and $k, m, n \in \mathbb{N}$ such that $m < k < n$. Then

$$\tau_m(z_1, \ldots, z_m) = \frac{(n-k)!}{(n-m)!} \int_{\Omega^{k-m}} \tau_k(z_1, \ldots, z_k)\, dz_{m+1} \ldots dz_k. \tag{5.5}$$

Proof. Fix a bounded Borel function $f : \Omega^m \to \mathbb{R}$, and define $\varphi : \Omega^k \to \mathbb{R}$ by $\varphi(z_1, \ldots, z_k) = f(z_1, \ldots, z_m)$. We know that f satisfies the equality in (2.1) for τ_m and that φ satisfies the analogous equality for τ_k. Analyzing the term on the right-hand side of (2.1), we have that

$$\sum_{(i_1,\ldots,i_k)\in\mathcal{I}_k} \varphi(w_{i_1}, \ldots, w_{i_k}) = \sum_{(i_1,\ldots,i_k)\in\mathcal{I}_k} f(w_{i_1}, \ldots, w_{i_m}) \tag{5.6}$$

$$= \sum_{(i_1,\ldots,i_m)\in\mathcal{I}_m} \frac{(n-m)!}{(n-k)!} f(w_{i_1}, \ldots, w_{i_m}). \tag{5.7}$$

Using (2.1), this implies that

$$\int_{\Omega^k} \varphi(z_1, \ldots, z_k)\tau_k(z_1, \ldots, z_k)\, dz_1 \ldots dz_k = \frac{(n-m)!}{(n-k)!} \int_{\Omega^m} f(z_1, \ldots, z_m)\tau_m(z_1, \ldots, z_m)\, dz_1 \ldots dz_m. \tag{5.8}$$

[3]For the application of [4, Theorem 30.1], we recall that the moment generating function of a Poisson distribution with rate λ is given by $\exp(\lambda(e^t - 1))$, whose power series converges for all $t \in \mathbb{R}$.

Additionally, by Fubini's theorem and the definition of φ,

$$\int_{\Omega^k} \varphi(z_1,\ldots,z_k)\tau_k(z_1,\ldots,z_k)\,dz_1\ldots dz_k = \int_{\Omega^m} f(z_1,\ldots,z_m)\left(\int_{\Omega^{k-m}} \tau_k(z_1,\ldots,z_k)\,dz_{m+1}\ldots dz_k\right) dz_1\ldots dz_m. \tag{5.9}$$

Since f was arbitrary, we conclude from comparing the right-hand sides of (5.8) and (5.9) that (5.5) holds. □

Proof of Lemma 4.2. Define

$$\mathcal{A} = \{\mathbf{z} \in \Omega^k : Q^{(k)}(\mathbf{z}) \text{ is positive definite}\} \tag{5.10}$$

and let $\mathcal{D} = \Omega^k \setminus \mathcal{A}$. We let M_i denote the i-th leading principal minor of $Q^{(k)}$, obtained by removing the last $k-i$ rows and columns. For $1 \le j \le k$, we define

$$\mathcal{D}_j = \{\boldsymbol{z} \in \Omega^k \mid \det M_j(\boldsymbol{z}) \le 0 \wedge \forall \ell < j \ \det M_\ell(\boldsymbol{z}) > 0\}. \tag{5.11}$$

In other words, $\mathcal{D}_j$ is the set of $\boldsymbol{z}$ where the smallest principal minor with a non-positive determinant is the j-th one. By Sylvester's criterion ([22, Theorem 7.2.5(b)]), we may write $\mathcal{D}$ as the disjoint union of $\mathcal{D}_j$ for $j = 1,\ldots,k$:

$$\mathcal{D} = \mathcal{D}_k \sqcup \mathcal{D}_{k-1} \sqcup \ldots \sqcup \mathcal{D}_1. \tag{5.12}$$

Note that membership in $\mathcal{D}_m$ is determined by the first m coordinates, meaning that if $(z_1,\ldots,z_k) \in \mathcal{D}_m$, then for all $(w_{m+1},\ldots,w_k) \in \Omega^{k-m}$, we have

$$(z_1,\ldots,z_m,w_{m+1},\ldots,w_k) \in \mathcal{D}_m. \tag{5.13}$$

With this in mind, we define

$$\mathcal{D}'_m = \{(z_1,\ldots,z_m) \in \Omega^m : \exists (z_{m+1},\ldots,z_k) \in \Omega^{k-m} \text{ with } (z_1,\ldots,z_k) \in \mathcal{D}_m\}. \tag{5.14}$$

Fix $m \le k$. By Lemma 2.6, there exists $c(k,d_\Omega) > 0$ such that for all $(z_1,\ldots,z_m) \in \mathcal{D}'_m$, we have

$$\tau_m(z_1,\ldots,z_m) \le c^{-1}e^{-cn}, \tag{5.15}$$

because $\tau_m \ge 0$ and $\det Q^{(m)}(z_1,\ldots,z_m) \le 0$ on $\mathcal{D}_m$ (since $Q^{(m)}$ is the m-th principal minor of $Q^{(k)}$).

By Lemma 5.1, it follows that there exists a constant $\hat{c} > 0$ such that $c > \hat{c}$, and for all $(z_1,\ldots,z_m) \in \mathcal{D}'_m$,

$$\begin{aligned}\int_{\Omega^{k-m}} \tau_k(z_1,\ldots,z_m,w_{m+1},\ldots,w_k)\,dw_{m+1}\ldots dw_k &= \frac{(n-m)!}{(n-k)!}\tau_m(z_1,\ldots,z_m), && (5.16)\\ &\le n^k \tau_m(z_1,\ldots,z_m) && (5.17)\\ &\le \hat{c}^{-1}e^{-\hat{c}n}. && (5.18)\end{aligned}$$

Now, if we define

$$D(z_1,\ldots,z_m) = \{(z_{m+1},\ldots,z_k) \in \Omega^{k-m} : \tau_k(z_1,\ldots,z_k) \ge \hat{c}^{-1}e^{-\hat{c}n/2}\}, \tag{5.19}$$

then by (5.18) and since $\tau_k \geq 0$, we have for all $(z_1, \ldots, z_m) \in \Omega^m$ that

$$\int_{D(z_1,\ldots,z_m)} \tau_k(z_1, \ldots, z_m, w_{m+1}, \ldots, w_k)\, dw_{m+1} \ldots dw_k \leq \hat{c}^{-1} e^{-\hat{c}n},$$

which implies by the definition of $D(z_1, \ldots, z_m)$ that

$$\mu(D(z_1, \ldots, z_m)) \leq e^{-\hat{c}n/2}.$$

Define $\mathcal{C}_k^{(m)} \subset \mathcal{D}_k$ by

$$\mathcal{C}_k^{(m)} = \big\{(z_1, \ldots, z_k) \in \mathcal{D}_k : \tau_k(z_1, \ldots, z_k) \geq \hat{c}^{-1} e^{-\hat{c}n/2}\big\}. \tag{5.20}$$

Then

$$\mu\big(\mathcal{C}_k^{(m)}\big) = \int_{(z_1,\ldots,z_m)\in\mathcal{D}'_m} \int_{D(z_1,\ldots,z_m)} d\boldsymbol{z} \leq |\Omega|^m \hat{c}^{-1} e^{-\hat{c}n} \leq 2^m \hat{c}^{-1} e^{-\hat{c}n}. \tag{5.21}$$

Since $m \leq k$, we deduce the existence of a constant $\widetilde{c} > 0$ such that $\hat{c}/2 > \widetilde{c}$ and

$$\sum_{m=1}^{k} \mu\big(\mathcal{C}_k^{(m)}\big) \leq \widetilde{c}^{-1} e^{-\widetilde{c}n}. \tag{5.22}$$

Then the desired conclusion holds after setting $c_k = \widetilde{c}$ and $C_k = \widetilde{c}^{-1}$, since with these choices, we have

$$\mathcal{C}_k \subset \bigcup_{m=1}^{k} \mathcal{C}_k^{(m)}, \tag{5.23}$$

and hence

$$\mu(\mathcal{C}_k) \leq \mu\left(\bigcup_{m=1}^{k} \mathcal{C}_k^{(m)}\right) \leq \sum_{m=1}^{k} \mu\big(\mathcal{C}_k^{(m)}\big) \leq \widetilde{c}^{-1} e^{-\widetilde{c}n} \tag{5.24}$$

by (5.22) and a union bound. $\square$

References

[1] Yacin Ameur. Repulsion in low temperature β-ensembles. *Communications in Mathematical Physics*, 359(3):1079–1089, 2018.

[2] Yacin Ameur and José Luis Romero. The planar low temperature Coulomb gas: separation and equidistribution. *Revista matemática iberoamericana*, 39(2):611–648, 2023.

[3] Gerard Ben Arous and Paul Bourgade. Extreme gaps between eigenvalues of random matrices. *The Annals of Probability*, 41(4):2648–2681, 2013.

[4] Patrick Billingsley. *Probability and Measure*. John Wiley & Sons, third edition, 2008.

[5] Pavel Bleher and Robert Mallison Jr. Zeros of sections of exponential sums. *International Mathematics Research Notices*, 2006:38937, 2006.

[6] Oriol Bohigas, Marie-Joya Giannoni, and Charles Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Physical Review Letters*, 52(1):1, 1984.

[7] Charles Bordenave and Djalil Chafaï. Around the circular law. *Probability Surveys*, 9:1–89, 2012.

[8] Alexei Borodin and Christopher D. Sinclair. The Ginibre ensemble of real random matrices and its scaling limits. *Communications in Mathematical Physics*, 291:177–224, 2009.

[9] Paul Bourgade. Extreme gaps between eigenvalues of Wigner matrices. *Journal of the European Mathematical Society*, 24(8):2823–2873, 2021.

[10] Sung-Soo Byun and Peter J. Forrester. Progress on the study of the Ginibre ensembles I: GinUE. *arXiv preprint arXiv:2211.16223*, 2022.

[11] Sung-Soo Byun and Peter J. Forrester. Progress on the study of the Ginibre ensembles II: GinOE and GinSE. *arXiv preprint arXiv:2301.05022*, 2023.

[12] Percy Deift, Tara Nanda, and Carlos Tomei. Ordinary differential equations and the symmetric eigenvalue problem. *SIAM Journal on Numerical Analysis*, 20(1):1–22, 1983.

[13] Renjie Feng, Gang Tian, and Dongyi Wei. Small gaps of GOE. *Geometric and Functional Analysis*, 29(6):1794–1827, 2019.

[14] Renjie Feng and Dongyi Wei. Small gaps of circular β-ensemble. *The Annals of Probability*, 49(2):997–1032, 2021.

[15] Alessio Figalli and Alice Guionnet. Universality in several-matrix models via approximate transport maps. *Acta Mathematica*, 217(1):115–159, 2016.

[16] Peter J. Forrester. *Log-Gases and Random Matrices*. Princeton University Press, 2010.

[17] Peter J. Forrester and Anthony Mays. A method to calculate correlation functions for $\beta = 1$ random matrices of odd size. *Journal of Statistical Physics*, 134(3):443–462, 2009.

[18] Peter J. Forrester and Taro Nagao. Eigenvalue statistics of the real Ginibre ensemble. *Physical Review Letters*, 99(5):050603, 2007.

[19] Peter J. Forrester and Taro Nagao. Skew orthogonal polynomials and the partly symmetric real Ginibre ensemble. *Journal of Physics A: Mathematical and Theoretical*, 41(37):375003, 2008.

[20] Martin Gebert and Mihail Poplavskyi. On pure complex spectrum for truncations of random orthogonal matrices and Kac polynomials. *arXiv preprint arXiv:1905.03154*, 2019.

[21] Advay Goel, Patrick Lopatto, and Xiaoyu Xie. Central limit theorem for the complex eigenvalues of Gaussian random matrices. *arXiv preprint arXiv:2306.10243*, 2023.

[22] Roger Horn and Charles Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

[23] Olav Kallenberg. *Random Measures, Theory and Applications*. Springer, 2017.

[24] Phil Kopel. Linear statistics of non-Hermitian matrices matching the real or complex Ginibre ensemble to four moments. *arXiv preprint arXiv:1510.02987*, 2015.

[25] T. Kriecherbauer, A.B.J. Kuijlaars, K.D. T.-R. McLaughlin, and P.D. Miller. Locating the zeros of partial sums of e^z with riemann–hilbert methods. In *Integrable Systems and Random Matrices: In Honor of Percy Deift: Conference on Integrable Systems, Random Matrices, and Applications in Honor of Percy Deift's 60th Birthday, May 22-26, 2006, Courant Institute of Mathematical Sciences, New York University, New York*, volume 458, page 183. American Mathematical Soc., 2008.

[26] Benjamin Landon, Patrick Lopatto, and Jake Marcinek. Comparison theorem for some extremal eigenvalue statistics. *The Annals of Probability*, 48(6):2894–2919, 2020.

[27] Patrick Lopatto and Kyle Luh. Tail bounds for gaps between eigenvalues of sparse random matrices. *Electronic Journal of Probability*, 26:1–26, 2021.

[28] Anthony Mays. A geometrical triumvirate of real random matrices. *arXiv preprint arXiv:1202.1218*, 2012.

[29] Hoi Nguyen, Terence Tao, and Van Vu. Random matrices: tail bounds for gaps between eigenvalues. *Probability Theory and Related Fields*, 167:777–816, 2017.

[30] Andrew M. Odlyzko. On the distribution of spacings between zeros of the zeta function. *Mathematics of Computation*, 48(177):273–308, 1987.

[31] Dai Shi and Yunjiang Jiang. Smallest gaps between eigenvalues of random matrices with complex Ginibre, Wishart and universal unitary ensembles. *arXiv preprint arXiv:1207.4240*, 2012.

[32] Christopher D. Sinclair. Correlation functions for $\beta = 1$ ensembles of matrices of odd size. *Journal of Statistical Physics*, 136(1):17–33, 2009.

[33] Hans-Jürgen Sommers and Waldemar Wieczorek. General eigenvalue correlations for the real Ginibre ensemble. *Journal of Physics A: Mathematical and Theoretical*, 41(40):405003, 2008.

[34] Alexander Soshnikov. Level spacings distribution for large random matrices: Gaussian fluctuations. *Annals of Mathematics*, pages 573–617, 1998.

[35] Alexander Soshnikov. Determinantal random point fields. *Russian Mathematical Surveys*, 55(5):923, 2000.

[36] Alexander Soshnikov. Statistics of extreme spacing in determinantal random point processes. *Moscow Mathematical Journal*, 5(3):705–719, 2005.

[37] Eric Thoma. Overcrowding and separation estimates for the coulomb gas. *Communications on Pure and Applied Mathematics*, 2022.

[38] Jade P. Vinson. *Closest Spacing of Eigenvalues*. PhD dissertation, Princeton University, 2001. arXiv:1111.2743.

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02906
Email address: `patrick_lopatto@brown.edu`

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02906
Email address: `matthew_meeker@brown.edu`