Acta Mathematicae Applicatae Sinica, English Series
Vol. 22, No. 2 (2006) 297–312

Acta Mathematicae Applicatae
Sinica, English Series

On the Spectral Distribution of Gaussian Random Matrices

B. Delyon, J. Yao

IRMAR, Universit de Rennes I, Campus de Beaulieu F-35042 Rennes Cedex, France
(E-mail: Jian-Feng.Yao@univ-rennes1.fr)

Abstract We consider the empirical spectral distribution (ESD) of a random matrix from the Gaussian Unitary Ensemble. Based on the Plancherel-Rotach approximation formula for Hermite polynomials, we prove that the expected empirical spectral distribution converges at the rate of $O(n^{-1})$ to the Wigner distribution function uniformly on every compact intervals $[u, v]$ within the limiting support $(-1, 1)$. Furthermore, the variance of the ESD for such an interval is proved to be $(\pi n)^{-2} \log n$ asymptotically which surprisingly enough, does not depend on the details (e.g. length or location) of the interval. This property allows us to determine completely the covariance function between the values of the ESD on two intervals.

Keywords Convergence rate, random matrices, empirical spectral distribution, Wigner distribution
2000 MR Subject Classification 60F15, 62H99

1 Introduction

The spectral analysis of large dimensional random matrices has been actively developed in the last few decades since the initial contributions of Wigner[18,19], see the monographs Mehta[12], Deift[6] or a recent review by Bai[2]. Various limiting distributions were discovered including the Wigner semicircular law, the Marčenko-Pastur law[11] and the circular law (see [1,8]). However, the convergence rates for these limits have not been fully identified yet. Johansson[10] has established a central limit theorem for linear integrals of the spectral distribution for a wide class of random matrices including the classical Gaussian Ensembles. As for the mean of the empirical spectral distribution function itself, successively improved rates have been proposed by Bai and his co-workers (see [2]).

In this paper, we consider a random matrix M_n from the so-called Gaussian Unitary Ensemble, that is a $n \times n$ Hermitian matrix whose off-diagonal elements are complex Gaussian with variance $\frac{1}{2}$ (independent real and imaginary parts with common variance $\frac{1}{4}$), and whose diagonal elements are real Gaussian with variance $\frac{1}{2}$. All the entries of M_n are independent.

Since the work of Wigner[20], it is well known that the joint distribution of the eigenvalues $x_1, \cdots, x_n$ of M_n has the following density w.r.t. the Lebesgue measure:

$$P(x_1, \cdots, x_n) = \exp\Big\{ -\sum_i x_i^2 + 2\sum_{i<j} \log|x_j - x_i| + C\Big\}$$

and that the distribution of two eigenvalues x, y of M_n drawn without replacement among

---

Manuscript received December 1, 2004. Revised September 2, 2005.

$x_1, \cdots, x_n$ is

$$s_n(x,y) = \frac{1}{n(n-1)}\big(K_n(x,x)K_n(y,y) - K_n(x,y)^2\big),$$
$$K_n(x,y) = \sum_{j=0}^{n-1} \varphi_j(x)\varphi_j(y) = \frac{\sqrt{n}}{\sqrt{2}(x-y)}\big(\varphi_n(x)\varphi_{n-1}(y) - \varphi_n(y)\varphi_{n-1}(x)\big).$$

Here φ_n is the n-th Hermite function

$$\varphi_n(x) = (2^n n!\sqrt{\pi})^{-1/2} e^{-x^2/2} H_n(x), \qquad n \geq 0,$$

where H_n is the the n-th Hermite polynomial

$$H_n(x) = e^{x^2}(-1)^n \frac{d^n}{dx^n}(e^{-x^2}).$$

If we denote by $\widetilde{f}_n(x)$ the function $f_n(x\sqrt{2n})$, the distribution of the normalized pair $(u,v) = (x,y)/\sqrt{2n}$ is thus

$$p_n(u,v) = 2n\widetilde{s}_n(u,v) = \frac{2}{n-1}(\widetilde{K}_n(u,u)\widetilde{K}_n(v,v) - \widetilde{K}_n(u,v)^2),$$
$$p_n(u) = \int p_n(u,v)dv = \sqrt{\frac{2}{n}}\widetilde{K}_n(u,u).$$

The last formula may be seen as a consequence of $\sqrt{2n}\int \widetilde{K}_n(u,v)^2 dv = \widetilde{K}_n(u,u)$ which follows from the orthonormality of the family $(\varphi_k)_{k\geq 0}$. We have also the following useful formula

$$p_n(u) = \big[\sqrt{2n}\varphi_n^2 - \sqrt{2(n+1)}\varphi_{n-1}\varphi_{n+1}\big](u\sqrt{2n}). \tag{1}$$

Wigner showed that $p_n(u)$ converges to the semi-circle law

$$\sigma(u) = \frac{2}{\pi}\sqrt{1-u^2}\,\mathbb{1}_{|u|\leq 1}.$$

We consider the empirical spectral measure

$$\sigma_n(dx) = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}(dx) \tag{2}$$

where $\lambda_1, ..\lambda_n$ are the eigenvalues of $M_n/\sqrt{2n}$ and δ_a the Dirac mass at a. For any function f, let

$$\sigma_n(f) = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i).$$

The *empirical spectral distribution function* F_n is simply $F_n(u) = \sigma_n(\mathbb{1}_{\lambda\leq u})$. As the behavior of F_n is rather different near the end points $\{-1,+1\}$ than in the interior of the interval, we will restrict our attention to the empirical process indexed by the set of close intervals contained in the (open) base interval $(-1,1)$, i.e. in the so-called "bulk" of the spectrum. More precisely, for $\overline{uv} = [u,v] \subset (-1,1)$, we define

$$F_n(\overline{uv}) = \sigma_n(\mathbb{1}_{\overline{uv}}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{u\leq\lambda_i\leq v} = F_n(v) - F_n(u), \tag{3}$$

the last identity being valid with probability 1.

In Section 3, we prove in Proposition 2 that the (interval-based) expected spectral distribution $\mathbb{E}F_n(\overline{uv})$ converges at the rate of $O(n^{-1})$ to $F(\overline{uv}) = F(v) - F(u)$, uniformly for all closed intervals in $(-1+\eta, 1-\eta)$ for some $\eta > 0$. This rate is interesting even in the examined Gaussian case. Indeed, in the case of general Wigner matrices with possibly non Gaussian entries, Bai et al.[3] established a rate of $\mathrm{O(n^{-1/2})}$ for

$$\|\mathbb{E}F_n - F\| := \sup_{x\in\mathbb{R}} |\mathbb{E}F_n(x) - F(x)|.$$

Although the exact rate for this sup-norm is still an open question, our result indicates that the rate should be $O(n^{-1})$ when restrcted to the bulk and the Gaussian case.

Our appoach relies on the celebrated Plancherel-Rotach's formula for asymptotic expansions of the Hermite polynomials. Indeed we need a higher order expansion than the usually refereed first-order expansion as given in Szegö[13p.199.], Consequently in Section 2 we first recall a second order exapnsion formula from the original paper of Plancherel and Rotach[14]. Next we establish an asymptotic expansion of the kernel $\widetilde{K}_n(u,v)$ (Proposition 1).

Based on this expansion for $\widetilde{K}_n$ we study the random fluctuations around the mean in Sections 4 and 5 and give the main results of the paper. Note that among many results established in Johansson[10], the author proved in particular that if f is a polynomial, the central limit theorem holds for $n\sigma_n(f)$. In particular

$$\lim_n \mathrm{Var}\big(n(\sigma_n(f) - \sigma(f))\big) = c_f, \tag{4}$$

for some positive constant c_f depending on f. In Section 4 of this paper, we prove for the (discontinuous) step function $f = \mathbb{1}_{\overline{uv}}$,

$$\mathrm{Var}\big(n(\sigma_n(f) - \sigma(f))\big) = \mathrm{Var}\big(n(F_n(\overline{uv}) - F(\overline{uv}))\big) = \pi^{-2}\log n + O(1).$$

This contrasts with the polynomial case. A surprising fact is that the dominating term $\pi^{-2}\log n$ is constant for *all* intervals $\overline{uv} \subset (-1,1)$. The result also confirms a widely-spread claim that any small neighborhood of the discontinuity points of f brings up an overwhelming contribution to the fluctuations (asymptotically).

Finally in Section 5, we identify the covariance function of the empirical process $\{F_n(\overline{uv})\}$. The covariance between $\{F_n(\overline{ab})\}$ and $\{F_n(\overline{cd})\}$ for two subintervals $\overline{ab}$ and $\overline{cd}$ of $(-1,1)$ has a simple asymptotic form. Only three values are possibles, namely $\pm\frac{1}{2}(\pi n)^{-2}\log n + O(n^{-2})$ or $O(n^{-2})$, depending on whether one of the intervals is included in the other. In particular, the details of these intervals such as their lengths or their location in $(-1,1)$ have no influence on this form. This asymptotic covariance function is very similar to the one found by Wieand[17] for random unitary matrices where a central limit theorem was also established. Actually we conjecture that such a CLT should also take place here.

We should also mention that a functional CLT for analytical linear spectral statistics of Wigner matrices with general entries is proposed in Bai and Yao[4]. In another related work of Diaconis and Evans[7], a multivariate CLT is proved for linear spectral statistics of random matrices from the complex unitary group. These authors have also recovered the CLT of Wieand[17] by a different approach based on Fourier analysis.

2 Asymptotic Expansion of the Kernel $\widetilde{K}_n$

2.1 The Plancherel-Rotach Formula

Throughout the paper we set for any $\psi \in (0,\pi)$,

$$a_\psi = \sin 2\psi - 2\psi,$$

$$b_\psi = \psi - \frac{1}{2}\sin 2\psi = -\frac{1}{2}a_\psi,$$
$$M_n(\psi) = \sin\Big(nb_\psi + \frac{\pi}{4} - \frac{\psi}{2}\Big).$$

Plancherel-Rotach Formula. *Let $\psi \in (0,\pi)$ and $x = 2\sqrt{n}\cos\psi$. Then for the Hermite polynomial*

$$\widetilde{H}_n(x) = (-1)^n e^{\frac{1}{2}x^2}\frac{d^n}{dx^n}\big(e^{-\frac{1}{2}x^2}\big)$$

we have

$$\frac{\widetilde{H}_{n-1}(x)}{(n-1)!} = \frac{e^{n(\frac{1}{2}+\cos^2\psi)}}{n^{n/2}\sqrt{\pi\sin\psi}}A_n(\psi), \tag{5}$$

where

$$A_n(\psi) = M_n(\psi) + \frac{1}{n}N_n(\psi) + \mathrm{O}([\mathrm{n}\sin^3\psi]^{-2}),$$
$$N_n(\psi) = \frac{3}{16}[\sin\psi]^{-2}\sin\Big(nb_\psi - \frac{3}{4}\pi - \frac{5}{2}\psi\Big) + \frac{5}{96}[\sin\psi]^{-3}\sin\Big(nb_\psi - \frac{5}{4}\pi - \frac{7}{2}\psi\Big).$$

This formula is exactly Formula (7) in [14] written for the case $k = 4$. Note that the form of $\widetilde{H}_n(x)$ used by Plancherel and Rotach is a scaled version of the usual Hermite polynomial $H_n(x)$ (we have indeed $\widetilde{H}_n(x) = (\sqrt{2})^{-n}H_n(x/\sqrt{2})$). Moreover, we slightly modify the formula by first following the Stirling's formula

$$n! = \Big(\frac{n}{e}\Big)^n\sqrt{2\pi n}\Big[1 + \frac{1}{12n} + O(n^{-2})\Big]$$

to get

$$e^{-\frac{1}{4}x^2}\widetilde{H}_{n-1}(x) = (\pi\sin\psi)^{-\frac{1}{2}}n^{-1}(n!)^{\frac{1}{2}}(2\pi n)^{\frac{1}{4}}\Big[1 + \frac{1}{24n} + O(n^{-2})\Big]A_n(\psi).$$

Next by substituting n for $n-1$, we have *for $x = 2\sqrt{n+1}\cos\psi$,*

$$e^{-\frac{1}{4}x^2}\widetilde{H}_n(x) = (\sin\psi)^{-\frac{1}{2}}(n!)^{\frac{1}{2}}2^{\frac{1}{4}}(\pi n)^{-\frac{1}{4}}\Big[1 - \frac{5}{24n} + O(n^{-2})\Big]A_{n+1}(\psi).$$

Coming back to H_n and with $y = \sqrt{2(n+1)}\cos\psi$, we have

$$\begin{aligned} e^{-\frac{1}{2}y^2}H_n(y) =& (\sin\psi)^{-\frac{1}{2}}(n!)^{\frac{1}{2}}2^{\frac{n}{2}+\frac{1}{4}}(\pi n)^{-\frac{1}{4}}\Big[1 - \frac{5}{24n} + O(n^{-2})\Big]A_{n+1}(\psi) \\ =& (\sin\psi)^{-\frac{1}{2}}(n!)^{\frac{1}{2}}2^{\frac{n}{2}+\frac{1}{4}}(\pi n)^{-\frac{1}{4}}\Big[M_{n+1}(\psi) + \frac{1}{n}C_n(\psi) + O(n^{-2})\Big], \end{aligned}$$

with

$$C_n(\psi) = -\frac{5}{24}M_{n+1}(\psi) + N_{n+1}(\psi)$$

and where the $O(n^{-2})$ term is uniform for all $\psi \in (\varepsilon, \pi-\varepsilon)$ for some $\varepsilon > 0$.

Finally we get the following second-order Plancherel-Rotach formula for the Hermite function $\varphi_n(y)$: for $y = \sqrt{2(n+1)}\cos\psi$

$$(2n)^{1/4}\varphi_n(y) = \sqrt{2}(\pi\sin\psi)^{-1/2}\Big\{M_{n+1}(\psi) + \frac{1}{n}C_n(\psi) + O(n^{-2})\Big\}, \tag{6}$$

where the error term is uniform on $(\varepsilon, \pi-\varepsilon)$.

It is worth noticing that if we keep only the main term $M_{n+1}(\psi)$ in the above expansion, and by considering the parametrization $y = \sqrt{2n+1}\cos\phi$, we readily arrive at the first order formula given in [13, p.199], with $M_{n+1}(\psi) = \sin\{\frac{1}{4}(2n+1)(\sin 2\phi - 2\phi) + \frac{3}{4}\pi\}$.

2.2 The Expansion of $\widetilde{K}_n$

The following family of functions $\mathcal{F}_n$ will play an important role in the derivations below. Fix a small $\varepsilon > 0$. A real function $\phi(\omega)$ defined on $[\varepsilon, \pi - \varepsilon]$ belongs to $\mathcal{F}_n$ if it is a (finite) linear combination of terms of type

$$P(\cos\omega, \sin\omega, \cos\omega/\sin\omega)\sin(nb_\omega + \alpha\pi + \beta)$$

or

$$Q(\cos\omega, \sin\omega, \cos\omega/\sin\omega)\sin\{(n+1)b_\omega + \alpha\pi + \beta\}$$

where P, Q are polynomials and α, β some constants. The functions M_n, N_{n+1} and C_n introduced in the previous section all belong to $\mathcal{F}_n$. Such a function ϕ is infinitely differentiable and $|\phi'(\omega)| = O(n)$.

Proposition 1. *Let $\eta \in (0,1]$ be fixed. For all $(u,v) = (\cos\omega, \cos\theta)$ such that $|u| \vee |v| \le 1 - \eta$, we have the following expansion*

$$\begin{aligned}
2\sqrt{2n}(u-v)\widetilde{K}_n(u,v) &= \frac{2}{\pi}(\sin\omega\sin\theta)^{-\frac{1}{2}} \\
&\times \Big\{ \sin\frac{\omega-\theta}{2}\cos\Big[\frac{n}{2}(a_\omega + a_\theta)\Big] + \sin\frac{\omega+\theta}{2}\sin\Big[\frac{n}{2}(a_\omega - a_\theta)\Big] \\
&+ \sum_{k=1}^{\infty}\frac{1}{n^k}\big[g_{n,k}(\omega,\theta) - g_{n,k}(\theta,\omega)\big]\Big\}
\end{aligned} \tag{7}$$

where the functions $g_{n,k}$ are linear combinations of terms

$$\phi_1(\omega)\cdots\phi_s(\omega)\psi_1(\theta)\cdots\psi_\ell(\theta)$$

with all ϕ_i and ψ_j's belonging to the family $\mathcal{F}_n$.

Proof. We will only prove the expansion up to the order $k = 1$ which is sufficient for the use in next sections.

In this proof we set $t_\omega = \cos\omega/\sin\omega$. Recall that

$$\begin{aligned}
\widetilde{K}_n(u,v) =& K_n(u\sqrt{2n}, v\sqrt{2n}) \\
=& \frac{1}{2}\frac{\varphi_n(u\sqrt{2n})\varphi_{n-1}(v\sqrt{2n}) - \varphi_{n-1}(u\sqrt{2n})\varphi_n(v\sqrt{2n})}{u-v}.
\end{aligned} \tag{8}$$

Clearly there exists $\varepsilon_0 > 0$ such that for large n, the angles ω_n and θ_n defined below all belong to the interval $[\varepsilon_0, \pi - \varepsilon_0]$:

$$\begin{aligned}
u\sqrt{2n} =& \sqrt{2n}\cos\omega = \sqrt{2(n+1)}\cos\omega_n, \\
v\sqrt{2n} =& \sqrt{2n}\cos\theta = \sqrt{2(n+1)}\cos\theta_n.
\end{aligned}$$

Application of the Plancherel-Rotach's formula in (6) gives

$$(2n)^{1/4}\varphi_n(u\sqrt{2n}) = \sqrt{2}(\pi\sin\omega_n)^{-1/2}\Big\{M_{n+1}(\omega_n) + \frac{1}{n}C_n(\omega_n) + O(n^{-2})\Big\}, \tag{9}$$

$$\begin{aligned}
[2(n-1)]^{1/4}\varphi_{n-1}(u\sqrt{2n}) &= \sqrt{2}(\pi\sin\omega)^{-1/2}\Big\{M_n(\omega) + \frac{1}{n-1}C_{n-1}(\omega) + O(n^{-2})\Big\} \\
&= \sqrt{2}(\pi\sin\omega)^{-1/2}\Big\{M_n(\omega) + \frac{1}{n}C_{n-1}(\omega) + O(n^{-2})\Big\}.
\end{aligned}$$

The main point is that we have to take care of the oscillation of the sinus function in M_n; otherwise the computation is standard.

First we want to substitute $\sin\omega$ for $\sin\omega_n$ in the r.h.s. of (9). As $\sqrt{2n}\cos\omega = \sqrt{2(n+1)}\cos\omega_n$, we have

$$\begin{aligned}\omega_n &= \omega + \frac{1}{2n}t_\omega - \frac{1}{8n^2}(t_\omega^3 + 3t_\omega) + O(n^{-3}),\\ (\sin\omega_n)^{-\frac{1}{2}} &= (\sin\omega)^{-\frac{1}{2}}\Big[1 - \frac{1}{4n}t_\omega^2 + O(n^{-2})\Big].\end{aligned} \tag{10}$$

Hence

$$(2n)^{1/4}\varphi_n(u\sqrt{2n}) = \sqrt{2}(\pi\sin\omega)^{-1/2}\Big\{M_{n+1}(\omega_n) + \frac{1}{n}D_n(\omega,\omega_n) + O(n^{-2})\Big\}, \tag{11}$$

where we have set

$$D_n(\omega,\omega_n) = -\frac{1}{4}t_\omega^2 M_{n+1}(\omega_n) + C_n(\omega_n).$$

Using the expansion (10) in M_{n+1} (see below for more details), and similarly in $C_n(\omega_n)$, it is clear that there is some function $G_n(\omega) \in \mathcal{F}_n$ such that

$$D_n(\omega,\omega_n) = G_n(\omega) + O(n^{-1}).$$

We thus have

$$(2n)^{1/4}\varphi_n(u\sqrt{2n}) = \sqrt{2}(\pi\sin\omega)^{-1/2}\Big\{M_{n+1}(\omega_n) + \frac{1}{n}G_n(\omega) + O(n^{-2})\Big\}.$$

The same expansions apply for $\varphi_n(v\sqrt{2n})$ and $\varphi_{n-1}(v\sqrt{2n})$, namely

$$\begin{aligned}(2n)^{1/4}\varphi_n(v\sqrt{2n}) &= \sqrt{2}(\pi\sin\theta)^{-1/2}\Big\{M_{n+1}(\theta_n) + \frac{1}{n}G_n(\theta) + O(n^{-2})\Big\},\\ [2(n-1)]^{1/4}\varphi_{n-1}(v\sqrt{2n}) &= \sqrt{2}(\pi\sin\theta)^{-1/2}\Big\{M_n(\theta) + \frac{1}{n}C_{n-1}(\theta) + O(n^{-2})\Big\}.\end{aligned}$$

Therefore by (8)

$$2(u-v)[4n(n-1)]^{1/4}\widetilde{K}_n(u,v) = \frac{2}{\pi}(\sin\omega\sin\theta)^{-1/2}\Big[\Delta_n + \frac{1}{n}\Gamma_n + O(n^{-2})\Big]$$

with

$$\Delta_n = M_{n+1}(\omega_n)M_n(\theta) - M_{n+1}(\theta_n)M_n(\omega),$$

and

$$\Gamma_n = [G_n(\omega)M_n(\theta) - G_n(\theta)M_n(\omega)] + [C_{n-1}(\theta)M_{n+1}(\omega) - C_{n-1}(\omega)M_{n+1}(\theta)]. \tag{12}$$

Note that the function Γ_n has the required form, namely a symmetric difference of some linear combination of products of functions from the family $\mathcal{F}_n$. Now we estimate precisely the term Δ_n. Recall that

$$\begin{aligned}M_{n+1}(\omega_n) &= \sin\Big\{(n+1)b_{\omega_n} + \frac{\pi}{4} - \frac{\omega_n}{2}\Big\}\\ &= \sin\Big\{(n+1)\Big(\omega_n - \frac{1}{2}\sin 2\omega_n\Big) + \frac{\pi}{4} - \frac{\omega_n}{2}\Big\}.\end{aligned}$$

By (10) we get

$$(n+1)b_{\omega_n} + \frac{1}{4}\pi - \frac{\omega_n}{2} = nb_\omega + \frac{\omega}{2} + \frac{\pi}{4} + \frac{1}{n}h_\omega + O(n^{-2}), \tag{13}$$

with

$$h_\omega = \frac{1}{4}\big(- [\sin\omega]^2 t_\omega^3 + [\sin 2\omega] t_\omega^2 - [\cos\omega]^2 t_\omega \big).$$

Therefore

$$M_{n+1}(\omega_n) = \sin\Big\{ nb_\omega + \frac{\omega}{2} + \frac{\pi}{4} \Big\} + \frac{1}{n} h_\omega \cos\Big\{ nb_\omega + \frac{\omega}{2} + \frac{\pi}{4} \Big\} + O(n^{-2}). \tag{14}$$

Recall that $M_n(\theta) = \sin(nb_\theta + \frac{1}{4}\pi - \frac{1}{2}\theta)$ we have then

$$\begin{aligned}\Delta_n =& \sin\Big\{ nb_\omega + \frac{\omega}{2} + \frac{\pi}{4} \Big\} \sin\Big\{ nb_\theta + \frac{\pi}{4} - \frac{\theta}{2} \Big\} \\ &- \sin\Big\{ nb_\theta + \frac{\theta}{2} + \frac{\pi}{4} \Big\} \sin\Big\{ nb_\omega + \frac{\pi}{4} - \frac{\omega}{2} \Big\} + \frac{1}{n} f_n(\omega,\theta) + O(n^{-2}),\end{aligned}$$

with

$$f_n(\omega,\theta) = h_\omega \cos\Big\{ nb_\omega + \frac{\omega}{2} + \frac{\pi}{4} \Big\} M_n(\theta) - h_\theta \cos\Big\{ nb_\theta + \frac{\theta}{2} + \frac{\pi}{4} \Big\} M_n(\omega). \tag{15}$$

On one hand the function f_n also has the required form of a symmetric difference. Moreover the first symmetric difference in (7) simply equals to

$$g_{n,1}(\omega,\theta) - g_{n,1}(\theta,\omega) = \Gamma_n + f_n.$$

On the other hand, by using

$$\sin(a-b)\sin(c+d) - \sin(c-d)\sin(a+b) = \sin(a+c)\sin(d-b) + \sin(a-c)\sin(b+d),$$

we obtain

$$\Delta_n = \cos[n(b_\theta + b_\omega)] \sin\frac{\omega-\theta}{2} + \sin[n(b_\theta - b_\omega)] \sin\frac{\omega+\theta}{2} + \frac{1}{n} f_n(\omega,\theta) + O(n^{-2}).$$

Taking into account that $b_\omega = -\frac{1}{2}a_\omega$, we finally obtain the first two terms in expansion (7). The proof is complete since substituting $[2n]^{1/2}$ for $[4n(n-1)]^{1/4}$ in the l.h.s. of this expansion does not affect the conclusion.

Corollary 1. *Let η, u and v as in Proposition 1. Then we have the following expansion*

$$\begin{aligned}&2\sqrt{2n}\widetilde{K}_n(u,v) = \pi^{-1}(\sin\omega\sin\theta)^{-\frac{1}{2}} \\ &\times \Big\{ - \frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} - \frac{\cos\frac{1}{2}n(a_\omega + a_\theta)}{\sin\frac{1}{2}(\omega+\theta)} + \sum_{k=0}^{\infty} \frac{1}{n^k} h_{n,k}(\omega,\theta) \Big\}\end{aligned}$$

where the functions $h_{n,k}$ satisfy $|h_{n,k}| \le d_k$ for some sequence of constants d_k. In particular,

$$\begin{aligned}&2\sqrt{2n}\widetilde{K}_n(u,v) = \pi^{-1}(\sin\omega\sin\theta)^{-\frac{1}{2}} \\ &\times \Big\{ - \frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} + \psi_n(\omega,\theta) + h_{n,0}(\omega,\theta) + O(n^{-1}) \Big\},\end{aligned} \tag{16}$$

with

$$\psi_n(\omega,\theta) = - \frac{\cos\frac{1}{2}n(a_\omega + a_\theta)}{\sin\frac{1}{2}(\omega+\theta)}. \tag{17}$$

Proof. Note that $u - v = \cos\omega - \cos\theta = -2\sin\frac{1}{2}(\omega+\theta)\sin\frac{1}{2}(\omega-\theta)$. From the expansion (7), we have

$$2\sqrt{2n}\widetilde{K}_n(u,v) = \pi^{-1}(\sin\omega\sin\theta)^{-\frac{1}{2}}$$
$$\Big\{ -\frac{\cos\frac{1}{2}n(a_\omega + a_\theta)}{\sin\frac{1}{2}(\omega+\theta)} - \frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} + \sum_{k=1}^{\infty}\frac{1}{n^{k-1}}h_{n,k-1}(\omega,\theta)\Big\}$$

where we have set for $k \geq 1$

$$h_{n,k-1}(\omega,\theta) := \frac{1}{n}\frac{g_{n,k}(\omega,\theta) - g_{n,k}(\theta,\omega)}{-2\sin\frac{1}{2}(\omega+\theta)\sin\frac{1}{2}(\omega-\theta)} \tag{18}$$

Let be $\varepsilon = \arccos(\eta)$, so that $\omega, \theta \in [\varepsilon, \pi-\varepsilon]$. The special form of $g_{n,k}$ implies that for some constant s_k we have

$$\Big|\frac{\partial}{\partial\omega}g_{n,k}(\omega,\theta)\Big| \vee \Big|\frac{\partial}{\partial\theta}g_{n,k}(\omega,\theta)\Big| \leq ns_k.$$

Hence

$$|h_{n,k-1}(\omega,\theta)| \leq \frac{2}{\sin\varepsilon}s_k\Big|\frac{\frac{1}{2}(\omega-\theta)}{\sin\frac{1}{2}(\omega-\theta)}\Big|.$$

Note that $\frac{1}{2}(\omega-\theta) \in [-\frac{1}{2}\pi+\varepsilon, \frac{1}{2}\pi-\varepsilon]$. As the function $x/\sin(x)$ is bounded on this interval, the boundedness of the $h_{n,k}$'s follows. Finally we obtain the expansion (16) by keeping only the first three terms.

3 Convergence Rate of the Expected Spectral Distribution Function

The so-called expected spectral distribution function is simply

$$\mathbb{E}F_n(u) = \int_{-\infty}^{u} p_n(s)ds$$

which converges to the Wigner distribution function $F(u)$. Similarly, for any interval $\overline{uv} = [u,v]$, the expected value $\mathbb{E}F_n(\overline{uv})$ converges to $F(\overline{uv}) = F(v) - F(u)$. Proposition 2 gives the rate of this convergence for the considered random matrices from Gaussian unitary ensemble. Note that similar results have been already established in [9], but our proof is completely different.

Proposition 2. *Let $\eta \in (0,1]$ be fixed. We have*

(1) $p_n(u) = \frac{2}{\pi}\sqrt{1-u^2} + O(n^{-1})$, *where the error term $O(n^{-1})$ is uniform for all $|u| \leq 1-\eta$.*

(2) $\sup_{\overline{uv}\subset[-1+\eta,1-\eta]} \big|\mathbb{E}F_n(\overline{uv}) - F(\overline{uv})\big| = O(n^{-1})$.

Proof. Let $u = \cos\omega$ and the angles ω_n and ω_n^+ be defined by the equalities

$$u\sqrt{2n} = \sqrt{2n}\cos\omega = \sqrt{2(n+1)}\cos\omega_n = \sqrt{2(n+2)}\cos\omega_n^+.$$

Let us from Eq.(1) that $p_n(u) = [\sqrt{2n}\varphi_n^2 - \sqrt{2(n+1)}\varphi_{n-1}\varphi_{n+1}](u\sqrt{2n})$. By the Plancherel-Rotach formula (6), we have

$$[2(n-1)]^{-\frac{1}{4}}\varphi_{n-1}(u\sqrt{2n}) = \sqrt{2}(\pi\sin\omega)^{-\frac{1}{2}}\big\{M_n(\omega) + O(n^{-1})\big\}.$$

As for $\varphi_n(u\sqrt{2n})$, by the computation in the proof of Proposition 1, namely Eqs.(11) and (14), we know that

$$[2n]^{-\frac{1}{4}}\varphi_n(u\sqrt{2n}) = \sqrt{2}(\pi\sin\omega)^{-\frac{1}{2}}\Big\{\sin\Big\{nb_\omega + \frac{\omega}{2} + \frac{\pi}{4}\Big\} + O(n^{-1})\Big\}.$$

For $\varphi_{n+1}(u\sqrt{2n})$, note that the computation between the pair (ω_n, ω_n^+) is exactly the same as for the one just recalled for the pair (ω, ω_n), except that we have to substitute $n+1$ for n. In other words,

$$[2(n+1)]^{-\frac{1}{4}}\varphi_{n+1}(u\sqrt{2n}) = \sqrt{2}(\pi \sin\omega_n)^{-\frac{1}{2}}\Big\{ \sin\Big\{(n+1)b_{\omega_n} + \frac{\omega_n}{2} + \frac{\pi}{4}\Big\} + O(n^{-1})\Big\}.$$

Recalling the expansion (13), we then find

$$(n+1)b_{\omega_n} + \frac{\omega_n}{2} + \frac{\pi}{4} = nb_\omega + \frac{3}{2}\omega + \frac{\pi}{4} + O(n^{-1}).$$

Collecting the three expansions we get

$$\begin{aligned} p_n(u) =& \frac{2}{\pi\sin\omega} \\ &\times \Big[\sin^2\Big\{nb_\omega + \frac{\omega}{2} + \frac{\pi}{4}\Big\} - \sin\Big\{nb_\omega - \frac{\omega}{2} + \frac{\pi}{4}\Big\}\sin\Big\{nb_\omega + \frac{3}{2}\omega + \frac{\pi}{4}\Big\} + O(n^{-1})\Big] \\ =& \frac{2}{\pi\sin\omega} \times \Big[\sin^2\omega + O(n^{-1})\Big] = \frac{2}{\pi}\sin\omega + O(n^{-1}). \end{aligned}$$

This proves the first assertion. The second assertion is then a straightforward consequence by integration.

4 Variance Function of the Empirical Process

We now study the variance function of the considered empirical process. The main result is the following

Theorem 1. *For all $\overline{uv} \subset (-1,1)$, the variance of $F_n(\overline{uv})$ has the following expansion*

$$Var(F_n(\overline{uv})) = \frac{1}{\pi^2}\frac{\log n}{n^2} + \mathrm{O}(\mathrm{n}^{-2}) \tag{19}$$

where the error term $\mathrm{O}(\mathrm{n}^{-2})$ is uniform for $\overline{uv} \subseteq [-1+\eta, 1-\eta]$ with any $\eta > 0$.

This theorem implies two important facts. First the variance of $F_n(\overline{uv})$ is on one hand much smaller than the "usual" $O(n^{-1})$ rate, and on the other hand much larger than the $O(n^{-2})$ order found for linear integrals $F_n(f)$ with f a polynomial (see Introduction). Moreover this variance is asymptotically independent of the details of the interval such as its length or its location in the base interval $(-1,1)$. This surprising fact will be the key for the computation of the covariance function of the process, see Section 5.

Proof Theorem 1. Let $Z_i = \mathbb{1}_{u\le\lambda_i\le v}$ for $1 \le i \le n$. These variables are (dependent) identically distributed Bernoulli variables with parameter $\alpha_n := \mathbb{E}F_n(\overline{uv}) = \mathbb{E}[F_n(v) - F_n(u)]$.

To get the covariance between two Z_i's, one has

$$\begin{aligned} \mathbb{E}(Z_1Z_2) =& \int_u^v\int_u^v p_n(s,t)dsdt\ , \\ =& \frac{n}{n-1}\int_u^v\int_u^v\Big[p_n(s)p_n(t) - \frac{2}{n}\widetilde{K}_n^2(s,t)\Big]dsdt\ , \\ =& \frac{n}{n-1}\Big[\alpha_n^2 - \frac{1}{n}\beta_n\Big] \end{aligned}$$

with

$$\beta_n = 2\int_u^v\int_u^v \widetilde{K}_n^2(s,t)dsdt. \tag{20}$$

It follows that

$$\mathrm{Cov}(Z_1, Z_2) = \frac{1}{n-1}[\alpha_n^2 - \beta_n],$$

and

$$n^2\ \mathrm{Var}(F_n(\overline{uv})) = n\ \mathrm{Var}(Z_1) + n(n-1)\ \mathrm{Cov}(Z_1, Z_2) = n(\alpha_n - \beta_n).$$

Let us set $\alpha = F(\overline{uv}) = F(v) - F(u)$. Since by Theorem 2, $n(\alpha_n - \alpha) = \mathrm{O}(1)$, it is sufficient to prove

$$\beta_n = \alpha - \frac{1}{\pi^2}\frac{\log n}{n} + O(n^{-1}). \tag{21}$$

Let $0 < \xi < \tau < \pi$ be defined by $u = \cos\tau,\ v = \cos\xi$. For any (small) $\eta > 0$, let $0 < \delta < \frac{\pi}{2}$ be defined by $1 - \eta = \cos\delta$. Thus $|u| \leq 1 - \eta$ if and only if $\tau \in [\delta, \pi - \delta]$.

Starting from Eq.(20), we have by the expansion (16) of $\widetilde{K}_n$

$$\begin{aligned}
4n\pi^2\beta_n =& \pi^2 \int_u^v \int_u^v 8n\widetilde{K}_n^2(s,t)dsdt \\
=& \int_u^v \int_u^v (\sin\omega \sin\theta)^{-1} \Big\{ \Big[\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} \Big]^2 \\
& - 2\psi_n(\omega,\theta)\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} - 2h_{n,0}(\omega,\theta)\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} \Big\} dsdt + \mathrm{O}(1) \\
=& \int_\xi^\tau \int_\xi^\tau \Big[k_n(\omega,\theta) + 2s_n(\omega,\theta) + 2t_n(\omega,\theta) \Big] d\theta d\omega + \mathrm{O}(1),
\end{aligned} \tag{22}$$

where

$$k_n(\omega,\theta) = \Big[\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)} \Big]^2, \tag{23}$$

$$s_n(\omega,\theta) = -\psi_n(\omega,\theta)\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)},$$

$$t_n(\omega,\theta) = -h_{n,0}(\omega,\theta)\frac{\sin\frac{1}{2}n(a_\omega - a_\theta)}{\sin\frac{1}{2}(\omega-\theta)}.$$

We will establish in next subsections that

$$A_1 := \int_\xi^\tau \int_\xi^\tau k_n(\omega,\theta)\ d\theta d\omega = 4n\pi^2\Big(\alpha - \frac{1}{\pi^2}\frac{\log n}{n}\Big) + \mathrm{O}(1), \tag{24}$$

$$A_2 := \int_\xi^\tau \int_\xi^\tau s_n(\omega,\theta) d\theta d\omega = \mathrm{O}(1), \tag{25}$$

$$A_3 := \int_\xi^\tau \int_\xi^\tau t_n(\omega,\theta)\ d\theta d\omega = \mathrm{O}(1). \tag{26}$$

The estimation (21) then follows and concludes the proof of the theorem.

4.1 Estimations of the Oscillating Integrals

We now establish the estimates (24)–(26). The functions to be integrated are all continuous and uniformly bounded on any off-diagonal region $\{|\omega-\theta|\geq\varepsilon\}$, but oscillate heavily near the diagonal.

As the computations of A_2 and A_3 are similar, we will only give full details for the integrals A_1 and A_2.

The integral $A_2 := \int\int s_n(\omega,\theta)$:

We start with the second integral. First note that

$$s_n(\omega,\theta)=\frac{\cos\frac{1}{2}n(a_\omega+a_\theta)\sin\frac{1}{2}n(a_\omega-a_\theta)}{\sin\frac{1}{2}(\omega+\theta)\sin\frac{1}{2}(\omega-\theta)}=\frac{\sin na_\omega-\sin na_\theta}{\cos\theta-\cos\omega}.$$

Let $D_\varepsilon=\{(\omega,\theta)\in[\xi,\tau]^2:|\theta-\omega|\geq\varepsilon\}$.

$$\begin{aligned}A_2=&\int_\xi^\tau\int_\xi^\tau s_n(\omega,\theta)\,d\omega d\theta=\lim_{\varepsilon\to0}\int\int_{D_\varepsilon}\frac{\sin na_\omega-\sin na_\theta}{\cos\theta-\cos\omega}d\omega d\theta\\=&2\lim_{\varepsilon\to0}\int\int_{D_\varepsilon}\frac{\sin na_\theta}{\cos\omega-\cos\theta}d\omega d\theta.\end{aligned}$$

Then

$$A_2^\varepsilon:=2\int\int_{D_\varepsilon}\frac{\sin na_\theta}{\cos\omega-\cos\theta}d\omega d\theta=2\int d\omega\int\frac{(\cos na_\theta-\cos na_\omega)'}{na'_\theta(\cos\theta-\cos\omega)}d\theta.$$

Here denotes derivative with respect to the variable θ. Therefore, integration by parts further yieds

$$\begin{aligned}A_2^\varepsilon=&2\int d\omega\frac{1}{n}\Big[\frac{1}{a'_\theta}\frac{\cos na_\theta-\cos na_\omega}{(\cos\theta-\cos\omega)}\Big]_{c_\omega}^{d_\omega}\\&-\frac{2}{n}\int\int_{D_\varepsilon}\frac{\cos na_\theta-\cos na_\omega}{(\cos\theta-\cos\omega)^2}\times\frac{a'_\theta\sin\theta-a''_\theta(\cos\theta-\cos\omega)}{{a'_\theta}^2}d\omega d\theta\\=:&\ A_{2a}-A_{2b},\end{aligned}\tag{27}$$

where the bounds c_ω and d_ω depend also on ε and belong to $[\xi,\tau]$.

Recall that $a'_\theta=-4\sin^2\theta$ is bounded away from zero on $[\delta,\pi-\delta]$. For the first term A_{2a}, note that for any continuously differentiable function $\psi(\omega,\theta)$ defined on $[\delta,\pi-\delta]^2$, we have

$$\Big|\frac{\cos n\psi(\theta)-\cos n\psi(\omega)}{\cos\theta-\cos\omega}\Big|\leq C_\psi n,\qquad(\omega,\theta)\in[\delta,\pi-\delta]^2,\tag{28}$$

for some constant C_ψ depending on the function ψ. Therefore the first term A_{2a} in (27) is uniformly bounded.

For the second term, let us set

$$G_\omega(\theta)=\frac{a'_\theta\sin\theta-a''_\theta(\cos\theta-\cos\omega)}{{a'_\theta}^2}.$$

By making use of the symmetry between ω and θ, we have

$$\begin{aligned}A_{2b}=&\frac{2}{n}\int\int_{D_\varepsilon}\frac{\cos na_\theta-\cos na_\omega}{(\cos\theta-\cos\omega)^2}G_\omega(\theta)\,d\omega d\theta\\=&\frac{1}{n}\int\int_{D_\varepsilon}\frac{\cos na_\theta-\cos na_\omega}{(\cos\theta-\cos\omega)^2}\big[G_\omega(\theta)-G_\theta(\omega)\big]\,d\omega d\theta\\=&\frac{1}{n}\int\int_{D_\varepsilon}\frac{\cos na_\theta-\cos na_\omega}{\cos\theta-\cos\omega}\,\frac{G_\omega(\theta)-G_\theta(\omega)}{\cos\theta-\cos\omega}d\omega d\theta.\end{aligned}$$

The first factor $(\cos na_\theta - \cos na_\omega)/(\cos\theta - \cos\omega)$ in the last integral is uniformly bounded by $C_1 n$ for some positive constant according to Eq.(28). The second factor $(G_\omega(\theta) - G_\theta(\omega))/(\cos\theta - \cos\omega)$ is a well defined continuous function on the compact domain $[\delta, \pi-\delta]^2$, hence also uniformly bounded. This proves that $A_{2b} = \mathrm{O}(1)$ uniformly, and hence by taking $\varepsilon \to 0$,

$$A_2 = \int_\xi^\tau \int_\xi^\tau s_n(\omega,\theta)\, d\omega d\theta = \mathrm{O}(1). \tag{29}$$

The integral $A_1 := \iint k_n(\omega,\theta)$: we have, by symmetry,

$$\begin{aligned} A_1 :=& \int_\xi^\tau \int_\xi^\tau k_n(\omega,\theta)\, d\omega d\theta = 2\int_\xi^\tau d\omega \int_\omega^\tau k_n(\omega,\theta) d\theta \\ =& 2\int_\xi^\tau d\omega \int_0^{\tau-\omega} k_n(\omega,\omega+x)dx \\ =& 2\int_\xi^\tau d\omega \int_0^{\tau-\omega} \Big[\frac{\sin\frac{1}{2}n(a_{\omega+x}-a_\omega)}{\sin\frac{1}{2}x}\Big]^2 dx. \end{aligned}$$

The following derivations largely depend on the family of functions g_ω indexed by $\omega \in [\delta, \pi-\delta]$ for some $\eta > 0$, which are defined as

$$g_\omega(x) = \frac{1}{2}[a_\omega - a_{\omega+x}] = x - \frac{1}{2}\big[\sin(2\omega+2x) - \sin(2\omega)\big], \qquad x \in [0,\pi].$$

It is easily checked that g_ω is a differentiable, strictly increasing function with derivative bounded away from 0 : $g'_\omega(x) \in [2\sin^2\delta, 2]$. Notice that $g_\omega(0) = 0$, $g_\omega(\pi) = \pi$ and $g'_\omega(0) = 2\sin^2\omega$. Its inverse function $h_\omega = g_\omega^{-1}$ is of similar nature : differentiable, strictly increasing with derivative bounded away from 0.

To abbreviate, let us set $b = b(\omega) = \tau - \omega$ and sometimes we denote g_ω simply by g and h_ω by h. We have

$$\begin{aligned} I_n(\omega) :=& \int_0^{b(\omega)} \Big[\frac{\sin ng_\omega(x)}{\sin\frac{1}{2}x}\Big]^2 dx = \frac{1}{n}\int_0^{ng(b)} \Big[\frac{\sin y}{\sin\frac{1}{2}h(\frac{y}{n})}\Big]^2 h'\Big(\frac{y}{n}\Big)dy \\ =& 4n\int_0^{ng(b)} \Big(\frac{\sin y}{y}\Big)^2 \cdot h'\Big(\frac{y}{n}\Big) \cdot \Big[\frac{\frac{1}{2}\frac{y}{n}}{\sin\frac{1}{2}h(\frac{y}{n})}\Big]^2 dy. \end{aligned}$$

Since $g(b) = g_\omega(b_\omega) > 0$, by the Lebesgue's dominated convergence theorem,

$$\frac{I_n(\omega)}{4n} \longrightarrow \int_0^\infty \Big(\frac{\sin y}{y}\Big)^2 \cdot \frac{1}{h'(0)}dy = \frac{\pi}{2h'(0)} = \pi\sin^2(\omega).$$

Again by that theorem,

$$\frac{A_1}{4n\pi^2} \longrightarrow \frac{2}{\pi}\int_\xi^\tau \sin^2(\omega)d\omega = \frac{2}{\pi}\int_u^v \sqrt{1-s^2}ds = \alpha.$$

Rate of convergence : Now we proceed to estimate the rate of the convergence above. We have

$$\frac{A_1}{4n\pi^2} - \alpha = \frac{2}{\pi^2}\int_\xi^\tau \Delta_n(\omega)d\omega, \tag{30}$$

where we have set

$$\begin{aligned}\Delta_n(\omega) :=&\frac{I_n(\omega)}{4n} - \pi\sin^2(\omega)\\ =&\int_0^{ng(b)} \Big(\frac{\sin y}{y}\Big)^2 \cdot \Big\{h'\Big(\frac{y}{n}\Big)\Big[\frac{\frac{1}{2}\frac{y}{n}}{\sin\frac{1}{2}h(\frac{y}{n})}\Big]^2 - \frac{1}{h'(0)}\Big\}dy - \int_{ng(b)}^{\infty}\Big(\frac{\sin y}{y}\Big)^2\frac{1}{h'(0)}dy\\ =:&\Delta_{n,1}(\omega) - \Delta_{n,2}(\omega).\end{aligned}$$

For the first term $\Delta_{n,1}$,

$$\Delta_{n,1} = \frac{1}{n}\int_0^{g(b)} t(\sin nz)^2 \cdot \Big\{z^{-2}\Big(h'(z)\Big[\frac{\frac{1}{2}z}{\sin\frac{1}{2}h(z)}\Big]^2 - \frac{1}{h'(0)}\Big)\Big\}dz.$$

By expanding $h(z)$ and $h'(z)$ near $z=0$, we have

$$h'(z)\Big[\frac{\frac{1}{2}z}{\sin\frac{1}{2}h(z)}\Big]^2 - \frac{1}{h'(0)} = \frac{1}{12}\,\frac{-3b^2+2ca+a^4}{a^3}z^2 + \mathrm{O}(\mathrm{z}^3)$$

where a, b, c are the first 3 derivatives of $h(z)$ at the origin. In particular, $\frac{1}{a} = 1/h'(0) = g'(0) = 2\sin^2(\omega)$ is bounded ; also the $\mathrm{O}(\mathrm{z}^3)$ term is uniform for all $\omega \in [\delta, \pi-\delta]$. Hence the above integral is bounded and

$$\Delta_{n,1}(\omega) = O(n^{-1}). \tag{31}$$

For the second term $\Delta_{n,2}(\omega)$, we get

$$\delta_n := \int_\xi^\tau \Delta_{n,2}(\omega)d\omega = \int_\xi^\tau \frac{1}{h'_\omega(0)}d\omega\int_{ng_\omega(b(\omega))}^{\infty}\Big(\frac{\sin y}{y}\Big)^2 dy.$$

Now let $\theta = \tau - \omega$, it turns out that

$$g_\omega(b(\omega)) = g_\omega(\theta) = \theta - \frac{1}{2}\sin 2\tau + \frac{1}{2}\sin(2\tau - 2\theta) = g_{\pi-\tau}(\theta)\ ,$$

and

$$\frac{1}{h'_\omega(0)} = g'_\omega(0) = 2\sin^2(\omega) = 2\sin^2(\tau-\theta) = 2\sin^2(\pi-\tau+\theta) = g'_{\pi-\tau}(\theta).$$

Hence

$$\delta_n = \int_0^{\tau-\xi} g'_{\pi-\tau}(\theta)d\theta\int_{ng_{\pi-\tau}(\theta)}^{\infty}\Big(\frac{\sin y}{y}\Big)^2 dy$$

Denote by $c = \tau - \xi > 0$ and $d = (\pi - \tau) \in [\delta, \pi-\delta]$.

$$\begin{aligned}\delta_n :=& \int_0^c g'_d(\theta)d\theta\int_{ng_d(\theta)}^{\infty}\Big(\frac{\sin y}{y}\Big)^2 dy\\ =&\int_0^\infty\Big(\frac{\sin y}{y}\Big)^2 dy\int_0^{c\wedge h_d(y/n)} g'_d(\theta)d\theta\\ =&\Big(\int_0^{ng_d(c)} dy + \int_{ng_d(c)}^{\infty} dy\Big)\Big(\frac{\sin y}{y}\Big)^2\Big[g_d(c)\wedge\frac{y}{n}\Big].\end{aligned} \tag{32}$$

The second part $\int_{ng_d(c)}^{\infty}(\cdot)$ is of order $\mathrm{O}(\mathrm{n}^{-1})$ since $g_d(c) > 0$ and for large positive X,

$$\int_X^\infty\Big(\frac{\sin y}{y}\Big)^2 dy = \mathrm{O}\Big(\frac{1}{\mathrm{X}}\Big).$$

For the first part,

$$\int_0^{ng_d(c)} \Big(\frac{\sin y}{y}\Big)^2 \frac{y}{n} dy = \frac{1}{n}\Big[\frac{1}{2}\log n + \mathrm{O}(1)\Big], \tag{33}$$

the last step being a consequence of the following classical expansion

$$\int_0^X \frac{\sin^2(y)}{y} dy = \frac{1}{2}\log X + \mathrm{O}(1)\ , \qquad \mathrm{X} \longrightarrow \infty.$$

Collecting results from Eqs.(30),(31),(33) gives

$$\frac{A_1}{4n\pi^2} - \alpha = \frac{2}{\pi^2}\Big[-\frac{1}{2n}\log n + O(n^{-1})\Big]$$

which is the desired result (24).

5 Covariance Function and a Conjecture

The aim of this section is to examine the covariance function of the process $\{F_n(\overline{uv})\}$ indexed by the set of closed intervals $\overline{uv} \subset (-1,1)$. Surprisingly, the asymptotic expansion of this function are easy to identify due to the particular form of the variance function established in Theorem 1.

Let

$$\Gamma_n(\overline{ab},\overline{cd}) := \ \mathrm{Cov}\big(F_n(\overline{ab}), F_n(\overline{cd})\big)\ , \qquad \overline{ab} \subset (-1,1), \quad \overline{cd} \subset (-1,1). \tag{34}$$

Theorem 2. *Without loss of generality, let us assume $b \le d$ for two closed intervals $\overline{ab}$, $\overline{cd}$ enclosed in $(-1,1)$. Then the covariance function Γ_n has the following expansions:*

		$\Gamma_n(\overline{ab},\overline{cd})$
1. $a < c < d$	*non inclusive intervals*	$-\frac{1}{2}\frac{\log n}{\pi^2 n^2} + O(n^{-2})$
2. $a = c < b < d$ *or*		
$c < a < b = d$	*inclusion with a common endpoint*	$+\frac{1}{2}\frac{\log n}{\pi^2 n^2} + O(n^{-2})$
3. $c < a < b < d$	*strict inclusion*	$O(n^{-2})$

All the $O(n^{-2})$ terms are uniform for intervals $\overline{ab}$, $\overline{cd}$ enclosed in $[-1+\eta, 1-\eta]$ for some $\eta > 0$.

Main information from this table are: first the covariance depend on whether or not one of the intervals is included in the other. Except Case 3 corresponding to a smaller covariance, they have the same asymptotic order than their variances; secondly, the asymptotic magnitude of the covariances do not depend explicitly on the size of the intervals, nor on their locations within the base interval $(-1,1)$.

Proof of Theorem 2 Case 1 consists in the following three sub-cases:

(1.a) neighboring intervals with $a < b = c < d$;
(1.b) disjoint intervals with $a < b < c < d$;
(1.c) overlapping intervals with $a < c < b < d$.

The basic case is (1.a) and conclusions for all others sub-cases as well as for Cases 2 and 3 follow from this special case. We thus give details in this special case and Case 3 only. In this proof $F_{\overline{ab}}$ stands for a shortcut of $F_n(\overline{ab})$ and let $\gamma_n = \frac{\log n}{\pi^2 n^2}$.

Case (1.a). We have $F_{\overline{ad}} = F_{\overline{ab}} + F_{\overline{cd}}$. Then

$$\mathrm{Var}(F_{\overline{ad}}) = \mathrm{Var}(F_{\overline{ab}}) + \mathrm{Var}(F_{\overline{cd}}) + 2\Gamma_n(F_{\overline{ab}}, F_{\overline{cd}}).$$

By Theorem 1, all the three variances are of order $\gamma_n + O(n^{-2})$. Thus the result follows.

Case 3. Here we have $c < a < b < d$. Thus $F_{\overline{cd}} = F_{\overline{ca}} + F_{\overline{ab}} + F_{\overline{bd}}$ and by using Case (1.a) and Theorem 1

$$\begin{aligned}\mathrm{Cov}(F_{\overline{ab}}, F_{\overline{cd}}) &= \mathrm{Cov}(F_{\overline{ab}}, F_{\overline{ca}}) + \mathrm{Cov}(F_{\overline{ab}}, F_{\overline{ab}}) + \mathrm{Cov}(F_{\overline{ab}}, F_{\overline{bd}}) \\ &= -\frac{1}{2}\gamma_n + \gamma_n - \frac{1}{2}\gamma_n + O(n^{-2}) = O(n^{-2}).\end{aligned}$$

Note that this asymptotic covariance function is very similar to the one found by Wieand[17] for random unitary matrices where this author also established a central limit theorem. Actually we conjecture that such a CLT should also take place here. Namely, given m subintervals $\overline{u_1v_1}, \cdots, \overline{u_mv_m}$ of $(-1,1)$, the random vector

$$\Big[\frac{F_n(\overline{u_jv_j}) - F(\overline{u_jv_j})}{\sqrt{\log n}/(\pi n)}, \quad j = 1, \cdots, m\Big]$$

converge weakly to a zero-mean Gaussian vector $(Z_1, \cdots, Z_m)$ with $\mathrm{Var}(Z_i) = 1$ and the covariances $\mathrm{Cov}(Z_i, Z_j)$ $(i \neq j)$ taking one of the three values, $\pm\frac{1}{2}$ or 0.

It is also worth noticing that from (7), we readily see that $\widetilde{K}_n$, once properly scaled, has a limit called the sine kernel. Random point field associated to this limiting kernel is then a special instance of random determinantal fields introduced by Soshnikov (see [16] and [15]). Indeed following Costin and Lebowitz[5], this author proved a central limit theorem for the empirical counts of such a field on an interval $[-L, L]$ when L goes to infinity.

References

[1] Bai, Z.D. Circular law. *Ann. Probab.*, 25: 494–529 (1997)

[2] Bai, Z.D. Methodologies in spectral analysis of large dimensional random matrices. A review. *Statistica Sinica*, 9: 611–677 (1998)

[3] Bai, Z.D., Miao, B.Q., Tsay, Jhishen. Convergence rate of the spectral distribution of large Wigner matrices. *International Mathematical J.*, 1: 65–90 (2001)

[4] Bai, Z.D., Yao, J.F. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli*, 11(6): 1059–1092 (2005)

[5] Costin, O., Lebowitz, J. Gaussian fluctuations in random matrices. *Phys. Rev. lett.*, 75: 69–72 (1995)

[6] Deift, P.A. Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. New York Univ., Courant Inst. Math. Sci., New York, 1999

[7] Diaconis, P., Evans, S.N. Linear functionals of eigenvalues of random matrices *Trans. Amer. Math. Soc.*, 353(7): 2615–2633 (2001)

[8] Girko, V.L. The circle law. *Theory Probab. Appl.*, 29: 694–706 (1984)

[9] Inoue, A., Nomura, Y. Some refinements of Wigner's semi-circle law for Gaussian random matrices using superanalyis. *Asymptotic Analysis*, 23: 329–375 (2000)

[10] Johansson, K. On fluctuations of eigenvalues of random hermitian matrices. *Duke Mathematical J.*, 91: 151–204 (1998)

[11] Marčenko, V.A., Pastur, L.A. Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sb*, 1: 457–483 (1967)

[12] Mehta, M.L. Random Matrices (second edition). Academic Press, San Diego, 1991

[13] Szegö G. Orthogonal polynomials. Colloquium Publ. 23, American Mathematical Society, New York, 1959

[14] Plancherel, M., Rotach W. Sur les valeurs asymtotiques des polynmes d'Hermite. *Comm. Math. Helv.*, 1: 227–254 (1929)

[15] Soshnikov, A. Gaussian limit for determinantal random point fields. *Ann. Probab.*, 30: 171–187 (2002)

[16] Soshnikov, A. Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields. *J. Statist. Phys.*, 100: 491–522 (2000)

[17] Wieand, K. Eigenvalue distributions of random unitary matrices. *Probab. Theory Related Fields*, 123: 202–224 (2002)

[18] Wigner, E.P. Characteristic vectors bordered matrices with infinite dimensions. *Ann. of Math.*, 62: 548–564 (1955)

[19] Wigner, E.P. On the distributions of the roots of certain symmetric matrices. *Ann. Math.*, 67: 325–327 (1958)

[20] Wigner E.P. Distribution laws for the roots of a random Hermitian matrix. In : *Statistical theory of spectra: Fluctuations*, C.E. Porter ed. Academic Press, ???? 1965, 446–461