Bernoulli **11**(6), 2005, 1059–1092

On the convergence of the spectral empirical process of Wigner matrices

Z.D. BAI[1,2] and J. YAO[3]

[1]*Department of Mathematics, Northern East Normal University, 5268 People's Road, 130024 Changchun, China*
[2]*Department of Statistics and Applied Probability, National University of Singapore, 10, Kent Ridge Crescent, Singapore 119260*
[3]*IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France.*
E-mail: jian-feng.yao@univ.rennes1.fr

It is well known that the spectral distribution F_n of a Wigner matrix converges to Wigner's semicircle law. We consider the empirical process indexed by a set of functions analytic on an open domain of the complex plane including the support of the semicircle law. Under fourth-moment conditions, we prove that this empirical process converges to a Gaussian process. Explicit formulae for the mean function and the covariance function of the limit process are provided.

Keywords: central limit theorem; linear spectral statistics; random matrix; spectral distribution; Wigner matrices

1. Introduction and main results

A *complex Wigner matrix* of size n is a Hermitian matrix $W_n = (x_{ij})_{1\leqslant i,j\leqslant n}$ where the upper-triangle entries $(x_{ij})_{1\leqslant i\leqslant j\leqslant n}$ are independent, zero-mean complex-valued random variables satisfying the following moment conditions:

(i) for all i, $\mathbb{E}|x_{ii}|^2 = \sigma^2 > 0$;
(ii) for all $i < j$, $\mathbb{E}|x_{ij}|^2 = 1$ and $\mathbb{E}x_{ij}^2 = 0$.

The set of these complex Wigner matrices is called the *unitary ensemble* (UE). Similarly, a *real Wigner matrix* of size n is a real symmetric matrix W_n where the upper-triangle entries $(x_{ij})_{1\leqslant i\leqslant j\leqslant n}$ are independent, zero-mean real-valued random variables satisfying the following moment conditions:

(i) for all i, $\mathbb{E}|x_{ii}|^2 = \sigma^2 > 0$;
(ii) for all $i < j$, $\mathbb{E}|x_{ij}|^2 = 1$.

The set of these real Wigner matrices is called the *orthogonal ensemble* (OE). In both cases, the entries are not necessarily identically distributed. If, in addition, the entries are Gaussian (with $\sigma^2 = 1$ and 2 for the UE and OE, respectively), the above ensembles are the classical Gaussian unitary ensemble (GUE) and Gaussian orthogonal ensemble (GOE) of random matrices.

1350–7265

The empirical spectral distribution F_n is the empirical distribution generated by the n eigenvalues of the normalized matrix $n^{-1/2}W_n$. This distribution is supported by the real line. Wigner (1955, 1958) first proved that as $n \to \infty$, $\mathbb{E}F_n$ converges to the semicircle law whose density function is given by

$$F(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x, \qquad x \in [-2, 2].$$

It was later established that F_n converges to F also in probability and almost surely. A review can be found in Bai (1999).

The problem of the convergence rate has been considered more recently, and several results are proposed in Bai (1993), Costin and Lebowitz (1995), Johansson (1998), Khorunzhy *et al.* (1996), Sinai and Soshnikov (1998) and Bai *et al.* (2002). However, the exact convergence rate remains unknown for Wigner matrices. Results from numerical simulations lead to a 'folklore conjecture' of a rate of the order of $O(1/n)$.

It thus seems natural to consider the asymptotics of the empirical process $G_n(x) = n(F_n(x) - F(x))$. However, there is plenty of evidence to show that the process G_n cannot converge in any metric space. Thus, we have to draw back a little and consider the linear functionals of the process $G_n(x)$.

More precisely, let $\mu(f)$ denote the integral of a function f with respect to a signed measure μ. An open set $\mathcal{U}$ of the complex plane including the interval $[-2, 2]$, the support of F, will be fixed throughout this paper. Next, define $\mathcal{A}$ to be the set of analytic functions $f : \mathcal{U} \to \mathbb{C}$. We then consider the empirical process $G_n := \{G_n(f)\}$ indexed by $\mathcal{A}$, that is,

$$G_n(f) := n\int_{-\infty}^{\infty} f(x)[F_n - F](\mathrm{d}x), \qquad f \in \mathcal{A}. \tag{1.1}$$

To study the weak limit of G_n, we need further conditions on the moments:

Condition 1.1 *Homogeneity of fourth moments.* $M = \mathbb{E}|x_{ij}|^4$ *for* $i \neq j$;

Condition 1.2 *Uniform tails. For any* $\eta > 0$, *as* $n \to \infty$,

$$\frac{1}{\eta^4 n^2}\sum_{i,j} \mathbb{E}\left[|x_{ij}|^4 \mathbb{1}_{\{|x_{ij}| \geqslant \eta\sqrt{n}\}}\right] = o(1).$$

Note that Condition 1.2 implies the existence of a sequence $\eta_n \downarrow 0$ such that

$$(\eta_n\sqrt{n})^{-4}\sum_{i,j} \mathbb{E}[|x_{ij}|^4 \mathbb{1}_{\{|x_{ij}| \geqslant \eta_n\sqrt{n}\}}] = o(1). \tag{1.2}$$

The main result of this paper is that the empirical process G_n converges to a Gaussian process. As a consequence, for any p elements $f_1, \ldots, f_p$ of $\mathcal{A}$, the finite-dimensional central limit theorem (CLT) holds, that is, the vector $[G_n(f_1), \ldots, G_n(f_p)]$ converges weakly to a p-dimensional Gaussian distribution.

Let $\{T_k\}$ be the family of Chebyshev polynomials and define, for $f \in \mathcal{A}$ and any integer $\ell \geqslant 0$,

$$\tau_\ell(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos(\theta))\mathrm{e}^{\mathrm{i}\ell\theta}\,\mathrm{d}\theta$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos(\theta))\cos(\ell\theta)\mathrm{d}\theta = \frac{1}{\pi}\int_{-1}^{1} f(2t)T_\ell(t)\frac{1}{\sqrt{1-t^2}}\mathrm{d}t. \tag{1.3}$$

In order to give a unified statement for both ensembles, we introduce the parameter κ with values 1 and 2 for the complex and real Wigner ensemble, respectively. Moreover, set $\beta = \mathbb{E}(|x_{12}|^2 - 1)^2 - \kappa$. In particular, for the GUE we have $\kappa = \sigma^2 = 1$ and for the GOE we have $\kappa = \sigma^2 = 2$, and in both cases $\beta = 0$.

Theorem 1.1. *Under Conditions 1.1 and 1.2, the spectral empirical process $G_n = (G_n(f))$ indexed by the set of analytic functions $\mathcal{A}$ converges weakly to a Gaussian process $G := \{G(f) : f \in \mathcal{A}\}$ with mean function*

$$\mathbb{E}[G(f)] = \frac{\kappa-1}{4}\{f(2)+f(-2)\} - \frac{\kappa-1}{2}\tau_0(f) + (\sigma^2-\kappa)\tau_2(f) + \beta\tau_4(f), \tag{1.4}$$

and covariance function $c(f, g) := \mathbb{E}[\{G(f) - \mathbb{E}G(f)\}\{G(g) - \mathbb{E}G(g)\}]$ given by

$$c(f, g) = (\sigma^2-\kappa)\tau_1(f)\tau_1(g) + 2\beta\tau_2(f)\tau_2(g) + \kappa\sum_{\ell=1}^{\infty}\ell\tau_\ell(f)\tau_\ell(g) \tag{1.5}$$
$$= \frac{1}{4\pi^2}\int_{-2}^{2}\int_{-2}^{2} f'(t)g'(s)V(t, s)\mathrm{d}t\,\mathrm{d}s, \tag{1.6}$$

where

$$V(t, s) = \left(\sigma^2 - \kappa + \frac{1}{2}\beta ts\right)(4-t^2)^{1/2}(4-s^2)^{1/2} + \kappa\log\left(\frac{4-ts+(4-t^2)^{1/2}(4-s^2)^{1/2}}{4-ts-(4-t^2)^{1/2}(4-s^2)^{1/2}}\right). \tag{1.7}$$

Note that our definition implies that the variance of $G(f)$ equals $c(f, \bar{f})$. Let $\delta_a(\mathrm{d}t)$ be the Dirac measure at a point a. The mean function can also be written as

$$\mathbb{E}[G(f)] = \int_{\mathbb{R}} f(2t)\mathrm{d}\nu(t), \tag{1.8}$$

with signed measure

$$\mathrm{d}\nu(t) = \frac{\kappa-1}{4}[\delta_1(\mathrm{d}t) + \delta_{-1}(\mathrm{d}t)]$$
$$+ \frac{1}{\pi}\left[-\frac{\kappa-1}{2} + (\sigma^2-\kappa)T_2(t) + \beta T_4(t)\right]\frac{1}{\sqrt{1-t^2}}\mathbb{1}_{[-1,1]}(t)\mathrm{d}t. \tag{1.9}$$

In the case of the GUE and GOE, the covariance reduces to the third term in (1.5). The mean $\mathbb{E}[G(f)]$ is thus always zero for the GUE, and for the GOE we have

$$\mathbb{E}[G(f)] = \frac{1}{4}\{f(2) + f(-2)\} - \frac{1}{2}\tau_0(f).$$

Therefore the limit process is not necessarily centred.

Gaussian fluctuations in random matrices are considered by various authors, starting with Costin and Lebowitz (1995). Johansson (1998) considers an extended random ensemble whose entries follow a density proportional to $\exp[-n\mathrm{tr}\{V(W)\}]$, where V is a polynomial of even degree with positive leading coefficient. He established a CLT for the linear spectral statistics. Therefore the Gaussian ensembles are special cases of both Johansson (1998) and the present paper. For these ensembles, our theorem applied to polynomials coincides with the results of Johansson; see Section 6 for a detailed comparison.

Another related work is Khorunzhy *et al.* (1996), where the authors consider the orthogonal ensemble (with general entries) and established a CLT for $n\{s_n(z_1) - \mathbb{E}s_n(z_1), \ldots, s_n(z_q) - \mathbb{E}s_n(z_q)\}$, where q is an arbitrary integer, z_j are complex numbers such that $|\Im(z_j)| \geqslant 2$ and s_n is the Stieltjes transform (or resolvent) $s_n(z) = \frac{1}{n}\mathrm{tr}(W_n/\sqrt{n} - zI)^{-1}$. This CLT is very close to Proposition 4.1 below, corresponding to the finite-dimensional convergence part of our Theorem 2.1. Note that Proposition 4.1 is applicable without the restriction $|\Im(z_j)| \geqslant 2$ so that the points z_j can approach the real axis in a well-controlled manner. This improvement is fundamental for the contour integration used for the derivation of the main Theorem 1.1 from Theorem 2.1. It is also worth noticing that in many applications the functional CLT given in Theorem 1.1 is more useful than the (finite-dimensional) CLT for the resolvent given in Proposition 4.1 (or in Khorunzhy *et al.* 1996).

Consider, for example, a bivariate function $f(x, t)$ and the stochastic process

$$Z_n(t) = \sum_{k=1}^{n} f(\lambda_k, t) - n\int_{-2}^{2} f(x, t)F(\mathrm{d}x).$$

If both f and $\partial f(x, t)/\partial t$ are analytic in x over a region containing $[-2, 2]$, it follows easily from Theorem 1.1 that $Z_n(t)$ converges to a Gaussian process. Its finite-dimensional convergence is exactly the same as in Theorem 1.1, while its tightness can be obtained as a simple consequence of the same theorem. However, such a result does not follow from the results of Khorunzhy *et al.* (1996). Processes like $Z_n(t)$ are of undoubted importance for applications of the random matrix theory.

A third related work is Sinai and Soshnikov (1998). Assume $p = p(n) \to \infty$ and $p/\sqrt{n} \to 0$. These authors establish a CLT for $\mathrm{tr}(2^{-1}n^{-1/2}W_n)^p - \mathbb{E}\mathrm{tr}(2^{-1}n^{-1/2}W_n)^p$ under the assumptions that the underlying variables are symmetric and have all moments satisfying an appropriate growth condition. They also prove that for f analytic on the disc of radius 2, the centred random variable $G_n(f) - \mathbb{E}[G_n(f)]$ has a Gaussian limit. However, neither the mean function nor variance function of the Gaussian limit is provided.

This paper is organized as follows. In Section 2 we give the main steps necessary to prove Theorem 1.1. In particular, Theorem 2.1 will be introduced as an intermediate result. We then introduce the truncation tool leading to a preliminary simplification of the proofs. Some useful inequalities and a standard formula are also presented. Sections 3 and 4 are devoted primarily to the proof of Theorem 2.1. Also a much shorter proof of Theorem 1.1 using Theorem 2.1 and a proof of (2.3) are given at different places in these two sections.

In Section 5 we derive the mean and covariance functions given in Theorem 1.1. Then in Section 6 we present applications to linear spectral statistics and to the Gaussian ensembles. Finally, Section 7 presents two frequently used lemmas.

2. Strategy of the proof, simplifications and known formula

2.1. Strategy of the proof

Let γ be the contour formed by the boundary of the rectangle with vertex $(\pm a \pm \mathrm{i}v_0)$ where $a > 2$ and $1 \geqslant v_0 > 0$. We can always assume that $\gamma \subset \mathcal{U}$ with sufficiently small (but fixed) a and v_0. Then, for every $x \in (-a, a)$, by Cauchy's theorem,

$$f(x) = \frac{1}{2\pi\mathrm{i}} \oint_\gamma \frac{f(z)}{z - x} \mathrm{d}z.$$

Recall that the Stieltjes transform $s_H(z)$ of any function H of bounded variation on $\mathbb{R}$ is defined by

$$s_H(z) = \int_{-\infty}^{\infty} \frac{\mathrm{d}H(x)}{x - z}, \qquad z \in \mathbb{C}^+ := \{u + \mathrm{i}v,\, v > 0\}.$$

This definition applies to a probability distribution function. Also, we may analytically extend the Stieltjes transform to the whole complex plane, except for the support of H.

Let $s_n(z)$ and $s(z)$ denote the extended Stieltjes transforms of F_n and F, respectively. Then

$$\begin{aligned} G_n(f) &= \int f(x) n[F_n - F](\mathrm{d}x) \\ &= \frac{1}{2\pi\mathrm{i}} \int \oint_\gamma \frac{f(z)}{z - x} n[F_n - F](\mathrm{d}x)\mathrm{d}z = \frac{1}{2\pi\mathrm{i}} \oint_\gamma f(z)\mathrm{d}z \int \frac{1}{z - x} n[F_n - F](\mathrm{d}x) \\ &= -\frac{1}{2\pi\mathrm{i}} \oint_\gamma f(z) n[s_n(z) - s(z)]\mathrm{d}z. \end{aligned} \tag{2.1}$$

The reader is reminded that the above equality may not be correct when some eigenvalues of $n^{-1/2}W_n$ fall outside the contour. However, the probability of this event decays rapidly to zero (see below).

This representation reduces our problem to showing that the following process, indexed by $z \notin [-2, 2]$,

$$\xi_n(z) = n[s_n(z) - s(z)], \tag{2.2}$$

converges in some appropriate space to a Gaussian process $\xi(z)$, $z \notin [-2, 2]$. We will show this conclusion by the following theorem. Throughout the paper we set $\mathbb{C}_0 = \{z = u + \mathrm{i}v : |v| \geqslant v_0\}$.

Theorem 2.1. *Under Conditions 1.1 and 1.2, the process $\{\xi_n(z); \mathbb{C}_0\}$ converges weakly to a Gaussian process $\{\xi(z); \mathbb{C}_0\}$ with the mean and covariance functions given in Propositions 3.1 and 4.1 below.*

Since the mean and covariance functions of $\xi(z)$ are independent of v_0, the process $\{\xi(z); \mathbb{C}_0\}$ in Theorem 2.1 can be taken as a restriction of a process $\{\xi(z)\}$ defined on the whole complex plane, except for the real line. Further, by noticing the symmetry $\xi(\bar{z}) = \overline{\xi(z)}$, and the continuity of the mean and covariance functions of $\xi(z)$ on the real axes except for $z \in [-2, 2]$, we may extend the process to $\{\xi(z); \Re z \notin [-2, 2]\}$.

Split the contour γ into the union $\gamma_u + \gamma_l + \gamma_r + \gamma_0$, where $\gamma_l = \{z = -a + \mathrm{i}y, n^{-2} < |y| \leqslant v_1\}$, $\gamma_r = \{z = a + \mathrm{i}y, n^{-2} < |y| \leqslant v_1\}$ and $\gamma_0 = \{z = \pm a + \mathrm{i}y, |y| \leqslant n^{-2}\}$. By Theorem 2.1 we obtain the weak convergence

$$\int_{\gamma_u} \xi_n(z)\mathrm{d}z \Rightarrow \int_{\gamma_u} \xi(z)\mathrm{d}z.$$

To prove Theorem 1.1, we need only show that for $j = l, r, 0$ and some event Q_n with $\mathbb{P}(Q_n) \to 1$,

$$\lim_{v_1 \downarrow 0} \limsup_{n\to\infty} \int_{\gamma_j} \mathbb{E}|\xi_n(z)\mathbb{1}_{Q_n}|^2 \,\mathrm{d}z = 0 \tag{2.3}$$

and

$$\lim_{v_1 \downarrow 0} \int_{\gamma_j} \mathbb{E}|\xi(z)|^2 \,\mathrm{d}z = 0. \tag{2.4}$$

The estimate (2.4) can be verified directly by the mean and variance functions of $\xi(z)$. The definition of the random event Q_n and the proof of (2.3) for the case $j = 0$ will be given in Section 2.3, and the proof of the non-random and random parts of the limits (2.3) with $j = l$ and r will be given in Sections 3.1 and 4.3, respectively.

2.2. Simplification by truncation

As proposed in Bai and Yin (1988) to control the fluctuations around the extreme eigenvalues, under Conditions 1.1 and 1.2, we will truncate the variables at a convenient rate without altering their weak limit.

Choosing η_n according to (1.2), we first truncate the variables as $\hat{x}_{ij} = x_{ij}\mathbb{1}_{|x_{ij}| \leqslant \eta_n\sqrt{n}}$. We must further normalize them by setting $\tilde{x}_{ij} = (\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij})/s_{ij}$, where s_{ij} is the standard deviation of $\hat{x}_{ij}$ for $i \neq j$ and σs_{ii} is the standard deviation of $\hat{x}_{ii}$.

Let $\hat{F}_n$ and $\tilde{F}_n$ be the empirical spectral distribution of the random matrices $n^{-1/2}(\hat{x}_{ij})$ and $n^{-1/2}(\tilde{x}_{ij})$, respectively. Define $\hat{G}_n$ and $\tilde{G}_n$ similarly by means of (1.1). First, observe that

$$\mathbb{P}(G_n \neq \hat{G}_n) \leqslant \mathbb{P}(F_n \neq \hat{F}_n) = o(1). \tag{2.5}$$

Indeed,

$$\begin{aligned}
\mathbb{P}(F_n \neq \hat{F}_n) &\leqslant \mathbb{P}\{\text{for some } i,\, j,\, \hat{x}_{ij} \neq x_{ij}\} \\
&\leqslant \sum_{i,j} \mathbb{P}\{|x_{ij}| \geqslant \eta_n \sqrt{n}\} \\
&\leqslant (\eta_n \sqrt{n})^{-4} \sum_{i,j} \mathbb{E}[\{|x_{ij}|^4 \mathbb{1}_{|x_{ij}| \geqslant \eta_n \sqrt{n}}\}] = o(1).
\end{aligned}$$

On the other hand, Condition 1.1 implies that

$$\begin{aligned}
\max_{i,j} |1 - s_{ij}| &\leqslant \max_{i,j} |1 - s_{ij}^2| \\
&= \max_{i,j} \left[\mathbb{E}(|x_{ij}|^2 \{\mathbb{1}_{|x_{ij}| \geqslant \eta_n \sqrt{n}}\}) + |\mathbb{E}(x_{ij} \mathbb{1}_{\{|x_{ij}| \geqslant \eta_n \sqrt{n}\}})|^2 \right] \\
&\leqslant (n^{-1} \eta_n^{-2} + M \eta_n^{-6} n^{-3}) \max_{i,j} [\mathbb{E}(|x_{ij}|^2 \{\mathbb{1}_{\{|x_{ij}| \geqslant \eta_n \sqrt{n}\}}\})] \to 0.
\end{aligned} \tag{2.6}$$

Therefore, by Conditions 1.1 and 1.2,

$$\sum_{i,j} \mathbb{E}(|x_{ij}|^2 |1 - s_{ij}^{-1}|^2) \leqslant C \sum_{i,j} (1 - s_{ij}^2)^2 \leqslant \frac{CM}{\eta_n^4 n^2} \sum_{i,j} \mathbb{E}(x_{ij}^4 \mathbb{1}_{\{|x_{ij}| \geqslant \eta_n \sqrt{n}\}}) \to 0.$$

Consequently, as f is analytic, we obtain

$$\begin{aligned}
\mathbb{E}|\tilde{G}_n(f) - \hat{G}_n(f)|^2 &\leqslant C\, \mathbb{E}\left(\sum_{j=1}^n |\tilde{\lambda}_{nj} - \hat{\lambda}_{nj}| \right)^2 \leqslant Cn\, \mathbb{E} \sum_{j=1}^n |\tilde{\lambda}_{nj} - \hat{\lambda}_{nj}|^2 \\
&= Cn\, \mathbb{E} \sum_{i,j} |n^{-1/2} (\tilde{x}_{ij} - \hat{x}_{ij})|^2 \\
&\leqslant C \left[\sum_{i,j} (\mathbb{E}|x_{ij}|^2) |1 - s_{ij}^{-1}|^2 + \sum_{i,j} |\mathbb{E}(\hat{x}_{ij})|^2 s_{ij}^{-2} \right] = o(1),
\end{aligned}$$

where $\tilde{\lambda}_{nj}$ and $\hat{\lambda}_{nj}$ are the jth largest eigenvalues of the Wigner matrices $n^{-1/2}(\tilde{x}_{ij})$ and $n^{-1/2}(\hat{x}_{ij})$, respectively. Therefore the weak limit of the variables $(G_n(f))$ is not affected if we substitute the normalized truncated variables $\tilde{x}_{ij}$ for the original x_{ij}.

From the normalization, the variables $\tilde{x}_{ij}$ all have mean 0 and the same absolute second moments as the original variables. However for the UE, the requirement $\mathbb{E}x_{ij}^2 = 0$ is no longer satisfied after these simplifications. Indeed, we now have $\mathbb{E}\tilde{x}_{ij}^2 = O(1/n)$.

We now assume that the above conditions hold and we use x_{ij} to denote the truncated and normalized variables $\tilde{x}_{ij}$.

Let λ_{ext} be the smallest or the largest eigenvalue of the matrix $n^{-1/2} W_n$ (defined by the truncated and normalized variables). An important consequence of the truncation (see the proof of Theorem 2.12 in Bai 1999) is that, for any $\eta > 0$ and $t > 0$,

$$\mathbb{P}(B_n) = o(n^{-t}), \qquad B_n = \{|\lambda_{\text{ext}}(n^{-1/2}W_n)| \geqslant 2 + \eta\}. \tag{2.7}$$

This property will be used in the proof of Corollary 7.3 below.

2.3. The proof of (2.3) for $j = 0$

If we choose $Q_n \subset B_n^c$, then, when Q_n happens, for any $z \in \gamma_0$ we have $|s_n(z)| \leqslant 2/(a-2)$ and $|s(z)| \leqslant 1/(a-2)$. Hence,

$$\int_{\gamma_0} \mathbb{E}|\xi_n(z)\mathbb{1}_{Q_n}|^2 \leqslant 4n(2/(a-2))^2\|\gamma_0\| = 4n^{-1}(2/(a-2))^2 \to 0.$$

2.4. Known formulae and easy consequences

The Stieltjes transform $s(z)$ of the semicircle law F is given by $s(z) = -\frac{1}{2}(z - \sqrt{z^2-4})$ which satisfies the equation $s^2 + sz + 1 = 0$. Here $\sqrt{z}$ is the square root of z with positive imaginary part. By the symmetry principle, the extended Stieltjes transform of the semicircle law can be found by solving $s(z) = \overline{s(\bar{z})}$, if $\Im z < 0$. To prove Theorem 2.1, we only need to show that $\xi_n(z)\mathbb{1}_{B_n^c} \Rightarrow \xi(z)$ on $\mathbb{C}_0$ with B_n defined in (2.7).

Define $D = (n^{-1/2}W_n - zI_n)^{-1}$. Let α_k be the kth column of W_n with x_{kk} removed and $W_n(k)$ the submatrix extracted from W_n by removing its kth row and kth column. Define $D_k = (n^{-1/2}W_n(k) - zI_{n-1})^{-1}$. Let A^* be the adjoint of A and define the auxiliary variables

$$\beta_k = -n^{-1/2}x_{kk} + z + n^{-1}\alpha_k^* D_k \alpha_k, \tag{2.8}$$

$$\delta(z) = -\frac{1}{n}\sum_{k=1}^{n} \frac{\varepsilon_k}{\beta_k(z + \mathbb{E}s_n(z))}, \tag{2.9}$$

$$\varepsilon_k = \frac{x_{kk}}{\sqrt{n}} - \frac{1}{n}\alpha_k^* D_k \alpha_k + \mathbb{E}s_n(z). \tag{2.10}$$

The Stieltjes transform $s_n(z)$ of F_n has the representation

$$\begin{aligned} s_n(z) &= \frac{1}{n}\operatorname{tr} D = \frac{1}{n}\operatorname{tr}\left(\frac{W_n}{\sqrt{n}} - zI_n\right)^{-1} \\ &= -\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\beta_k} = -\frac{1}{z + \mathbb{E}s_n(z)} + \delta(z). \end{aligned} \tag{2.11}$$

In particular,

$$\mathbb{E}s_n(z) = s(z + \mathbb{E}\delta(z)) + \mathbb{E}\delta(z). \tag{2.12}$$

Moreover, under Condition 1.1, we have for $z \in \mathbb{C}_0$ and some generic constant K,

$$|\mathbb{E}\delta(z)| \leqslant Kn^{-1}. \tag{2.13}$$

This entails, in particular,

$$|\mathbb{E}s_n(z) - s(z)| \leqslant Kn^{-1}. \tag{2.14}$$

Secondly, for any $p > 1$ such that $\sup_{i,j} \mathbb{E}|x_{ij}|^p < \infty$,

$$\mathbb{E}|\alpha_k^* D_k \alpha_k - \mathrm{tr} D_k|^p \leqslant Kn^{p/2}, \tag{2.15}$$

$$\mathbb{E}|\mathrm{tr} D_k - \mathbb{E}\mathrm{tr} D_k|^p \leqslant Kn^{p/2}. \tag{2.16}$$

Inequality (2.13) follows from (4.26) of Bai (1993), and (2.15) and (2.16) follow from Lemmas 2.1 and 2.6 of Bai *et al.* (2002).

Finally, let us mention a 'trick' frequently played on the matrix D or D_k: as W_n is Hermitian (or real symmetric), the eigenvalues of D are of the form $1/(\lambda_j - z)$ with real λ_j. All these values are then bounded by $1/v$, so that $|\mathrm{tr}D| \leqslant n/v$ and $|\mathrm{tr}D\overline{D}| \leqslant n/v^2$. Another useful equality from this spectral decomposition of D is the following differentiation rule (with respect to z): $\mathrm{d}D(z)/\mathrm{d}z = D^2(z)$.

3. The mean function of ξ_n

An expansion from (2.12) gives for the mean function $b_n(z) := \mathbb{E}\xi_n(z)$,

$$b_n(z) := n[\mathbb{E}s_n(z) - s(z)] = [1 + s'(z)]n\,\mathbb{E}\delta(z)\{1 + o(1)\}.$$

Let the index k always run from 1 to n. When we derive a bound or a limit for some expression g_k depending on k (e.g. $\mathrm{tr}D_k$), the result holds *uniformly in* k (i.e. the bound or the limit is independent of k), so that the same is true for the mean $n^{-1}(g_1 + \ldots + g_n)$.

Proposition 3.1. *The mean function* $b_n(z)$ *uniformly tends to*

$$b(z) = [1 + s'(z)]s^3(z)[\sigma^2 - 1 + (\kappa - 1)s'(z)\beta s^2(z)]$$

for $z \in \mathbb{C}_0$ *and for both the UE and OE.*

Proof. By definition,

$$n\delta(z) = -\sum_{k=1}^{n} \frac{\varepsilon_k}{\beta_k[z + \mathbb{E}s_n(z)]}.$$

We aim to prove that $n\,\mathbb{E}\delta$ tends to a limit $d(z) = s^3(z)[\sigma^2 - 1 + (\kappa - 1)s'(z)\beta s^2(z)]$. Using the identity

$$\frac{1}{u - \varepsilon} = \frac{1}{u}\left[1 + \frac{\varepsilon}{u} + \ldots + \frac{\varepsilon^p}{u^p} + \frac{\varepsilon^{p+1}}{u^p(u - \varepsilon)}\right]$$

for any integer p, we obtain

$$n\delta(z) = -\sum_{k=1}^{n} \frac{\varepsilon_k}{[z + \mathbb{E}s_n(z)]^2} - \sum_{k=1}^{n} \frac{\varepsilon_k^2}{[z + \mathbb{E}s_n(z)]^3} - \sum_{k=1}^{n} \frac{\varepsilon_k^3}{\beta_k[z + \mathbb{E}s_n(z)]^3}$$

$$= S_1 + S_2 + S_3.$$

First, we prove that $\mathbb{E}S_3 = o(1)$. We have $|\beta_k| \geqslant v$ because

$$|\beta_k| = \left| z - \frac{x_{kk}}{\sqrt{n}} + n^{-1}\alpha_k^* D_k \alpha_k \right|$$

$$\geqslant \operatorname{Im}\left(z - \frac{x_{kk}}{\sqrt{n}} + n^{-1}\alpha_k^* D_k \alpha_k \right)$$

$$= v(1 + n^{-1}\alpha_k^* D_k \overline{D}_k \alpha_k).$$

Therefore, by Lemma 7.2 below,

$$|\mathbb{E}S_3| \leqslant v^{-1}|z + \mathbb{E}s_n(z)|^{-3}\, \mathbb{E}\sum_{k=1}^{n} |\varepsilon_k^3| = o(1).$$

For $\mathbb{E}S_1$ we have

$$\mathbb{E}\varepsilon_k = \mathbb{E}\left(\frac{x_{kk}}{\sqrt{n}} - n^{-1}\alpha_k^* D_k \alpha_k \right) + \mathbb{E}s_n(z)$$

$$= n^{-1}[\mathbb{E}\mathrm{tr}D - \mathbb{E}\mathrm{tr}D_k].$$

On the other hand,

$$\mathrm{tr}D - \mathrm{tr}D_k = \frac{1 + n^{-1}\alpha_k^* D_k^2 \alpha_k}{n^{-1/2}x_{kk} - z - n^{-1}\alpha_k^* D_k \alpha_k}.$$

By inequalities (2.15) and (2.16), it is easy to see that

$$n^{-1/2}x_{kk} - z - n^{-1}\alpha_k^* D_k \alpha_k \xrightarrow{L_2} -z - s(z),$$

$$1 + n^{-1}\alpha_k^* D_k^2 \alpha_k = 1 + [n^{-1}\alpha_k^* D_k \alpha_k]'$$

$$\xrightarrow{L_2} 1 + s'(z).$$

Therefore, as $|\beta_k| \geqslant v$ it follows that

$$\mathrm{tr}D - \mathrm{tr}D_k \xrightarrow{L_1} -\frac{1 + s'(z)}{z + s(z)} = s(z)[1 + s'(z)].$$

Hence,

$$\mathbb{E}S_1 \xrightarrow{L_1} -s^3(z)[1 + s'(z)]. \tag{3.1}$$

For the term $\mathbb{E}S_2$, by the previous estimate for $\mathbb{E}\varepsilon_k$, we have

$$\mathbb{E}\varepsilon_k^2 = \mathbb{E}(\varepsilon_k - \mathbb{E}\varepsilon_k)^2 + O(n^{-2}).$$

Furthermore, by the definition of ε_k, we have

$$\begin{aligned}\varepsilon_k - \mathbb{E}\varepsilon_k &= x_{kk}/\sqrt{n} - n^{-1}[\alpha_k^* D_k \alpha_k - \mathbb{E}\mathrm{tr}D_k] \\ &= x_{kk}/\sqrt{n} - n^{-1}[\alpha_k^* D_k \alpha_k - \mathrm{tr}D_k] + n^{-1}[\mathrm{tr}D_k - \mathbb{E}\mathrm{tr}D_k].\end{aligned}$$

Therefore

$$\mathbb{E}[\varepsilon_k - \mathbb{E}\varepsilon_k]^2 = \frac{\sigma^2}{n} + \frac{1}{n^2}\mathbb{E}[\alpha_k^* D_k \alpha_k - \mathrm{tr}D_k]^2 + \frac{1}{n^2}\ \mathbb{E}[\mathrm{tr}D_k - \mathbb{E}\mathrm{tr}D_k]^2. \qquad (3.2)$$

It will be proved later that

$$\mathbb{E}[\mathrm{tr}D_k - \mathbb{E}\mathrm{tr}D_k]^2 \to c(z),$$

for some function $c(z)$, so that we neglect the last term in (3.2). To evaluate the second term, let us use the notation $\alpha_k = (\xi_i)$ and $D_k = (d_{ij})$. We have

$$\begin{aligned}\mathbb{E}[\alpha_k^* D_k \alpha_k - \mathrm{tr}D_k]^2 &= \mathbb{E}\left[\sum_{i\neq j} d_{ij}\bar{\xi}_i\xi_j + \sum_i d_{ii}(|\xi_i|^2 - 1)\right]^2 \\ &= \mathbb{E}\left[\sum_{i\neq j}\sum_{s\neq t} d_{ij}d_{st}\bar{\xi}_i\xi_j\bar{\xi}_s\xi_t\right] + \mathbb{E}\left[\sum_i d_{ii}^2(|\xi_i|^2-1)^2\right] \\ &= \mathbb{E}(\bar{\xi}_1)^2\,\mathbb{E}(\xi_2)^2\,\mathbb{E}\left[\sum_{i\neq j} d_{ij}^2\right] + \mathbb{E}|\xi_1|^2\,\mathbb{E}|\xi_2|^2\,\mathbb{E}\left[\sum_{i\neq j} d_{ij}d_{ji}\right] \\ &\quad + \mathbb{E}(|\xi_1|^2-1)^2\,\mathbb{E}\left[\sum_i d_{ii}^2\right] \\ &= \kappa\mathbb{E}\left[\sum_{i,j} d_{ij}d_{ji}\right] + \beta\,\mathbb{E}\left[\sum_i d_{ii}^2\right] + o(n). \qquad (3.3)\end{aligned}$$

Here we see the main difference between the UE and OE. For the UE, we need the assumption $\mathbb{E}(x_{ij}^2) = 0$ for the original variables, which implies $\mathbb{E}(\xi_i^2) = o(1)$ for the truncated ones. Without this assumption, it is difficult to deal with the limit of $\sum_{i,j} d_{ij}^2 = \mathrm{tr}(D_k D_k^{\mathrm{T}})$. The introduction of the parameter κ, taking different values on the two ensembles, allows us to give a unified expression for the computation above.

Noticing that

$$\sum_{i,j} d_{ij}d_{ji} = \mathrm{tr}D_k^2 = \frac{\mathrm{d}}{\mathrm{d}z}\mathrm{tr}D_k,$$

we then have

$$n\,\mathbb{E}\varepsilon_k^2 = \sigma^2 + \kappa\,\mathbb{E}[n^{-1}\mathrm{tr}D_k^2] + \beta\,\mathbb{E}\left[n^{-1}\sum_i d_{ii}^2\right] + o(1).$$

Furthermore, by Lemma 7.1 below,

$$\lim_n \frac{1}{n}\mathrm{tr}D_k^2 = \lim_n \left(\frac{1}{n}\mathrm{tr}D_k\right)' = s'(z), \quad \text{in } L_1,$$

$$\lim_n \frac{1}{n}\sum_i d_{ii}^2 = \frac{1}{[-z-s(z)]^2} = s^2(z), \quad \text{in } L_1.$$

Hence

$$\sum_{k=1}^n \mathbb{E}\varepsilon_k^2 = \sigma^2 + \kappa s'(z) + \beta s^2(z) + o(1).$$

Summing the three terms, we obtain

$$n\,\mathbb{E}\delta(z) = s^3\left(\sigma^2 - 1 + (\kappa-1)s' + \beta s^2\right) + o(1).$$

The proposition is proved. □

3.1. Proof of the non-random part of (2.3) for $j = l,\, r$

Using the notation defined in this section, we proceed to prove, for $j = l$ or r

$$\lim_{v_1\downarrow 0}\limsup_{n\to\infty}\int_{\gamma_j} |\mathbb{E}\xi_n(z)\mathbb{1}_{Q_n}|^2\,\mathrm{d}z = 0. \tag{3.4}$$

By symmetry, we need only consider the case $j = l$. Let $z \in \gamma_l$, that is, $z = -a + \mathrm{i}v$ with $n^{-2} < |v| < v_1$. Without loss of generality we may assume that $v_1 < a/2 - 1$.

Note that $\Re s(z) > 0$ for all $z \in \gamma_l$. Thus, we have

$$\nu := \inf_{z=-a+\mathrm{i}v,|v|<v_0} \Re\{-(z+s(z))\} = \inf_{z=-a+\mathrm{i}v,|v|<v_0} \Re\left\{\frac{1}{s(z)}\right\} > 0.$$

Let B_n be the event defined in (2.7) with $\eta = \frac{1}{2}(a-2)$. Now define $Q_n = B_n^c \cap_k \{|\beta_k| > \nu/3)\}$ and $B_{nk} = \{|\lambda_{\mathrm{ext}}(n^{-1/2}W_k)| \geqslant 1 + a/2\}$. By the interlacing theorem (see Rao and Rao, 2001, p. 328), we have $B_{nk} \subset B_n$. Multiplying both sides of (2.11) by $\mathbb{1}_{Q_n}$ gives the following expressions analogous to those used in the proof of Proposition 3.1:

$$\tilde{s}_n(z) = -\frac{1}{z + \mathbb{E}\tilde{s}_n(z)} + \tilde{\delta}(z),$$

$$\begin{aligned} n\tilde{\delta}(z) &= \frac{n\mathbb{1}_{Q_n^c}}{z+\mathbb{E}\tilde{s}_n(z)} - \sum_{k=1}^n \frac{\tilde{\varepsilon}_k\mathbb{1}_{Q_n}}{[z+\mathbb{E}\tilde{s}_n(z)]^2} - \sum_{k=1}^n \frac{\tilde{\varepsilon}_k^2\mathbb{1}_{Q_n}}{\beta_k[z+\mathbb{E}\tilde{s}_n(z)]^2} \\ &= \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3, \end{aligned} \tag{3.5}$$

where $\tilde{s}_n(z) = s_n(z)\mathbb{1}_{Q_n}$ and $\tilde{\varepsilon}_k = n^{-1/2}x_{kk} - n^{-1}\alpha_k^* D_k \alpha_k + \mathbb{E}\tilde{s}_n(z)$.

Note that $\mathbb{E}\tilde{s}_n \to s$ uniformly on γ_l. From (3.5), we find that $\mathbb{E}\tilde{\delta}(z) \to 0$ uniformly. Our next goal is to show that for any fixed $t > 0$,

$$\mathbb{P}(Q_n^c) = o(n^{-t}), \tag{3.6}$$

which is a consequence of

$$\sum_{k=1}^n \mathbb{P}(|\beta_k| \leqslant \nu/3) = o(n^{-t}). \tag{3.7}$$

Observe that $|z + \mathbb{E}\tilde{s}_n| \geqslant 2\nu/3$, for all large n, which, together with $|\beta_k| \leqslant \nu/3$, implies that $|\tilde{\varepsilon}_k| \geqslant \nu/3$. Thus (3.7) follows from Corollary 7.3 below. Hence, (3.6) is proved.

On the other hand, we have $\mathbb{E}\tilde{s}_n(z) = s(z + \mathbb{E}\tilde{\delta}(z)) + \mathbb{E}\tilde{\delta}(z)$, which, together with facts that $\mathbb{E}\tilde{\delta} \to 0$ and that $s'(z)$ is uniformly bounded when $|\Re(z)| \geqslant a/2 + 1$, implies that

$$|\mathbb{E}\tilde{s}_n(z) - s(z)| \leqslant K|\mathbb{E}\tilde{\delta}(z)|.$$

Therefore,

$$\begin{aligned} n|\mathbb{E}(s_n(z) - s(z))\mathbb{1}_{Q_n}| &\leqslant n|s(z)|\mathbb{P}(Q_n^c) + n|\mathbb{E}\tilde{s}_n(z) - s(z)| \\ &\leqslant o(n^{-t}) + Kn|\mathbb{E}(\tilde{\delta}(z))|. \end{aligned}$$

This reduces the proof of (3.4) to showing that $\mathbb{E}\tilde{S}_j$, $j = 1, 2, 3$, are all uniformly bounded on γ_l.

We obviously have that, for all large n,

$$|\mathbb{E}\tilde{S}_1| \leqslant (3/2\nu)n\mathbb{P}(Q_n^c) = o(n^{-t}),$$

$$|\mathbb{E}\tilde{S}_3| \leqslant (3/\nu)^3 2^{-2} \sum_{k=1}^n \mathbb{E}|\tilde{\varepsilon}_k|^2 \mathbb{1}_{Q_n} = O(1).$$

To complete the proof of (3.4), we need to prove that $\mathbb{E}\tilde{S}_2$ is uniformly bounded for $z \in \gamma_l$. First, we have

$$\begin{aligned} |n\,\mathbb{E}\tilde{\varepsilon}_k \mathbb{1}_{B_{nk}^c}| &= |\,\mathbb{E}\mathrm{tr}D\mathbb{1}_{Q_n}\mathbb{P}(B_{nk}^c) - \mathbb{E}\mathrm{tr}D_k\mathbb{1}_{B_{nk}^c}| \\ &\leqslant |\,\mathbb{E}(\mathrm{tr}D - \mathrm{tr}D_k)\mathbb{1}_{B_n^c}| + |\mathbb{E}\mathrm{tr}D\mathbb{1}_{Q_n}|\mathbb{P}(B_{nk}) \\ &\quad + |\mathbb{E}\mathrm{tr}D\mathbb{1}_{Q_n^c B_n^c}| + |\mathbb{E}\mathrm{tr}D_k\mathbb{1}_{B_{nk}^c B_n}| \\ &\leqslant K + Kn\left[2\mathbb{P}(B_n) + \sum_{k=1}^n \mathbb{P}(|\tilde{\varepsilon}_k| \geqslant \nu/3,\ B_n^c)\right] \leqslant K. \end{aligned}$$

Next, we have

$$\begin{aligned} |n\,\mathbb{E}\tilde{\varepsilon}_k\mathbb{1}_{B_{nk}^c} - n\,\mathbb{E}\tilde{\varepsilon}_k\mathbb{1}_{Q_n}| &\leqslant n\mathbb{E}|\tilde{\varepsilon}_k[\mathbb{1}_{Q_n B_{nk}} + \mathbb{1}_{Q_n^c B_{nk}^c}]| \\ &\leqslant Kn[\mathbb{P}(B_{nk}) + \mathbb{P}(Q_n^c)] \to 0. \end{aligned}$$

Hence,

$$|\mathbb{E}\tilde{S}_1| \leqslant (2\nu/3)^{-2} \sum_{k=1}^{n} |\mathbb{E}\varepsilon_k \mathbb{1}_{Q_n}| \leqslant K.$$

The proof of (3.4) is then complete.

4. Convergence of the process $\zeta_n = \xi_n - \mathbb{E}\xi_n$

In this section we consider the convergence of the random part of the process ξ_n given by

$$\zeta_n = \{\zeta_n(z) = \xi_n(z) - \mathbb{E}\xi_n(z),\, z \in \mathbb{C}_0\}.$$

To this end we establish the finite-dimensional convergence and the tightness of $\zeta_n(z)$.

4.1. Finite-dimensional convergence of (ζ_n)

Let $\mathcal{F}_k = \sigma(x_{ij},\, k+1 \leqslant i, j \leqslant n)$ for $0 \leqslant k \leqslant n$ and $\mathbb{E}_k(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_k)$. Note that the filtration $(\mathcal{F}_k)_k$ is decreasing in k. Let us also mention that for each n, the entries x_{ij} could depend on n; in that case we would have a sequence of filtrations $\mathcal{F}^{(n)} := (\mathcal{F}_k^{(n)})_{1\leqslant k\leqslant n}$. This possible dependence on n will be assumed throughout the paper.

The following martingale decomposition is well known (see (2.12) in Bai *et al.* 2002):

$$\zeta_n(z) := \xi_n(z) - \mathbb{E}\xi_n(z) = \mathrm{tr}D - \mathbb{E}\mathrm{tr}D = \sum_{k=1}^{n} u_k,$$

where

$$\begin{aligned} u_k &= (\mathbb{E}_{k-1} - \mathbb{E}_k)\mathrm{tr}D = (\mathbb{E}_{k-1} - \mathbb{E}_k)(\mathrm{tr}D - \mathrm{tr}D_k) \\ &= (\mathbb{E}_{k-1} - \mathbb{E}_k)a_k - \mathbb{E}_{k-1}b_k, \\ a_k &= -\frac{(1 + n^{-1}\alpha_k^* D_k^2 \alpha_k)g_k}{\beta_k(z + n^{-1}\mathrm{tr}D_k)}, \qquad b_k = \frac{h_k}{(z + n^{-1}\mathrm{tr}D_k)}, \\ g_k &:= n^{-1/2}x_{kk} - n^{-1}(\alpha_k^* D_k \alpha_k - \mathrm{tr}D_k) \\ h_k &:= n^{-1}(\alpha_k^* D_k^2 \alpha_k - \mathrm{tr}(D_k^2)). \end{aligned} \tag{4.1}$$

We have

$$
\begin{aligned}
a_k &= -\frac{1+n^{-1}\alpha_k^* D_k^2 \alpha_k}{z+n^{-1}\mathrm{tr} D_k}\frac{g_k}{\beta_k} \\
&= -\frac{1+n^{-1}\mathrm{tr}(D_k^2) g_k}{(z+n^{-1}\mathrm{tr} D_k)^2} - \frac{h_k g_k}{(z+n^{-1}\mathrm{tr} D_k)^2} - \frac{(1+n^{-1}\alpha_k^* D_k^2 \alpha_k) g_k^2}{(z+n^{-1}\mathrm{tr} D_k)^2 \beta_k} \\
&:= a_{k1} + a_{k2} + a_{k3}.
\end{aligned}
$$

By Lemma 7.2, for $z \in \mathbb{C}_0$,

$$
\begin{aligned}
\mathbb{E}\left| \sum_{k=1}^{n} (\mathbb{E}_{k-1} - \mathbb{E}_k) a_{k3} \right|^2 &= \sum_{k=1}^{n} \mathbb{E}|(\mathbb{E}_{k-1} - \mathbb{E}_k) a_{k3}|^2 \\
&\leqslant \sum_{k=1}^{n} \mathbb{E}\left| \frac{(1+n^{-1}\alpha_k^* D_k^2 \alpha_k) g_k^2}{(z+n^{-1}\mathrm{tr} D_k)^2 (g_k - (z+n^{-1}\mathrm{tr} D_k))} \right|^2 \\
&\leqslant v^{-6} \sum_{k=1}^{n} \mathbb{E}|g_k|^4 = o(1). \qquad (4.2)
\end{aligned}
$$

We also have

$$
\begin{aligned}
\mathbb{E}\left| \sum_{k=1}^{n} (\mathbb{E}_{k-1} - \mathbb{E}_k) a_{k2} \right|^2 &= \sum_{k=1}^{n} \mathbb{E}|(\mathbb{E}_{k-1} - \mathbb{E}_k) a_{k2}|^2 \\
&\leqslant \sum_{k=1}^{n} \mathbb{E}\left| \frac{h_k g_k}{(z+n^{-1}\mathrm{tr} D_k)^2} \right|^2 \\
&\leqslant v^{-4} \sum_{k=1}^{n} (\mathbb{E}|h_k|^4 \, \mathbb{E}|g_k|^4)^{1/2} = o(1), \qquad (4.3)
\end{aligned}
$$

where we have used the facts that $|(1+n^{-1}\alpha_k^* D_k^2 \alpha_k)/(g_k - (z+n^{-1}\mathrm{tr} D_k))| \leqslant v^{-1}$ and $\Im(z+n^{-1}\mathrm{tr} D_k) \geqslant v$. Furthermore,

$$
\begin{aligned}
(\mathbb{E}_{k-1} - \mathbb{E}_k) a_{k1} &= (\mathbb{E}_{k-1} - \mathbb{E}_k)\left[-\frac{1+n^{-1}\mathrm{tr}(D_k^2)}{(z+n^{-1}\mathrm{tr} D_k)^2} g_k \right] \\
&= -\mathbb{E}_{k-1} \frac{1+n^{-1}\mathrm{tr}(D_k^2)}{(z+n^{-1}\mathrm{tr} D_k)^2} g_k,
\end{aligned}
$$

where the last equality follows from the fact that the conditional expectation with respect to $\mathbb{E}_k$ is zero. Hence,

$$\zeta_n(z) = \sum_{k=1}^{n} \mathbb{E}_{k-1}\left[-\frac{1+n^{-1}\mathrm{tr}(D_k^2)}{(z+n^{-1}\mathrm{tr}D_k)^2}g_k - b_k\right] + o_p(1)$$

$$= \sum_{k=1}^{n} \mathbb{E}_{k-1}\phi_k(z) + o_p(1),$$

where we have denoted by $\phi_k(z)$ the term in the square brackets.

Let $\{z_s, s = 1, \dots, p\}$ be p different points belonging to $\mathbb{C}_0$. The problem is then reduced to seeking the weak convergence of the vector martingale

$$\mathbf{A}_n := \sum_{k=1}^{n} \mathbb{E}_{k-1}(\phi_k(z_1), \dots, \phi_k(z_p)) =: \sum_{k=1}^{n} \mathbb{E}_{k-1}\phi_k. \tag{4.4}$$

Proposition 4.1. *Assume that Conditions 1.1 and 1.2 are satisfied. For any set of p points* $\{z_s, s = 1, \dots, p\}$ *of* $\mathbb{C}_0$*, the random vector* $(\xi(z_1), \dots, \xi(z_p))$ *converges weakly to a p-dimensional zero-mean Gaussian distribution with covariance matrix given by*

$$\Gamma(z_j, z_s) = \frac{\partial^2}{\partial z_j \partial z_s}\left[(\sigma^2 - \kappa)s_j s_s + \frac{1}{2}\beta(s_j s_s)^2 - \kappa\log(1 - s_j s_s)\right]$$

$$= s_j' s_s'\left[\sigma^2 - \kappa + 2\beta s_j s_s + \frac{\kappa}{(1 - s_j s_s)^2}\right], \tag{4.5}$$

with $s_j = s(z_j)$.

Proof (Part 1). We apply a CLT to the martingale $\mathbf{A}_n$ defined in (4.4). Consider its *hook* process:

$$\Gamma_n(z_i, z_j) := \sum_{k=1}^{n} \mathbb{E}_k[\mathbb{E}_{k-1}\phi_k(z_i)\mathbb{E}_{k-1}\phi_k(z_j)].$$

Then we have to check the following two conditions:

Condition 4.1. Γ_n *converges in probability to the matrix* Γ*;*

Condition 4.2 Lyapunov condition. *For some* $a > 2$*,*

$$\sum_{k=1}^{n} \mathbb{E}_k[\|\mathbb{E}_{k-1}\phi_k\|^a] \xrightarrow{p} 0.$$

The verification of Condition 4.1 and the computation of the limit Γ are lengthy and delayed to Section 4.4. Here we prove that Condition 4.2 is satisfied with $a = 4$. We will prove that

$$\mathbb{E}\sum_{k=1}^{n}\mathbb{E}_k[\|\mathbb{E}_{k-1}\phi_k\|^4] = \sum_{k=1}^{n}\mathbb{E}[\|\mathbb{E}_{k-1}\phi_k\|^4] \to 0.$$

Note that

$$\|\mathbb{E}_{k-1}\phi_k\|^4 \leqslant K[|\mathbb{E}_{k-1}\phi_k(z_1)|^4 + \ldots + |\mathbb{E}_{k-1}\phi_k(z_p)|^4],$$

so that it will be sufficient to establish that for any $z \in \mathbb{C}_0$,

$$S_n := \sum_{k=1}^{n}\mathbb{E}|\mathbb{E}_{k-1}\phi_k(z)|^4 \to 0.$$

By Jensen's inequality, we obtain $S_n \leqslant \sum_{k=1}^{n}\mathbb{E}|\phi_k(z)|^4$. By definition of ϕ_k and b_k, we have, by the 'trick' explained at the end of Section 2,

$$\begin{aligned}|\phi_k(z)| &= \left|\frac{1+n^{-1}\mathrm{tr}(D_k^2)}{(z+n^{-1}\mathrm{tr}D_k)^2}g_k + \frac{h_k}{(z+n^{-1}\mathrm{tr}D_k)}\right| \\ &\leqslant \frac{1+v^{-1}}{v^2}|g_k| + \frac{1}{v}|h_k|,\end{aligned}$$

Under Condition 1.1 and by Lemma 7.2, we thus have $\mathbb{E}|\phi_k(z)|^4 \leqslant Kn^{-1}\eta_n^4$. Hence $S_n = o(1)$ and Condition 4.2 is satisfied. □

4.2. Tightness of (ζ_n)

It is enough to establish the following Hölder condition: for some positive constant K,

$$\mathbb{E}|\zeta_n(z_1) - \zeta_n(z_2)|^2 \leqslant K|z_1 - z_2|^2, \qquad z_1, z_2 \in \mathbb{C}_0. \tag{4.6}$$

By the definition of ζ_n, we have

$$\begin{aligned}\mathbb{E}|\zeta_n(z_1) - \zeta_n(z_2)|^2 &= |z_1 - z_2|^2\,\mathbb{E}|\mathrm{tr}D(z_1)D(z_2) - \mathbb{E}\mathrm{tr}D(z_1)D(z_2)|^2 \\ &= |z_1 - z_2|^2\sum_{k=1}^{n}\mathbb{E}|(\mathbb{E}_{k-1} - \mathbb{E}_k)(\mathrm{tr}D(z_1)D(z_2))|^2.\end{aligned}$$

Using the formula

$$\begin{pmatrix}\sigma & \alpha^* \\ \alpha & \Sigma\end{pmatrix}^{-1} = \theta\begin{pmatrix}1 & -\alpha^*\Sigma^{-1} \\ -\Sigma^{-1}\alpha & (\theta\Sigma)^{-1} + \Sigma^{-1}\alpha\alpha^*\Sigma^{-1}\end{pmatrix}, \qquad \theta := \frac{1}{\sigma - \alpha^*\Sigma^{-1}\alpha},$$

we obtain

$$\begin{aligned}\mathrm{tr}D(z_1)D(z_2) - \mathrm{tr}D_k(z_1)D_k(z_2) \leqslant{}& \frac{|1 + n^{-1}\alpha_k^*D_k(z_1)D_k(z_2)\alpha_k|^2}{\beta_k(z_1)\beta_k(z_2)} \\ &+ \frac{\alpha_k^*D_k(z_1)D_k(z_2)D_k(z_1)\alpha_k}{n\beta_k(z_1)} + \frac{\alpha_k^*D_k(z_2)D_k(z_1)D_k(z_2)\alpha_k}{n\beta_k(z_2)}.\end{aligned}$$

If $M = D_k(z_1)D_k(z_2)D_k(z_1)$, then by Lemma 7.2 we have

$$\begin{aligned}
&\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{\alpha_k^* M\alpha_k}{n\beta(z_1)}\right|^2\\
&\quad=\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{\alpha_k^* M\alpha_k-\mathrm{tr}M}{n\beta(z_1)}-\frac{\mathrm{tr}M(\alpha_k^* D_k(z_1)\alpha_k-\mathrm{tr}D_k(z_1))}{n^2\beta_k(z_1)(z+n^{-1}\mathrm{tr}D_k)}\right|^2\\
&\quad\leqslant Kn^{-2}[\mathbb{E}|\alpha_k^* M\alpha_k-\mathrm{tr}M|^2+\mathbb{E}|\alpha_k^* D_k(z_1)\alpha_k-\mathrm{tr}D_k(z_1)|^2]\\
&\quad\leqslant Kn^{-1}.
\end{aligned}$$

Therefore,

$$\sum_{k=1}^{n}\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{\alpha_k^* D_k(z_1)D_k(z_2)D_k(z_1)\alpha_k}{n\beta_k(z_1)}\right|^2\leqslant K.$$

Similarly,

$$\sum_{k=1}^{n}\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{\alpha_k^* D_k(z_2)D_k(z_1)D_k(z_2)\alpha_k}{n\beta_k(z_2)}\right|^2\leqslant K.$$

By a similar decomposition approach, one may prove that

$$\begin{aligned}
&\sum_{k=1}^{n}\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{|1+n^{-1}\alpha_k^* D_k(z_1)D_k(z_2)\alpha_k|^2}{\beta_k(z_1)\beta_k(z_2)}\right|^2\\
&=\sum_{k=1}^{n}\mathbb{E}\left|(\mathbb{E}_{k-1}-\mathbb{E}_k)\frac{|1+n^{-1}\alpha_k^* D_k(z_1)D_k(z_2)\alpha_k|^2}{\beta_k(z_1)\beta_k(z_2)}-\frac{|1+n^{-1}\mathrm{tr}D_k(z_1)D_k(z_2)|^2}{(z_1+n^{-1}\mathrm{tr}D_k(z_1))(z_2+n^{-1}\mathrm{tr}D_k(z_2))}\right|^2\\
&\leqslant K.
\end{aligned}$$

Then (4.6) follows from the above estimates and the fact that

$$(\mathbb{E}_{k-1}-\mathbb{E}_k)\mathrm{tr}(D(z_1)D(z_2))=(\mathbb{E}_{k-1}-\mathbb{E}_k)[\mathrm{tr}(D(z_1)D(z_2))-\mathrm{tr}(D_k(z_1)D_k(z_2))].$$

4.3. Conclusion of the proof of (2.3) for $j = l, r$

By (3.4), to complete the proof of (2.3), we need only show that

$$\lim_{v_1\downarrow 0}\limsup_{n\to\infty}\int_{\gamma_l}\mathbb{E}|(\xi_n(z)\mathbb{1}_{Q_n}-\mathbb{E}\xi_n(n)\mathbb{1}_{Q_n})|^2\,\mathrm{d}z=0. \tag{4.7}$$

Using the same expansion as (4.1) and a slightly different decomposition, we have

$$\xi_n(z)\mathbb{1}_{Q_n} - \mathbb{E}\xi_n(n)\mathbb{1}_{Q_n} = \sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)\mathrm{tr}D\mathbb{1}_{Q_n}$$

$$= \sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)u_k\mathbb{1}_{Q_n} + \sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)\mathrm{tr}D_k\mathbb{1}_{Q_n},$$

$$u_k = -\frac{1 + n^{-1}\alpha_k^* D_k^2\alpha_k}{\beta_k} = u_{k0} + u_{k1} + u_{k2}$$

where

$$u_{k0} = -\frac{1 + n^{-1}\mathrm{tr}(D_k^2)}{z + \mathbb{E}\tilde{s}_n(z)},$$

$$u_{k1} = -\frac{\alpha_k^* D_k^2\alpha_k - \mathrm{tr}(D_k^2)}{n(z + \mathbb{E}\tilde{s}_n(z))},$$

$$u_{k2} = -\frac{(1 + n^{-1}\alpha_k^* D_k^2\alpha_k)\tilde{\varepsilon}_k}{\beta_k(z + \mathbb{E}\tilde{s}_n(z))}.$$

When $z \in \gamma_l$, we have $|\mathrm{tr}D_k(z)| \leqslant nv^{-1} \leqslant n^3$, and therefore

$$\mathbb{E}\left|\sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)\mathrm{tr}D_k\mathbb{1}_{Q_n}\right|^2 = \sum_{k=1}^{n}\mathbb{E}|(\mathbb{E}_{k-1} - \mathbb{E}_k)\mathrm{tr}D_k\mathbb{1}_{Q_n^c}|^2 \leqslant n^4\mathbb{P}(Q_n^c) = o(1).$$

Also, noticing that $(\mathbb{E}_{k-1} - \mathbb{E}_k)u_{k0} = 0$, to prove (4.7), we need only verify that

$$\mathbb{E}\left|\sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)u_{kj}\mathbb{1}_{Q_n}\right|^2 \leqslant K, \qquad j = 1, 2.$$

Since $z + \mathbb{E}\tilde{s}_n(z)$ is bounded away from 0, we have

$$\mathbb{E}\left|\sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)u_{k1}\mathbb{1}_{Q_n}\right|^2 \leqslant Kn^{-2}\sum_{k=1}^{n}\mathbb{E}|\alpha_k^* D_k^2\alpha_k - \mathrm{tr}(D_k^2)|^2\mathbb{1}_{Q_n}$$

$$\leqslant Kn^{-2}\sum_{k=1}^{n}\mathbb{E}|\alpha_k^* D_k^2\alpha_k - \mathrm{tr}(D_k^2)|^2\mathbb{1}_{B_{nk}^c}$$

$$\leqslant K/(a/2 - 1)^4$$

and

$$\mathbb{E}\left|\sum_{k=1}^{n}(\mathbb{E}_{k-1}-\mathbb{E}_k)u_{k2}\mathbb{1}_{Q_n}\right|^2 \leqslant K\sum_{k=1}^{n}\mathbb{E}\big(1+n^{-1}|\alpha_k|^2\big)|\tilde{\varepsilon}_k|^2\mathbb{1}_{Q_n}$$

$$\leqslant K\sum_{k=1}^{n}\mathbb{E}|\tilde{\varepsilon}_k|^2\mathbb{1}_{Q_n}+Kn\eta_n^2\sum_{k=1}^{n}\mathbb{E}|\tilde{\varepsilon}_k|^2\mathbb{1}_{Q_n\cap\{|\alpha_k|^2\geqslant 2n\}}$$

$$\leqslant K+Kn\sum_{k=1}^{n}\mathbb{E}^{1/2}|\tilde{\varepsilon}_k|^4\mathbb{1}_{Q_n}P^{1/2}(|\alpha_k|^2\geqslant 2n)\leqslant K.$$

Here we have used the fact that $\mathbb{P}(|\alpha_k|^2 \geqslant 2n) = o(n^{-t})$ for any fixed $t > 0$ and uniformly in $k \leqslant n$. This fact can be regarded as a consequence of Lemma 7.2.

4.4. Conclusion of the proof of Proposition 4.1

The goal is to check Condition 4.1 introduced in Part 1 of this proof. Recall that $\Gamma_n(z_1, z_2) = \sum_{k=1}^n \mathbb{E}_k[\mathbb{E}_{k-1}\phi_k(z_1)\mathbb{E}_{k-1}\phi_k(z_2)]$. First, we have

$$\phi_k(z) = -\frac{1+n^{-1}\mathrm{tr}D_k^2}{(z+n^{-1}\mathrm{tr}D_k)^2}g_k - b_k = \frac{\partial}{\partial z}\left(\frac{g_k}{z+n^{-1}\mathrm{tr}D_k}\right).$$

When z_z, $z_2 \in \mathbb{C}_0$ and $z_1 \neq z_2$, by exchanging the expectations and the derivation (here, the exchangibility is a consequence of the dominated convergence theorem), we obtain

$$\Gamma_n(z_1, z_2) = \frac{\partial^2}{\partial z_1 \partial z_2}\sum_{k=1}^{n}\mathbb{E}_k\left[\mathbb{E}_{k-1}\frac{g_k(z_1)}{z_1+n^{-1}\mathrm{tr}D_k(z_1)}\mathbb{E}_{k-1}\frac{g_k(z_2)}{z_2+n^{-1}\mathrm{tr}D_k(z_2)}\right]$$

and $\Gamma_n(z_1, z_1) = \lim_{z_2\to z_1}\Gamma_n(z_1, z_2)$.

As $n^{-1}\mathrm{tr}D_k(z_j) \to s(z_j)$ in L_2, we obtain by substitution and using $s^2(z) + zs(z) + 1 = 0$,

$$\Gamma_n(z_1, z_2) = \frac{\partial^2}{\partial z_1 \partial z_2}\left(s(z_1)s(z_2)\sum_{k=1}^{n}\mathbb{E}_k[\,\mathbb{E}_{k-1}g_k(z_1)\,\mathbb{E}_{k-1}g_k(z_2)] + o_p(1)\right)$$

$$:= \frac{\partial^2}{\partial z_1 \partial z_2}\big(s(z_1)s(z_2)\tilde{\Gamma}_n(z_1, z_2) + o_p(1)\big).$$

To find the limit of Γ_n, by the Vitali theorem (see Titchmarsh, 1939, p. 168), it is sufficient to find the limit of $\tilde{\Gamma}_n$.

By the definition of g_k, we have

$$\mathbb{E}_k[\mathbb{E}_{k-1}g_k(z_1)\mathbb{E}_{k-1}g_k(z_2)]$$

$$= \mathbb{E}_k\left[\frac{x_{kk}^2}{n} + \frac{1}{n^2}\,\mathbb{E}_{k-1}(\alpha_k^*D_k\alpha_k - \mathrm{tr}D_k)(z_1)\,\mathbb{E}_{k-1}(\alpha_k^*D_k\alpha_k - \mathrm{tr}D_k)(z_2).\right]$$

To evaluate the second term, by a computation similar to that leading to (3.3), we obtain

$$\mathbb{E}_k[\mathbb{E}_{k-1}(\alpha_k^* D_k \alpha_k - \mathrm{tr} D_k)(z_1)\ \mathbb{E}_{k-1}(\alpha_k^* D_k \alpha_k - \mathrm{tr} D_k)(z_2)]$$

$$= \kappa \sum_{i,j>k} [\mathbb{E}_{k-1} D_k(z_1)]_{ij} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} + \beta\ \mathbb{E}_k \sum_{i>k} [\mathbb{E}_{k-1} D_k(z_1)]_{ii} [\mathbb{E}_{k-1} D_k(z_2)]_{ii}.$$

Therefore,

$$\begin{aligned}\tilde{\Gamma}_n(z_1, z_2) &= \sigma^2 + \frac{\kappa}{n^2} \sum_{k=1}^{n} \sum_{i,j>k} [\mathbb{E}_{k-1} D_k(z_1)]_{ij} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ &\quad + \frac{\beta}{n^2} \sum_{k=1}^{n} \mathbb{E}_k \sum_{i=k+1}^{n} [\mathbb{E}_{k-1} D_k(z_1)]_{ii} [\mathbb{E}_{k-1} D_k(z_2)]_{ii} \\ &= \sigma^2 + S_1 + S_2. \end{aligned} \tag{4.8}$$

By Lemma 7.1, we find that

$$S_2 \to \frac{1}{2} \beta s(z_1) s(z_2), \qquad \text{in } L_2.$$

In the following, let us find the limit of S_1.

4.4.1. *A decomposition tool*

To evaluate the sum S_1 in (4.8), we need the following decomposition. Let e_j $(j = 1, \ldots, k-1, k+1, \ldots, n)$ be the $(n-1)$-vector whose jth (or $(j-1)$th) element is 1, the rest being 0 if $j < k$ (or $j > k$). By definition,

$$D_k^{-1}(z) = n^{-1/2} W_n(k) - zI = \sum_{i,j \neq k} n^{-1/2} x_{ij} e_i e_j' - zI.$$

Multiplying both sides by D_k gives a useful identity,

$$zD_k(z) + I = \sum_{i,j \neq k} n^{-1/2} x_{ij} e_i e_j' D_k(z). \tag{4.9}$$

Let (i, j) be two indices different from k. To make D_k 'independent' of x_{ij}, we introduce the matrix D_{kij}:

$$D_{kij} = \left(n^{-1/2} [W_n(k) - \delta_{ij} (x_{ij} e_i e_j' + x_{ji} e_j e_i')] - zI \right)^{-1}, \tag{4.10}$$

where $\delta_{ij} = 1$ for $i \neq j$ and $\delta_{ii} = \frac{1}{2}$. The idea is that D_{kij} is a perturbation of D_k independent of x_{ij}. It is easy to verify that

$$D_k - D_{kij} = -D_{kij} n^{-1/2} \delta_{ij} (x_{ij} e_i e_j' + x_{ji} e_j e_i') D_k. \tag{4.11}$$

4.4.2. The limit of S_1

From (4.9) and (4.11) we obtain

$$\begin{aligned} zD_k = & -I + \sum_{i,j\neq k} n^{-1/2} x_{ij} e_i e_j' D_{kij} - \sum_{i,j\neq k} n^{-1/2} x_{ij} e_i e_j' D_{kij} n^{-1/2} \delta_{ij} (x_{ij} e_i e_j' + x_{ji} e_j e_i') D_k \\ = & -I + \sum_{i,j\neq k} n^{-1/2} x_{ij} e_i e_j' D_{kij} - s(z) \frac{n-3/2}{n} \sum_{i\neq k} e_i e_i' D_k \\ & - \sum_{i,j\neq k} \delta_{ij} \big(n^{-1} |x_{ij}|^2 [(D_{kij})_{jj} - s(z)] + n^{-1}(|x_{ij}|^2 - 1) s(z) \big) e_i e_i' D_k \\ & - \sum_{i,j\neq k} n^{-1} \delta_{ij} x_{ij}^2 (D_{kij})_{ji} e_i e_j' D_k. \end{aligned} \tag{4.12}$$

Therefore,

$$\begin{aligned} & z_1 \sum_{i,j>k} [\mathbb{E}_{k-1} D_k(z_1)]_{ij} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ & = -\sum_{i>k} [\mathbb{E}_{k-1} D_k(z_2)]_{ii} \\ & \quad + n^{-1/2} \sum_{i,j,\ell>k} x_{i\ell} [\mathbb{E}_{k-1} D_{ki\ell}(z_1)]_{\ell,j} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ & \quad - s(z_1) \frac{n-3/2}{n} \sum_{i,j>k} [\mathbb{E}_{k-1} D_k(z_1)]_{ij} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ & \quad - \sum_{\substack{i,j>k \\ \ell\neq k}} \delta_{i\ell}\ \mathbb{E}_{k-1} \left[\left(\frac{|x_{i\ell}|^2 - 1}{n} s(z_1) + \frac{|x_{i\ell}|^2}{n} (D_{ki\ell}(z_1)_{\ell\ell} - s(z_1)) \right) [D_k(z_1)]_{ij} \right] [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ & \quad - \frac{1}{n} \sum_{\substack{i,j>k \\ \ell\neq k}} \delta_{ij}\ \mathbb{E}_{k-1} x_{i\ell}^2 [D_{ki\ell}(z_1)]_{\ell i} [D_k(z_1)]_{\ell,j} [\mathbb{E}_{k-1} D_k(z_2)]_{ji} \\ & = T_1 + T_2 + T_* + T_3 + T_4. \end{aligned}$$

First, note that the term T_* is proportional to the left-hand side. We now evaluate the contributions of the remaining four terms to the sum S_1 in (4.8). For T_1 we have

$$n^{-2} \sum_k \sum_{i>k} ([\mathbb{E}_{k-1} D_k(z_2)]_{ii} - s(z_2)) \to 0, \qquad \text{in } L_2.$$

T_3 and T_4 turn to be negligible (we do not provide more details here as the computations are lengthy but elementary). T_2 is not negligible and we simplify it progressively. We have

$$\sum_{i,j,\ell>k} \mathbb{E}_{k-1} \frac{x_{i\ell}}{\sqrt{n}} [D_{ki\ell}(z_1)]_{j\ell} [\mathbb{E}_{k-1} D_k(z_2)]_{ji}$$

$$\begin{aligned} T_2 &= \sum_{i,j,\ell>k} \mathbb{E}_{k-1} \frac{x_{i\ell}}{\sqrt{n}} [D_{ki\ell}(z_1)]_{\ell j} [\mathbb{E}_{k-1}(D_k - D_{ki\ell})(z_2)]_{ji} \\ &\quad + \sum_{i,j,\ell>k} \mathbb{E}_{k-1} \frac{x_{i\ell}}{\sqrt{n}} [D_{ki\ell}(z_1)]_{\ell j} [E_{k-1} D_{ki\ell}(z_2)]_{ji}. \\ &= T_{2\mathrm{a}} + T_{2\mathrm{b}}. \end{aligned}$$

Again the contribution from $T_{2\mathrm{b}}$ can be proved to be negligible. As for the remaining term $T_{2\mathrm{a}}$, we have

$$\begin{aligned} &n^{-1/2} \sum_{i,j,\ell>k} x_{i\ell} [\mathbb{E}_{k-1} D_{ki\ell}(z_1)]_{\ell j} [\mathbb{E}_{k-1}(D_k - D_{ki\ell})(z_2)]_{ji} \\ &= -n^{-1} \sum_{i,j,\ell>k} [\mathbb{E}_{k-1} D_{ki\ell}(z_1)]_{\ell j} [\mathbb{E}_{k-1} D_{ki\ell}(z_2) \delta_{i\ell} (x_{i\ell}^2 e_i e_\ell' + |x_{ij}|^2 e_\ell e_i') D_k(z_2)]_{ji} \\ &:= J_1 + J_2, \end{aligned}$$

where

$$\begin{aligned} \mathbb{E}|J_1| &= n^{-1} \sum_{i,j,\ell>k} \mathbb{E}|x_{i\ell}^2| [\mathbb{E}_{k-1} D_{ki\ell}(z_1)]_{\ell j} [D_{ki\ell}(z_2)]_{ji} [D_k(z_2)]_{\ell i}| \\ &\leqslant n^{-1} \left(\sum_{i,j_1,j_2,\ell>k} \mathbb{E}|x_{i\ell}|^4 |[\mathbb{E}_{k-1} D_{ki\ell}(z_1)]_{\ell j_1}|^2 |[D_{ki\ell}(z_2)]_{j_2 i}|^2 | \sum_{i,\ell>k} \mathbb{E}|(D_k(z_2))_{\ell i}|^2 \right)^{1/2} \\ &\leqslant K n^{1/2}. \end{aligned}$$

Hence, the contribution of this term is negligible. Finally, we have

$$\begin{aligned} J_2 &= -\sum_{i,j,\ell>k} \mathbb{E}_{k-1} \frac{|x_{i\ell}|^2}{n} [D_{kij}(z_1)]_{\ell j} [\mathbb{E}_{k-1} D_{kij}(z_2)]_{j\ell} [D_k(z_2)]_{ii} \\ &\simeq -s(z_2) \sum_{i,j,\ell>k} \mathbb{E}_{k-1} \frac{1}{n} [D_{ki\ell}(z_1)]_{\ell j} [\mathbb{E}_{k-1} D_{ki\ell}(z_2)]_{j\ell} \\ &= -\frac{n-k}{n} s(z_2) \sum_{j,\ell>k} [\mathbb{E}_{k-1} D_k(z_1)]_{\ell j} [\mathbb{E}_{k-1} D_k(z_2)]_{j\ell} + o_p(n), \end{aligned}$$

where the last approximation follows from

$$[D_{ki\ell}(z_1)]_{\ell j}[\mathbb{E}_{k-1}D_{ki\ell}(z_2)]_{j\ell} = [\mathbb{E}_{k-1}D_k(z_1)]_{\ell j}[\mathbb{E}_{k-1}D_k(z_2)]_{j\ell} + o_p(1).$$

Summing the estimates of T_i, $i = 1, \ldots, 4$, we have proved that

$$\begin{aligned} z_1 \sum_{i,j>k} &[\mathbb{E}_{k-1}D_k(z_1)]_{ij}[\mathbb{E}_{k-1}D_k(z_2)]_{ji} \\ &= -s(z_1)\sum_{i,j>k}[\mathbb{E}_{k-1}D_k(z_1)]_{ij}[\mathbb{E}_{k-1}D_k(z_2)]_{ji} - (n-k)s(z_2) \\ &\quad - \frac{n-k}{n}s(z_2)\sum_{i,j>k}[\mathbb{E}_{k-1}D_k(z_1)]_{ij}[\mathbb{E}_{k-1}D_k(z_2)]_{ji} + R_k, \end{aligned} \tag{4.13}$$

where the residual term R_k is of order $o_p(n)$ uniformly in $k = 1, \ldots, n$. Let us define

$$X_k = \sum_{i,j>k}[\mathbb{E}_{k-1}D_k(z_1)]_{ij}[\mathbb{E}_{k-1}D_k(z_2)]_{ji}.$$

So (4.13) becomes

$$z_1 X_k = -s(z_1)X_k - (n-k)s(z_2) - s(z_2)X_k\frac{n-k}{n} + R_k. \tag{4.14}$$

By $z_1 + s(z_1) = -1/s(z_1)$, identity (4.14) is equivalent to

$$X_k = (n-k)s(z_1)s(z_2) + \frac{n-k}{n}s(z_1)s(z_2)X_k - s(z_1)R_k. \tag{4.15}$$

Consequently,

$$X_k = \frac{(n-k)s(z_1)s(z_2) - s(z_1)R_k}{1 - n^{-1}(n-k)s(z_1)s(z_2)}.$$

Summing and letting $n \to \infty$ yields

$$\frac{1}{n^2}\sum_{k=1}^{n} X_k \xrightarrow{p} s_1 s_2 \int_0^1 \frac{t}{1 - t s_1 s_2}\,\mathrm{d}t = -1 - (s_1 s_2)^{-1}\log(1 - s_1 s_2),$$

with $s_j = s(z_j)$. Finally, $\tilde{\Gamma}_n(z_1, z_2)$ converges in probability to

$$\tilde{\Gamma}(z_1, z_2) = \sigma^2 - \kappa + \frac{1}{2}\beta s_1 s_2 - \kappa(s_1 s_2)^{-1}\log(1 - s_1 s_2).$$

The proof of Proposition 4.1 is then complete.

5. Computation of the mean and covariance function of $G(f)$

5.1. The mean

Let γ be a contour as defined in Section 1. By (2.1) and Proposition 3.1, we have

$$\begin{aligned}\mathbb{E}(G_n(f)) &= -\frac{1}{2\pi i}\oint_\gamma f(z)\mathbb{E}\xi_n(z)dz \\ \to \mathbb{E}(G(f)) &= -\frac{1}{2\pi i}\oint_\gamma f(z)\mathbb{E}\xi_n(z)dz \\ &= -\frac{1}{2\pi i}\oint_\gamma f(z)[1+s'(z)]s^3(z)[\sigma^2-1+(\kappa-1)s'(z)+\beta s^2]dz.\end{aligned}$$

Select $\rho < 1$ but so close to 1 that the contour

$$\gamma' = \{z = -(\rho e^{i\theta} + \rho^{-1}e^{-i\theta}) : 0 \leqslant \theta < 2\pi\}$$

is completely contained in the analytic region of f. Note that when z runs a cycle along γ' counterclockwise, s runs a cycle along the circle $|s| = \rho$ counterclockwise because $z = -(s+s^{-1})$.[1] By Cauchy's theorem, the above integral along γ equals the integral along γ'. Thus, by changing variable z to s and noticing that $s' = s^2/(1-s^2)$, we obtain

$$\mathbb{E}(G(f)) = -\frac{1}{2\pi i}\oint_{|s|=\rho} f(-s-s^{-1})s\left[\sigma^2 - 1 + (\kappa-1)\frac{s^2}{1-s^2} + \beta s^2\right]ds$$

By setting $s = -e^{i\theta}$ and then $t = \cos\theta$, using $T_k(\cos\theta) = \cos(k\theta)$,

$$\begin{aligned}&-\frac{1}{2\pi i}\oint_{|s|=1} f(-s-s^{-1})s\left[\sigma^2-1+(\kappa-1)\frac{s^2}{1-s^2}+\beta s^2\right]ds \\ &= -\frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos\theta)\left[(\sigma^2-1)e^{2i\theta}+(\kappa-1)\frac{e^{Ai\theta}}{1-e^{2i\theta}}+\beta e^{Ai\theta}\right]d\theta \\ &= -\frac{1}{\pi}\int_0^{\pi} f(2\cos\theta)\left[(\sigma^2-1)\cos 2\theta - \frac{1}{2}(\kappa-1)(1+2\cos 2\theta)+\beta\cos 4\theta\right]d\theta \\ &= \frac{1}{\pi}\int_{-1}^{1} f(2t)\left[-\frac{1}{2}(\kappa-1)+(\sigma^2-\kappa)T_2(t)+\beta T_4(t)\right]\frac{1}{\sqrt{1-t^2}}dt \\ &= -\frac{1}{2}(\kappa-1)\tau_0(f)+(\sigma^2-\kappa)\tau_2(f)+\beta\tau_4(f).\end{aligned}$$

Let us evaluate the difference

[1]The reason for choosing $|s| = \rho < 1$ is that the mode of the Stieltjes transform of the semicircle law is less than 1; see Bai (1993).

$$\frac{1}{2\pi i}\left[\oint_{|s|=1}-\oint_{|s|=\rho}\right]f(-s-s^{-1})s\left[\sigma^2-1+(\kappa-1)\frac{s^2}{1-s^2}+\beta s^2\right]ds.$$

Note that the integrand has two poles on the circle $|s|=1$ with residuals $-\frac{1}{2}f(\pm 2)$ at points $s=\mp 1$. By contour integration, we have

$$\frac{1}{2\pi i}\left[\oint_{|s|=1}-\oint_{|s|=\rho}\right]f(-s-s^{-1})s\left[\sigma^2-1+(\kappa-1)\frac{s^2}{1-s^2}+\beta s^2\right]ds$$
$$=\frac{\kappa-1}{4}(f(2)+f(-2)).$$

Putting together these two results gives formula (1.4) for $\mathbb{E}[G(f)]$.

5.2. The covariance

Let γ_j, $j=1,2$, be two disjoint contours with vertex $\pm(2+\varepsilon_j)\pm i v_j$. The positive values of ε_j and v_j are chosen so small that the two contours are contained in $\mathcal{U}$. By (2.1) and Theorem 2.1, we have

$$\begin{aligned}\operatorname{cov}(G_n(f),G_n(g))&=-\frac{1}{4\pi^2}\oint_{\gamma_1}\oint_{\gamma_2}f(z_1)g(z_2)\operatorname{cov}(\xi_n(z_1),\xi_n(z_2))dz_1\,dz_2\\&=-\frac{1}{4\pi^2}\oint_{\gamma_1}\oint_{\gamma_2}f(z_1)g(z_2)\Gamma_n(z_1,z_2)dz_1\,dz_2+o(1)\\\to c(f,g)&=-\frac{1}{4\pi^2}\oint_{\gamma_1}\oint_{\gamma_2}f(z_1)g(z_2)\Gamma(z_1,z_2)dz_1\,dz_2,\end{aligned}$$

where $\Gamma(z_1,z_2)$ is given by (4.5).

By the proof of Proposition 4.1, we have

$$\Gamma(z_1,z_2)=\frac{\partial^2}{\partial z_1\partial z_2}s(z_1)s(z_2)\tilde{\Gamma}(z_1,z_2).$$

Integrating by parts, we obtain

$$\begin{aligned}c(f,g)&=-\frac{1}{4\pi^2}\oint_{\gamma_1}\oint_{\gamma_2}f'(z_1)g'(z_2)s(z_1)s(z_2)\tilde{\Gamma}(z_1,z_2)dz_1\,dz_2\\&=-\frac{1}{4\pi^2}\oint_{\gamma_1}\oint_{\gamma_2}A(z_1,z_2)dz_1\,dz_2,\end{aligned}$$

where

$$A(z_1,z_2)=f'(z_1)g'(z_2)\left[s(z_1)s(z_2)(\sigma^2-\kappa)+\frac{1}{2}\beta s^2(z_1)s^2(z_2)-\kappa\log(1-s(z_1)s(z_2))\right].$$

Let $v_j \to 0$ first and then $\varepsilon_j \to 0$. It is easy to show that the integral along the vertical edges of the two contours tends to 0 when $v_j \to 0$. Therefore, it follows that

$$c(f, g) = -\frac{1}{4\pi^2}\int_{-2}^{2}\int_{-2}^{2}[A(t_1^-, t_2^-) - A(t_1^-, t_2^+) - A(t_1^+, t_2^-) + A(t_1^+, t_2^+)]\mathrm{d}t_1\,\mathrm{d}t_2,$$

where $t_j^{\pm} := t_j \pm \mathrm{i}0$. We first consider the case where f and g are real-valued on the real axes. Recalling that $s(t \pm \mathrm{i}0) = \frac{1}{2}(-t \pm \mathrm{i}\sqrt{4 - t^2})$, we have

$$\begin{aligned}
&f'(t_1)g'(t_2)[s(t_1^-)s(t_2^-) - s(t_1^+)s(t_2^-) - s(t_1^-)s(t_2^+) + s(t_1^+)s(t_2^+)]\\
&\quad = -f'(t_1)g'(t_2)\sqrt{4 - t_1^2}\sqrt{4 - t_2^2},\\
&f'(t_1)g'(t_2)[s^2(t_1^-)s^2(t_2^-) - s^2(t_1^+)s^2(t_2^-) - s^2(t_1^-)s^2(t_2^+) + s^2(t_1^+)s^2(t_2^+)]\\
&\quad = -f'(t_1)g'(t_2)t_1t_2\sqrt{4 - t_1^2}\sqrt{4 - t_2^2},\\
&f'(t_1)g'(t_2)[\log(1 - s(t_1^-)s(t_2^-))\\
&\qquad - \log(1 - s(t_1^+)s(t_2^-)) - \log(1 - s(t_1^-)s(t_2^+)) + \log(1 - s(t_1^+)s(t_2^+))]\\
&\quad = f'(t_1)g'(t_2)\log\left|\frac{1 - s(t_1^-)s(t_2^-)}{1 - s(t_1^-)s(t_2^+)}\right|^2 = -f'(t_1)g'(t_2)\log\left(\frac{4 - t_1t_2 - \sqrt{(4 - t_1^2)(4 - t_2^2)}}{4 - t_1t_2 + \sqrt{(4 - t_1^2)(4 - t_2^2)}}\right).
\end{aligned}$$

We thus recover formula (1.6). Moreover, if f and g can take complex values on the real axes, the above expression remains true since we have the decomposition $f(z) = f_{\mathrm{r}}(z) + \mathrm{i}f_{\mathrm{i}}(z)$ and $g(z) = g_{\mathrm{r}}(z) + \mathrm{i}g_{\mathrm{i}}(z)$ where f_{r}, f_{i}, g_{r} and g_{i} are analytic on $\mathcal{U}$ and are real-valued on the real axes.

To derive the second representation, formula (1.5), let $\rho_1 < \rho_2 < 1$ and define contours γ_j' as in the previous subsection. Then

$$\begin{aligned}
c(f, g) &= -\frac{1}{4\pi^2}\oint_{\gamma_1'}\oint_{\gamma_2'} f(z_1)g(z_2)\Gamma(z_1, z_2)\mathrm{d}z_1\,\mathrm{d}z_2\\
&= -\frac{1}{4\pi^2}\oint_{|s_1|=\rho_1}\oint_{|s_2|=\rho_2} f(-s_1 - s_1^{-1})g(-s_2 - s_2^{-1})\left(\sigma^2 - \kappa + 2\beta s_1s_2 + \frac{\kappa}{(1 - s_1s_2)^2}\right)\mathrm{d}s_1\,\mathrm{d}s_2.
\end{aligned}$$

By the Cauchy integral, we may set $\rho_2 = 1$ without changing the value of the integral. Rewriting $\rho_1 = \rho$, expanding the fraction with a Taylor series and then making the changes of variable $s_1 = -\rho\mathrm{e}^{\mathrm{i}\theta_1}$ and $s_2 = -\mathrm{e}^{\mathrm{i}\theta_2}$, we obtain

$$c(f, g) = \frac{1}{4\pi^2}\int_{[-\pi,\pi]^2} f(\rho e^{i\theta_1} + \rho^{-1}e^{-i\theta_1})g(2\cos\theta_2)\Big[(\sigma^2 - \kappa)\rho e^{i(\theta_1+\theta_2)}$$

$$+ 2\beta\rho^2 e^{i2(\theta_1+\theta_2)} + \kappa\sum_{k=1}^{\infty} k\rho^k e^{ik(\theta_1+\theta_2)}\Big] d\theta_1\, d\theta_2$$

$$= (\sigma^2 - \kappa)\rho\tau_1(f, \rho)\tau_1(g) + 2\beta\rho^2\tau_2(f, \rho)\tau_2(g) + \kappa\sum_{k=1}^{\infty} k\rho^k\tau_k(f, \rho)\tau_k(g),$$

where

$$\tau_k(f, \rho) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\rho e^{i\theta} + \rho^{-1}e^{-i\theta})e^{ik\theta}\, d\theta.$$

By integration by parts, for $k \geqslant 3$ we have

$$\tau_k(f, \rho) = \frac{\rho^{-1}}{k}\tau_{k-1}(f', \rho) - \frac{\rho}{k}\tau_{k+1}(f', \rho)$$

$$= \frac{\rho^2}{k(k+1)}\tau_{k+2}(f'', \rho) - \frac{2}{k^2-1}\tau_k(f'', \rho) + \frac{\rho^{-2}}{k(k-1)}\tau_{k-2}(f'', \rho).$$

Since f'' is uniformly bounded in $\mathcal{U}$, we have $|\tau_k(f, \rho)| \leqslant K/k(k-1)$ uniformly for all ρ close to 1. Then (1.5) follows by the dominated convergence theorem and letting $\rho \to 1$ under the summation.

6. Application to linear spectral statistics and related results

First, note that $W_n/(2\sqrt{n})$ is a scaled Wigner matrix in the sense that the limit law is the scaled Wigner semicircular law $2\pi^{-1}\sqrt{1-x^2}\,dx$ on the interval $[-1, 1]$. To deal with this scaling, we define for any function f, its scaled copy $\tilde{f}$ by the relation $f(2x) = \tilde{f}(x)$ for all x.

6.1. Chebyshev polynomials

Consider first a Chebyshev polynomial T_k with $k \geqslant 1$ and define ϕ_k such that $\tilde{\phi}_k = T_k$. Set $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ otherwise. Using the orthogonality property

$$\frac{1}{\pi}\int_{-1}^{1} T_i(t)T_j(t)\frac{1}{\sqrt{1-t^2}}\, dt = \begin{cases} \delta_{ij}, & \text{if } i = 0, \\ \frac{1}{2}\delta_{ij}, & \text{otherwise,} \end{cases}$$

it is easily seen that $\tau_\ell(\phi_k) = \frac{1}{2}\delta_{k\ell}$ for any integer $\ell \geqslant 0$. Thus by (1.4) we have for the mean

$$m_k := \mathbb{E}[G(\phi_k)] = \frac{\kappa - 1}{4}(T_k(1) + T_k(-1)) + \frac{1}{2}(\sigma^2 - \kappa)\delta_{k2} + \frac{1}{2}\beta\delta_{k4}$$
$$= \frac{1}{2}[(\kappa - 1)e(k) + (\sigma^2 - \kappa)\delta_{k2} + \beta\delta_{k4}], \tag{6.1}$$

with $e(k) = 1$ if k is even and $e(k) = 0$ otherwise.

For two Chebyshev polynomials T_k and T_ℓ, by (1.5) the asymptotic covariance between $G_n(\phi_k)$ and $G_n(\phi_\ell)$ equals 0 for $k \neq \ell$, and

$$\Sigma_{\ell\ell} = \left(\frac{1}{2}\right)^2 \left[(\sigma^2 - \kappa)\delta_{\ell 1} + 2\beta\delta_{\ell 2} + \kappa\ell\right] \tag{6.2}$$

for $k = \ell$. Application of Theorem 1.1 readily yields the following result.

Corollary 6.1. *Assume that Conditions 1.1 and 1.2 are satisfied. Let $T_1, \ldots, T_p$ be p first Chebyshev polynomials and define the ϕ_k such that $\tilde{\phi}_k = T_k$. Then the vector $[G_n(\phi_1), \ldots, G_n(\phi_p)]$ converges in distribution to a Gaussian vector with mean $w_p = (m_k)$ and a diagonal covariance matrix $D_p = (\Sigma_{kk})$ with their elements defined in* (6.1) *and* (6.2)*, respectively.*

In particular, these Chebyshev polynomial statistics are asymptotically independent. Now consider the Gaussian case. For the GUE, we have $\kappa = \sigma^2 = 1$ and $\beta = 0$. Then $m_k = 0$ (already known!) and $\Sigma_{kk} = k(\frac{1}{2})^2$. As for the GOE, since $\kappa = \sigma^2 = 2$ and $\beta = 0$, we obtain $m_k = \frac{1}{2}e(k)$ and $\Sigma_{kk} = 2k(\frac{1}{2})^2$. Therefore with Corollary 6.1 we have recovered the CLT established by Johansson for linear spectral statistics of Gaussian ensembles (see Johansson 1998, Theorem 2.4 and Corollary 2.8).

6.2. Linear spectral statistics

Let Q be an arbitrary polynomial of degree k and define ϕ be such that $\tilde{\phi} = Q$. Then Q has an unique expansion on the basis of Chebyshev polynomials:

$$Q = a_k(Q)T_k + \ldots + a_0(Q)T_0.$$

Furthermore, it is easily seen that $a_j(Q) = 2\tau_j(\phi)$ for $j \geqslant 1$ and $a_0(Q) = \tau_0(\phi)$.

Now consider s polynomials $Q_1, \ldots, Q_s$ and denote by p the maximum of their degrees. Define the ϕ_k such that $\tilde{\phi}_k = Q_k$. There is a unique matrix decomposition

$$\begin{pmatrix} Q_1 \\ \vdots \\ Q_s \end{pmatrix} = A \begin{pmatrix} T_p \\ \vdots \\ T_1 \end{pmatrix} + u\,T_0,$$

where A is an $s \times p$ matrix and u an $s \times 1$ vector. Application of Corollary 6.1 gives the following result.

Corollary 6.2. *Assume that Conditions 1.1 and 1.2 are satisfied. Then the vector*

$[G_n(\phi_1), \ldots, G_n(\phi_p)]$ converges in distribution to a Gaussian vector with mean $Aw_p + u$ and covariance matrix AD_pA^{T}, where w_p and D_p are defined in Corollary 6.1 for Chebyshev polynomials.

7. Two useful lemmas

With the notation defined in the previous sections, we prove the following two lemmas.

Lemma 7.1. *Suppose $v_0 > 0$ is a fixed constant. Then, for any $z \in \mathbb{C}_0$, we have*

$$\max_{i,j,k,\ell} |\mathbb{E}_k(D_{kij})_{\ell\ell} - s(z)| \to 0, \qquad \text{in } L_2,$$

where the maximum is taken over all k, i, $j \neq k$ and all ℓ.

Proof. Recall the identity

$$D_{kij}^{-1} = D_k^{-1} - n^{-1/2}\delta_{ij}(x_{ij}e_ie_j' + x_{ji}e_je_i').$$

Since $|(D_k^{-1})_{\ell\ell}| \geqslant v > 0$ and $|x_{ij}| \leqslant \eta_n\sqrt{n}$, we have $|(D_{kij})_{\ell\ell} - (D_k)_{\ell\ell}| \leqslant K\eta_n$ for some constant K independent of the indices. On the other hand, we know that $(D_k)_{\ell\ell} = (D)_{\ell\ell} + O(n^{-1})$. Moreover, by definition,

$$\begin{aligned}(D)_{\ell\ell} &= \frac{1}{n^{-1/2}x_{\ell\ell} - z - n^{-1}\alpha_\ell^* D_\ell \alpha_\ell} \\ &= \frac{1}{-z - s(z)} + \frac{-s(z) - [n^{-1/2}x_{\ell\ell} - n^{-1}\alpha_\ell^* D_\ell\alpha_\ell]}{(n^{-1/2}x_{\ell\ell} - z - n^{-1}\alpha_\ell^* D_\ell\alpha_\ell)[-z - s(z)]}.\end{aligned}$$

By writing

$$|n^{-1}\alpha_\ell^* D_\ell\alpha_\ell - s(z)| \leqslant n^{-1}|\alpha_\ell^* D_\ell\alpha_\ell - \mathbb{E}\mathrm{tr}D_\ell| + |n^{-1}\,\mathbb{E}\mathrm{tr}D_\ell - s(z)|,$$

for any $z \in \mathbb{C}_0$, the first term here converges to zero in L_2 (see (2.15)) and the second also tends to 0.

Noting that $|-z-s(z)|^{-1} = |s(z)|$ has a positive lower bound, as does $|n^{-1/2}x_{\ell\ell} - z - n^{-1}\alpha_\ell^* D_\ell\alpha_\ell|$, we obtain

$$|(D)_{\ell\ell} - s(z)| \leqslant K(\eta_n + |n^{-1}\alpha_\ell^* D_\ell\alpha_\ell - s(z)|).$$

The conclusion follows from the fact that $s(z)^2 + zs(z) + 1 = 0$. □

Lemma 7.2. *Suppose that $\mathbb{E}x_i = 0$, $\mathbb{E}|x_i|^2 = 1$, $\sup E|x_i|^4 < \infty$ and $|x_i| \leqslant \eta_n\sqrt{n}$ with $\eta_n \to 0$ slowly. Assume that A is a Hermitian matrix of order n bounded in norm by M. Then, for any given $2 \leqslant p \leqslant b\log(n\eta_n^2)$* with some $b > 1$, *there exists a constant K such that*

$$\mathbb{E}|u^* Au - \mathrm{tr}(A)|^p \leqslant n^p(n\eta_n^4)^{-1}(MK\eta_n^2)^p,$$

where $u = (x_1, \cdots, x_n)^{\mathrm{T}}$.

Proof. Without loss of generality, we may assume that $p = 2s$ is an even integer. Write $A = (a_{ij})$. We first consider

$$S_1 = \sum_{i=1} a_{ii}(|x_{\rm i}|^2 - 1) = \sum_{i=1}^{n} a_{ii}\xi_{\rm i}.$$

By noticing that $|a_{ii}| \leqslant M$, we have, for $p/\log(n\eta_n^4) \geqslant 1$,

$$\begin{aligned}
\mathbb{E}|S_1|^p &\leqslant \sum_{\ell=1}^{s} \sum_{1\leqslant j_1<\ldots<j_\ell\leqslant n} \sum_{\substack{i_1+\ldots+i_\ell=p\\ i_1,\ldots,i_\ell\geqslant 2}} p! \prod_{t=1}^{\ell} \frac{|a_{j_t j_t}^{i_t}|\mathrm{E}|\xi_{j_t}^{i_t}|}{(i_t)!} \\
&\leqslant M^p \sum_{\ell=1}^{s} n^\ell \sum_{\substack{i_1+\ldots+i_\ell=p\\ i_1,\ldots,i_\ell\geqslant 2}} (n\eta_n^2)^{p-2\ell} \frac{p!}{(i_1)!\cdots(i_\ell)!} \\
&\leqslant n^p(M\eta_n^2)^p \sum_{\ell=1}^{s} (n\eta_n^4)^{-\ell}\ell^p \\
&\leqslant s(M\eta_n^2)^p(n\eta_n^4)^{-p/\log(n\eta_n^4)}(p/\log(n\eta_n^4))^p \\
&\leqslant n^p(MK\eta_n^2)^p(n\eta_n^4)^{-1}.
\end{aligned}$$

Here we have used the facts that $a^\ell \ell^p$ is maximum for $\ell = -p/\log a$ if $a \in (0, 1)$, $p \geqslant 2$ and $s^{1/p}p/2\log(n\eta_n^2) \leqslant K$ as $p \leqslant b\log n$. Note that the last inequality is still true when $p/\log(n\eta_n^4) < 1$ since $(n\eta_n^4)^{-\ell}\ell^p \leqslant (n\eta_n^4)^{-1}$.

Next, let us consider

$$S_2 = \sum_{1\leqslant i\neq j\leqslant n} a_{ij}x_{\rm i}\bar{x}_j.$$

Then we have

$$\mathbb{E}|S_2|^p = \sum a_{i_1 j_1}\bar{a}_{k_1,\ell_1}\cdots a_{i_s j_s}\bar{a}_{k_s,\ell_s}\mathrm{E}x_{i_1}\bar{x}_{k_1}\bar{x}_{j_1}x_{\ell_1}\cdots x_{i_s}\bar{x}_{k_s}\bar{x}_{j_s}x_{\ell_s}.$$

Draw a directional graph G of $p = 2s$ edges which link i_t to j_t and ℓ_t to k_t, $t = 1, \ldots, s$. Note that if G has a vertex of degree 1, then the graph corresponds to a term with expectation 0. That is, for the non-zero term, the vertices of the graph have degree either 0 or greater than 1. For vertices of non-zero degree, denote by $p_1, \ldots, p_m$ the degree of vertex $v_1, \ldots, v_m$. We have $m \leqslant s$. By assumption,

$$|\mathrm{E}x_{i_1}\bar{x}_{k_1}\bar{x}_{j_1}x_{\ell_1}\cdots x_{i_s}\bar{x}_{k_s}\bar{x}_{j_s}x_{\ell_s}| \leqslant K(n\eta_n^2)^{p-m-\phi/2},$$

where ϕ is the number of vertices of degree 3 plus twice the number of vertices of degree greater than 3. Now, suppose that the graph can be split into q disjoint connected subgraphs $G_1, \ldots, G_q$ with $m_1, \ldots, m_q$ vertices, respectively. As an example, consider the contribution by G_1 to $\mathbb{E}|S_2|^p$. Assume that G_1 has s_1 edges, $e_1, \ldots, e_{s_1}$. Choose tree subgraph G_1' of G_1 and assume its edges are $e_1, \ldots, e_{m_1-1}$, without loss of generality. Note that

$$\sum_{v_1,\ldots,v_{m_1}\leqslant n}\prod_{t=1}^{m_1-1}|a_{e_t}|^2 \leqslant M^{2m_1-2}n$$

and

$$\sum_{v_1,\ldots,v_{m_1}\leqslant n}\prod_{t=m_1}^{s_1}|a_{e_t}|^2 \leqslant M^{2s_1-2m_1+2}n^{m_1-1}.$$

Here the first inequality follows from the fact that $\sum_{v_1}|a_{v_1v_2}|^2 \leqslant M^2$ since it is a diagonal element of AA^*. The second inequality follows from the facts that $\sum_{v_1}|a_{v_1v_2}|^\ell \leqslant M^\ell$, for any $\ell \geqslant 2$, and that $s_1 \geqslant m_1$ since all vertices have degree no less than 2. Therefore, the contribution of G_1 is bounded by

$$\sum_{v_1,\ldots,v_{m_1}\leqslant n}\prod_{t=1}^{s_1}|a_{e_t}| \leqslant \left(\sum_{v_1,\ldots,v_{m_1}\leqslant n}\prod_{t=1}^{m_1-1}|a_{e_t}|^2\sum_{v_1,\ldots,v_{m_1}\leqslant n}\prod_{t=m_1}^{s_1}|a_{e_t}|^2\right)^{1/2} \leqslant M^{s_1}n^{m_1/2}.$$

Noticing that $m_1+\ldots+m_q = m$ and $s_1+\ldots+s_q = s$, we eventually obtain the contribution from all graphs isomorphic to a given in-homogeneous subgraph, which is $M^{2s}n^{m/2}$. Because the two vertices of each edge are not equal, we have $q \leqslant m/2$. The number of inhomogeneous graphs is less than $\binom{m}{2}^p \leqslant m^{2p}$. We finally obtain

$$\begin{aligned}\mathbb{E}|S_2|^p &\leqslant M^{2s}\sum_{m=2}^{s} n^{m/2}(n\eta_n^2)^{p-m-\phi/2}m^{2p}\\ &\leqslant n^p(M\eta_n^2)^p\sum_{m=2}^{s}(n^{1/2}\eta_n^2)^{-m}(n\eta_n^2)^{-\phi/2}m^{2p} \leqslant n^p(n\eta_n^4)^{-1}(MK\eta_n^2)^p.\end{aligned}$$

Combining the estimates of $\mathbb{E}|S_1|^p$ and $\mathbb{E}|S_2|^p$, the proof of of the lemma is complete. □

Corollary 7.3. *For any positive constants ν and t, when $z \in \mathbb{C}_0$, all the following probabilities have order $o(n^{-t})$:*

$$\mathbb{P}(|\varepsilon_k| \geqslant \nu), \qquad \mathbb{P}(|\tilde{\varepsilon}_k| \geqslant \nu), \qquad \mathbb{P}(|g_k| \geqslant \nu), \qquad \mathbb{P}(|h_k| \geqslant \nu).$$

When $z \notin \mathbb{C}_0$ but $|\Re(z)| \geqslant a$, the same estimates remain true.

Proof. Recalling the definition of ε_k, we have

$$\begin{aligned}|\varepsilon_k| &= \left|n^{-1/2}x_{kk} - \frac{1}{n}\alpha_k^* D_k\alpha_k + \mathbb{E}s_n(z)\right|\\ &\leqslant \eta_n + n^{-1}|\alpha_k^* D_k\alpha_k - \mathrm{tr}(D_k)| + n^{-1}|\mathrm{tr}(D_k) - \mathbb{E}\mathrm{tr}(D)|.\end{aligned}$$

By the decomposition (4.1), the Burkholder inequality (Burkholder 1973) and the fact that $z \in \mathbb{C}_0$, we have

$$\mathbb{P}(n^{-1}|\mathrm{tr}(D) - \mathbb{E}(\mathrm{tr}(D))| \geq \nu/3) \leq 3^p \nu^{-p} n^{-p}\, \mathbb{E}|\mathrm{tr}(D) - \mathbb{E}(\mathrm{tr}(D))|^p$$
$$\leq 3^p \nu^{-p} K_p n^{-p}\, \mathbb{E}\left(\sum_{k=1}^{n} \mathrm{E}_k(u_k)^2\right)^{p/2} \leq 3^p \nu^{-p} K_p n^{-p/2} v_0^{-p}. \tag{7.3}$$

Choosing p fixed and $p > 2t$, we then have

$$\mathbb{P}\{n^{-1}|\mathrm{tr}(D) - \mathbb{E}(\mathrm{tr}(D))| \geq \nu\} = o(n^{-t}).$$

When $z \in \mathbb{C}_0$, the norm of D_k is bounded by $1/v_0$. Thus, by Lemma 7.2 with $p = [\log n]$, we have

$$\mathbb{P}\{n^{-1}|\alpha_k^* D_k \alpha_k - \mathrm{tr}(D_k)| \geq \nu/3\} \leq 3^p \nu^{-p}\, \mathbb{E}\{n^{-1}|\alpha_k^* D_k \alpha_k - \mathrm{tr}(D_k)|^p\}$$
$$\leq (3\nu^{-1} MK\eta_n)^p = o(n^{-t}). \tag{7.4}$$

Therefore $\mathbb{P}(|\varepsilon_k| \geq \nu) = o(n^{-t})$ since $\eta_n \to 0$ and $n^{-1}|\mathrm{tr}(D) - \mathrm{tr}(D_k)| \leq /nv$.

When $z \notin \mathbb{C}_0$ but $\Re(z)| \geq a$, the eigenvalues of D_k are bounded by $2/(a-2)$ when B_{nk}^c happens. Thus, inequalities (7.3) and (7.4) can be modified as

$$\mathbb{P}\{n^{-1}|\mathrm{tr}(D) - \mathbb{E}\mathrm{tr}(D)| \geq \nu/3,\ B_n^c\} \leq \mathbb{P}\left\{n^{-1}\left|\sum_{k=1}^{n}(\mathbb{E}_{k-1} - \mathbb{E}_k)u_k \mathbb{1}_{B_{nk}^c}\right| \geq \nu/3\right\} = o(n^{-t})$$

and

$$\mathbb{P}\{n^{-1}|\alpha_k^* D_k \alpha_k - \mathrm{tr}(D_k)| \geq \nu/3,\ B_{nk}^c\} = o(n^{-t}).$$

This also proves $\mathbb{P}(|\varepsilon_k| \geq \nu/3) = o(n^{-t})$ since $\mathbb{P}(B_n) = o(n^{-t})$.

The proofs of the other probabilities are similar and hence omitted. The proof of the lemma is complete. □

Acknowledgements

The research of the first author was supported in part by grant R-155-000-014-112 from the National University of Singapore. He is grateful to the research institute IRMAR for a scientific visit which is at the origin of the present work.

References

Bai, Z. (1993) Convergence rate of expected spectral distributions of large random matrices. Part I. Wigner matrices. *Ann. Probab.*, **21**, 625–648.

Bai, Z. (1999) Methodologies in spectral analysis of large dimensional random matrices. A review. *Statist. Sinica*, **9**, 611–677.

Bai, Z. and Yin, Y. (1988) Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.*, **16**, 1729–1741.

Bai, Z.D., Miao, B. and Tsay, J. (2002) Convergence rates of the spectral distributions of large Wigner matrices. *Int. Math. J.*, **1**(1), 65–90.

Burkholder, D. (1973) Distribution function inequalities for martingales. *Ann. Probab.*, **1**, 19–42.

Costin, O. and Lebowitz, J. (1995) Gaussian fluctuations in random matrices. *Phys. Rev. Lett.*, **75**, 69–72.

Johansson, K. (1998) On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, **91**, 151–204.

Khorunzhy, A.M., Khoruzhenko, B.A. and Pastur, L.A. (1996) Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.*, **37**, 5033–5060.

Rao, C.R. and Rao, M.B. (2001) *Matrix Algebra and Its Applications to Statistics and Econometrics*. River Edge, NJ: World Scientific.

Sinai, Y. and Soshnikov, A. (1998) Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat. (N.S.)*, **29**, 1–24.

Titchmarsh, E. (1939) *The Theory of Functions* (2nd edn). London: Oxford University Press.

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

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

Received November 2003 and revised March 2005