The Annals of Probability
1997, Vol. 25, No. 1, 494–529

CIRCULAR LAW

By Z. D. Bai[1]

National Sun Yat-sen University

It was conjectured in the early 1950's that the empirical spectral distribution of an $n \times n$ matrix, of iid entries, normalized by a factor of $1/\sqrt{n}$, converges to the uniform distribution over the unit disc on the complex plane, which is called the circular law. Only a special case of the conjecture, where the entries of the matrix are standard complex Gaussian, is known. In this paper, this conjecture is proved under the existence of the sixth moment and some smoothness conditions. Some extensions and discussions are also presented.

1. Introduction. Suppose that Ξ_n is an $n \times n$ matrix with entries $\xi_{kj} = (1/\sqrt{n})x_{kj}$ and $\{x_{kj},\ k, j = 1, 2, \dots,\}$ forms an infinite double array of iid complex random variables of mean zero and variance one. Using the complex eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of Ξ_n, we can construct a two-dimensional empirical distribution by

$$\mu_n(x, y) = \frac{1}{n}\#\{i \le n\colon \mathrm{Re}(\lambda_k) \le x, \mathrm{Im}(\lambda_k) \le y\},$$

which is called the empirical spectral distribution of the matrix Ξ_n.

The motivation for the study of spectral analysis of large-dimensional random matrices comes from quantum mechanics. The energy level of a quantum is not directly observable and it is known that the energy levels of quantums can be described by the eigenvalues of a matrix of observations. Since the 1960's, the spectral analysis of large-dimensional random matrices has attracted considerable interest from probabilists, mathematicians and statisticians. For a general review, the reader is referred to, among others, Bai (1993a, b), Bai and Yin (1993, 1988a, b, 1986), Geman (1980, 1986), Silverstein and Bai (1995), Wachter (1978, 1980) and Yin, Bai and Krishnaiah (1988).

Most of the important existing results are on symmetric large-dimensional random matrices. Basically, two powerful tools are used in this area. The first is the moment approach which was successfully used in finding the limiting spectral distributions of large-dimensional random matrices and in establishing the strong convergence of extreme eigenvalues. See, for example, Bai and Yin (1993, 1988a, b, 1986), Geman (1980, 1986), Jonsson (1982) and Yin, Bai

Received March 1996.
[1]Supported by ROC Grant NSC84-2816-M110-009L.
AMS 1991 *subject classifications*. Primary 60F15; secondary 62H99.
Key words and phrases. Circular law, complex random matrix, noncentral Hermitian matrix, largest and smallest eigenvalue of random matrix, spectral radius, spectral analysis of large-dimensional random matrices.

and Krishnaiah (1988). The second is the Stieltjes transform which was used in Bai (1993a, b), Bai and Silverstein (1995), Marčenko and Pastur (1967), Pastur (1972, 1973), Silverstein and Choi (1995) and Wachter (1978, 1980). Unfortunately, these two approaches are not suitable for dealing with nonsymmetric random matrices. Due to lack of appropriate methodologies, very few results were known about nonsymmetric random matrices. The only known result is about the spectral radius of the matrix Ξ_n. Bai and Yin [(1986), under the fourth moment] and Geman [(1986), under some growth restrictions on all moments], independently proved that with probability 1, the upper limit of the spectral radius of Ξ_n is not greater than 1.

Since the early 1950's, it has been conjectured that the distribution $\mu_n(x, y)$ converges to the so-called circular law, that is, the uniform distribution over the unit disk in the complex plane. This problem has been unsolved, except where the entries are complex normal variables [given in an unpublished paper of Silverstein in 1984 but reported in Hwang (1986)]. Silverstein's proof relies on the explicit expression of the joint distribution density of the eigenvalues of Ξ_n [see, e.g., Ginibre (1965)]. Hence his approach cannot be extended to the general case. Girko presented (1984a, b) a proof of this conjecture under some conditions. However, the paper contained too many mathematical gaps, leaving the problem still open. After Girko's flaw was found, "many have tried to understand Girko's 'proofs' without success," [Edelman (1995)]. When the entries are iid real normal random variables, Edelman (1995) found the conditional joint distribution of the complex eigenvalues when the number of real eigenvalues are given and showed that the expected empirical spectral distribution of Ξ_n tends to the circular law.

In spite of mathematical gaps in his arguments, Girko had come up with an important idea (his Lemma 1), which established a relation between the characteristic function of the empirical spectral distribution of Ξ_n and an integral involving the empirical spectral distribution of a Hermitian matrix. Girko's Lemma 1 is presented below for easy reference.

GIRKO'S LEMMA 1. *For any $uv \neq 0$, we have*

$$\begin{aligned} m_n(u, v) &= \iint_2 \exp(iux + ivy)\,\mu_n(dx, dy) \\ &= \frac{u^2+v^2}{4iu\pi} \iint \frac{\partial}{\partial s}\left[\int_0^\infty \ln x\,\nu_n(dx, z)\right] \exp(ius + ivt)\,dt\,ds, \end{aligned} \tag{1.1}$$

where $z = s + it$, $i = \sqrt{-1}$ and $\nu_n(x, z)$ is the empirical spectral distribution of the nonnegative definite Hermitian matrix $\mathbf{H}_n = \mathbf{H}_n(z) = (\Xi_n - z\mathbf{I})^*(\Xi_n - z\mathbf{I})$. Here and throughout this paper, Ξ^* denotes the complex conjugate and transpose of the matrix Ξ.

It is easy to see that $m_n(u, v)$ is an entire function in both u and v. By Bai and Yin (1986) or Geman (1986), the family of distributions $\mu_n(x, y)$ is tight. And hence, every subsequence of $\mu_n(x, y)$ contains a completely convergent subsequence and the characteristic function $m(u, v)$ of the limit must be also entire. Therefore, to prove the circular law, applying Girko's Lemma 1, one needs only show that the right-hand side of (1.1) converges to its counterpart generated by the circular law. Note that the function $\ln x$ is not bounded at both infinity and zero. Therefore, the convergence of the right hand side of (1.1) cannot be simply reduced to the convergence of ν_n. In view of the results of Yin, Bai and Krishnaiah (1988), there would not be a serious problem for the upper limit of the inner integral, since the support of ν_n is a.s. eventually bounded from the right by $(2 + \varepsilon + |z|)^2$ for any positive ε. In his 1984 papers, Girko failed only in dealing with the lower limit of the integral.

In this paper, making use of Girko's lemma, we shall provide a proof of the famous circular law.

THEOREM 1.1 (Circular law). *Suppose that the entries of* $\mathbf{X}$ *have finite sixth moment and that the joint distribution of the real and imaginary part of the entries has a bounded density. Then, with probability* 1, *the empirical distribution* $\mu_n(x, y)$ *tends to the uniform distribution over the unit disc in two-dimensional space.*

The proof of the theorem will be rather tedious. Thus, for ease of understanding, an outline of the proof is provided first.

The proof of the theorem will be presented by showing that with probability 1, $m_n(u, v) \to m(u, v)$ for every (u, v) such that $uv \neq 0$. To this end, we need the following steps.

1. Reduce the range of integration. First we need to reduce the range of integration to a finite rectangle, so that the dominated convergence theorem is applicable. As will be seen, proof of the circular law reduces to showing that for every large $A > 0$ and small $\varepsilon > 0$,

$$\begin{aligned}\int\int_T \left[\frac{\partial}{\partial s}\int_0^\infty \ln x\,\nu_n(dx, z)\right]\exp(ius + ivt)\,ds\,dt \\ \to \iint_T\left[\frac{\partial}{\partial s}\int_0^\infty \ln x\,\nu(dx, z)\right]\exp(ius + ivt)\,ds\,dt,\end{aligned} \tag{1.2}$$

   where $T = \{(s, t);\ |s| \le A,\ |t| \le A^2,\ |\sqrt{s^2 + t^2} - 1| \ge \varepsilon\}$ and $\nu(x, z)$ is the limiting spectral distribution of the sequence of matrices $\mathbf{H}_n$ which determines the circular law.
2. Find the limiting spectrum $\nu(\cdot, z)$ of $\nu_n(\cdot, z)$ and show that it determines the circular law.
3. Find a convergence rate of $\nu_n(x, z)$ to $\nu(x, z)$ uniformly in every bounded region of z. Then, we will be able to apply the convergence rate to establish (1.2). As argued earlier, it is sufficient to show the following.

4. Show that for suitably defined sequence ε_n, with probability 1:

$$\limsup_{n\to\infty} \iint_T \left| \int_{\varepsilon_n}^{\infty} \ln x(\nu_n(dx, z) - \nu(dx, z)) \right| = 0, \tag{1.3}$$

and

$$\limsup_{n\to\infty} \left| \iint_T \int_0^{\varepsilon_n} \ln x \nu_n(dx, z)\, ds\, dt \right| = 0 \tag{1.4}$$

The convergence rate of $\nu_n(\cdot, z)$ will be used in proving (1.3). The proof of (1.4) will be specifically treated. The proofs of the above four steps are rather long and thus the paper is organized into several sections. For convenience, a list of symbols and their definitions are given in Section 2. Section 3 is devoted to the reduction of the integral range. In Section 4, we shall present some lemmas discussing the properties of the limiting spectrum ν and its Stieltjes transform, and some lemmas establishing a convergence rate of ν_n. The most difficult part of this work, namely, the proof of (1.4), is given in Section 5 and the proof of Theorem 1.1 is present in Section 6. Some discussions and extensions are given in Section 7. Some technical lemmas are presented in the Appendix.

2. List of notations. The definitions of the notations presented below will be given again when the notations appear.

$\{x_{kj}\}$: a double array of iid complex random variables with $\mathrm{E}(x_{kj}) = 0$, $\mathrm{E}|x_{kj}|^2 = 1$ and $\mathrm{E}|x_{kj}|^6 < \infty$;

$\mathbf{X}_n = (x_{kj})_{k,\, j=1,2,\dots,n}$. Its kth column vector is denoted by $\mathbf{x}_k$.

$\boldsymbol{\Xi}_n = (1/\sqrt{n})\mathbf{X}_n = (\xi_{jk}) = (\boldsymbol{\xi}_k)$.

$\mathbf{R}(z) = \boldsymbol{\Xi}_n - z\mathbf{I}_n$ with $z = s + it$ and $i = \sqrt{-1}$. Its kth column vector is denoted by $\mathbf{r}_k$.

$\mathbf{H}_n = \mathbf{R}^*(z)\mathbf{R}(z)$.

$\mathbf{A}^*$ denotes the complex conjugate and transpose of the matrix $\mathbf{A}$.

$m_n(u, v)$ and $m(u, v)$ denote the characteristic functions of the distributions μ_n and the circular law μ.

$F^{\mathbf{X}}$ denotes the empirical spectral distribution of $\mathbf{X}$ if $\mathbf{X}$ is a matrix. However, we do not use this notation for the matrix $\boldsymbol{\Xi}_n$ since it is traditionally and simply denoted as F_n.

$\alpha = x + iy$. In most cases, $y = y_n = n^{-1/60} \ln^{-1} n$. But in some places, y denotes a fixed positive number.

$\nu_n(x, z)$ denotes the empirical spectral distribution of $\mathbf{H}_n$ and $\nu(x, z)$ denotes its limiting spectral distribution.

$\Delta_n(\alpha)$ and $\Delta(\alpha)$ are the Stieltjes transforms of $\nu_n(x, z)$ and $\nu(x, z)$ respectively.

Boldface capitals will be used to denote matrices and boldface lower case used for vectors.

The symbol K_d denotes the upper bound of the joint density of its real and imaginary parts of the entries x_{kj}. In Section 7, it is also used for the upper bound of the conditional density of the real part of the entry of $\mathbf{X}$ when its imaginary part is given.

$\varepsilon_n = \exp(-n^{1/120})$, a constant.

$\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts of a complex number.

$I(\cdot)$ denotes the indicator function of the set in parentheses.

$\|f\|$ denotes the uniform norm of the function f, that is, $\|f\| = \sup_{\mathbf{x}}|f(\mathbf{x})|$.

$\|\mathbf{A}\|$ denotes the operation norm of the matrix $\mathbf{A}$, that is, its largest singular value.

3. Integral range reduction. Let $\mu_n(x, y)$ denote the empirical spectral distribution of the matrix $\Xi_n = (1/\sqrt{n})X_n$ and $\nu_n(x, z)$ denote the empirical distribution of the Hermitian matrix $\mathbf{H} = \mathbf{H}_n = (\Xi_n - z\mathbf{I})^*(\Xi_n - z\mathbf{I})$, for each fixed $z = s + it \in C$. The following lemma is the same as Girko's Lemma 1. We present a proof here for completeness; this proof is easier to understand than that provided by Girko (1984a, b).

LEMMA 3.1. *For all $u \neq 0$ and $v \neq 0$, we have*

$$\begin{aligned} m_n(u, v) &= \iint \exp(ius + ivy)\mu_n(dx, dy) \\ &= \frac{u^2 + v^2}{4iu\pi} \iint g_n(s, t)\exp(ius + ivt)\, dt\, ds \end{aligned} \tag{3.1}$$

where $\iint \cdots dt\, ds$ denotes the iterated integral $\int[\int \cdots dt]\, ds$ and

$$g_n(s, t) = \frac{1}{n}\sum_{k=1}^{n} \frac{2(s - \mathrm{Re}(\lambda_k))}{\left(s - \mathrm{Re}(\lambda_k)\right)^2 + \left(t - \mathrm{Im}(\lambda_k)\right)^2} = \frac{\partial}{\partial s}\int_0^\infty \ln x\, \nu_n(dx, z).$$

REMARK 3.1. When $z = \lambda_k$ for some $k \leq n$, $\nu_n(x, z)$ will have a positive measure of $1/n$ at $x = 0$ and hence the inner integral of $\ln x$ is not well defined. Therefore, the iterated integral in (3.1) should be understood as the generalized integral. That is, we cut off the n discs with centers $[\mathrm{Re}(\lambda_k), \mathrm{Im}(\lambda_k)]$ and radius ε from the s, t plane. Take the integral outside the n discs in the s, t plane and then take $\varepsilon \to 0$. Then, the outer integral in (3.1) is defined to be the limit [w.r.t. (with respect to) $\varepsilon \to 0$]) of the integral over the reduced integration range.

REMARK 3.2. Note that $g_n(s, t)$ is twice the real part of the Stieltjes transform of the two-dimensional empirical distribution μ_n, that is,

$$\iint \frac{\mu_n(dx, dy)}{x + iy - z} = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\lambda_k - z},$$

which has exactly n simple poles at the n eigenvalues of Ξ_n. The function $g_n(s, t)$ uniquely determines the n eigenvalues of the matrix Ξ_n. On the other hand, $g_n(s, t)$ can also be regarded as the derivative (w.r.t. s) of the logarithm of the determinant of $\mathbf{H}$ which can be expressed as an integral w.r.t. the empirical spectral distribution of $\mathbf{H}$, as given in the second equality in the definition of $g_n(s,t)$. In this way, the problem of the spectrum of a non-Hermitian matrix is transformed as one of the spectrum of a Hermitian matrix, so that the approach via Stieltjes transforms can be applied to this problem.

PROOF. Note that for all $uv \neq 0$,

$$\begin{aligned}
&\frac{u^2+v^2}{2iu\pi}\iint \frac{s}{s^2+t^2}\exp(ius+ivt)\,dt\,ds \\
&\quad = \frac{u^2+v^2}{2iu\pi}\iint \frac{\operatorname{sign}(s)}{1+t^2}\exp(ius+iv|s|t)\,dt\,ds \\
&\quad = \frac{u^2+v^2}{2iu}\int \operatorname{sign}(s)\exp(ius-|vs|)\,ds \\
&\quad = \frac{u^2+v^2}{2|u|}\int \sin|us|\exp(-|vs|)\,ds = 1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\iint \exp(iux+ivy)\mu_n(dx,dy) \\
&\quad = \frac{u^2+v^2}{2iu\pi}\iint \frac{1}{n}\sum_{k=1}^{n}\frac{s}{s^2+t^2} \\
&\qquad\qquad \times\exp\big(ius+ivt+iu\operatorname{Re}(\lambda_k)+iv\operatorname{Im}(\lambda_k)\big)\,dt\,ds \\
&\quad = \frac{u^2+v^2}{4iu\pi}\iint \frac{1}{n}\sum_{k=1}^{n}\frac{2(s-\operatorname{Re}(\lambda_k))}{(s-\operatorname{Re}(\lambda_k))^2+(t-\operatorname{Im}(\lambda_k))^2} \\
&\qquad\qquad \times\exp(ius+ivt)\,dt\,ds \\
&\quad = \frac{u^2+v^2}{4iu\pi}\iint \left[\frac{\partial}{\partial s}\int_0^{\infty}\ln x\,\nu_n(dx,z)\right]\exp(ius+ivt)\,dt\,ds.
\end{aligned}$$

The proof of Lemma 3.1 is complete. □

LEMMA 3.2. *For all* $uv \neq 0$, *we have*

$$\begin{aligned}
m(u,v) &= \frac{1}{\pi}\iint_{x^2+y^2\le 1}\exp(iux+ivy)\,dx\,dy \\
&= \frac{u^2+v^2}{4iu\pi}\iint g(s,t)\exp(ius+ivt)\,dt\,ds,
\end{aligned} \tag{3.2}$$

where

$$g(s,t)=\begin{cases}\dfrac{2s}{s^2+t^2}, & \text{if } s^2+t^2>1,\\ 2s, & \text{otherwise}.\end{cases}$$

PROOF. As in the proof of Lemma 3.1, we have, for all $uv \neq 0$,

$$m(u,v)=\frac{u^2+v^2}{4iu\pi}\iint\left[\frac{1}{\pi}\iint_{x^2+y^2\le 1}\frac{2(s-x)}{(s-x)^2+(t-y)^2}\,dx\,dy\right] \times\exp(ius+ivt)\,ds\,dt. \tag{3.3}$$

Then, the lemma follows from the fact that the inner integral on the right-hand side of (3.3) equals $g(s,t)$, using Green's formula. □

LEMMA 3.3. *For any $uv \neq 0$ and $A > 2$, with probability 1, when n is large, we have*

$$\left|\int_{|s|\ge A}\int_{-\infty}^{\infty} g_n(s,t)\exp(ius+ivt)\,ds\,dt\right| \le \frac{4\pi}{|v|}\exp\left(-\frac{1}{2}|v|A\right) \tag{3.4}$$

and

$$\left|\int_{|s|\le A}\int_{|t|\ge A^2} g_n(s,t)\exp(ius+ivt)\,ds\,dt\right| \le \frac{8(A+1+\varepsilon)}{A^2}. \tag{3.5}$$

Furthermore, the two inequalities above hold if the function $g_n(s,t)$ is replaced by $g(s,t)$.

PROOF. From Bai and Yin (1986), it follows that with probability 1, when n is large, we have $\max_k\{|\lambda_k|\} \le 1+\varepsilon$. Hence,

$$\begin{aligned}&\left|\int_{|s|\ge A}\int_{-\infty}^{\infty} g_n(s,t)\exp(ius+ivt)\,ds\,dt\right|\\ &\quad=\left|\int_{|s|\ge A}\int_{-\infty}^{\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{2(s-\mathrm{Re}(\lambda_k))}{(s-\mathrm{Re}(\lambda_k))^2+(t-\mathrm{Im}(\lambda_k))^2}\right.\\ &\qquad\qquad\left.\times\exp(ius+ivt)\,ds\,dt\right|\\ &\quad=\left|\frac{\pi}{n}\sum_{k=1}^{n}\int_{|s|\ge A}\mathrm{sign}(s-\mathrm{Re}(\lambda_k))\exp\big(ius-|v(s-\mathrm{Re}(\lambda_k))|\big)\,ds\right|\\ &\quad\le\frac{4\pi}{|v|}\exp\left(-\frac{1}{2}|v|A\right),\end{aligned} \tag{3.6}$$

and

$$\left|\int_{|s|\le A}\int_{|t|\ge A^2} g_n(s,t)\exp(ius+ivt)\,ds\,dt\right| \le \frac{8(A+1+\varepsilon)}{A^2}. \tag{3.7}$$

Similarly, one can prove the above two inequalities for $g(s,t)$. The proof of Lemma 3.3 is complete. □

From Lemma 3.3, one can see that the right-hand sides of (3.4) and (3.5) can be made arbitrarily small by making A large enough. The same is true when $g_n(s,t)$ is replaced by $g(s,t)$. Therefore, the proof of the circular law is reduced to showing

$$\int_{|s|\le A}\int_{|t|\le A^2}[g_n(s,t)-g(s,t)]\exp(ius-ivt)\,ds\,dt \to 0. \tag{3.8}$$

Finally, define sets

$$T = \{(s,t): |s|\le A, |t|\le A^2 \text{ and } ||z|-1|\ge\varepsilon\}$$

and

$$T_1 = \{(s,t): ||z|-1|<\varepsilon\},$$

where $z = s+it$.

LEMMA 3.4. *For all fixed A and* $0<\varepsilon<1$, *for all* n,

$$\left|\iint_{T_1} g_n(s,t)\,ds\,dt\right| \le 24\pi\sqrt{\varepsilon}. \tag{3.9}$$

Furthermore, when $g_n(s,t)$ *in* (3.9) *is replaced by* $g(s,t)$, *the estimation* (3.9) *remains true.*

PROOF. For any fixed u and v, by a polar transformation, we obtain

$$\left|\iint_{T_1}\frac{(s-u)\,dt\,ds}{(s-u)^2+(t-v)^2}\right| = \left|\int_0^{2\pi}2D(\theta)\cos\theta\,d\theta\right| \le 24\pi\sqrt{\varepsilon},$$

where $D(\theta)$ is the sum of lengths of at most two segments which are the intersection of the ring T_1 and the straight line $(s-u)\cos\theta+(t-v)\sin\theta = 0$. In the above, we have used the fact that $\max_\theta D(\theta)\le 2\sqrt{4\varepsilon+2\varepsilon^2}\le 6\sqrt{\varepsilon}$.

This completes the proof of (3.9) for $g_n(s,t)$. The proof of (3.9) for $g(s,t)$ is similar and thus omitted. The proof of Lemma 3.4 is complete. □

Note that the right-hand side of (3.9) can be made arbitrarily small by choosing ε small. Thus, by Lemmas 3.3 and 3.4, to prove the circular law, one needs only to show that, for each fixed $A>0$ and $\varepsilon\in(0,1)$,

$$\iint_T (g_n(s,t)-g(s,t))\,ds\,dt \to 0 \quad \text{a.s.} \tag{3.10}$$

4. Convergence of $\nu_n(x,z)$ and the limiting spectrum $\nu(x,z)$. In this section, we shall establish a convergence rate of $\nu_n(x, z)$ and discuss properties of the limiting distribution $\nu(x, z)$ of $\nu_n(x, z)$. Throughout the remainder of this paper, we shall use the notations $o(1)$ and $O(1)$ in the sense of "almost surely." Furthermore, if the quantities represented by the symbols $o(1)$ or $O(1)$ are involved with indices j, l or k, or variables α or z, then the orders are uniform about these indices and variables.

Suppose that $\nu(\cdot, z)$ is the limiting spectral distribution of some convergent subsequence of $\nu_n(\cdot, z)$. Denote by $\Delta_n(\alpha, z)$ and $\Delta(\alpha, z)$, $\alpha = x + iy$, $y > 0$, the Stieltjes transforms of $\nu_n(\cdot, z)$ and $\nu(\cdot, z)$, respectively, that is,

$$\Delta_n(\alpha, z) = \int \frac{1}{x-\alpha}\nu_n(dx, z) = \frac{1}{n}\operatorname{tr}(\mathbf{H} - \alpha\mathbf{I})^{-1}$$

and

$$\Delta(\alpha, z) = \int \frac{1}{x-\alpha}\nu(dx, z),$$

where α is a complex number with positive imaginary part. The variable z in these symbols will be omitted when there is no confusion. We will prove the following lemmas.

LEMMA 4.1. *Suppose that the conditions of Theorem* 1.1 *are true. Write*

$$\Delta_n(\alpha)^3 + 2\Delta_n(\alpha)^2 + \frac{\alpha + 1 - |z|^2}{\alpha}\Delta_n(\alpha) + \frac{1}{\alpha} = r_n, \tag{4.1}$$

where $r_n = r_n(\alpha, z)$. *Then, we have*

$$\Delta^3 + 2\Delta^2 + \frac{\alpha + 1 - |z|^2}{\alpha}\Delta + \frac{1}{\alpha} = 0 \tag{4.2}$$

the remainder term r_n *satisfies*

$$\begin{gathered}\sup\{|r_n|: \alpha = x + iy \text{ with } -\infty < x < \infty, u \ge y_n, |z| \le M\} = o(\delta_n),\\ y_n = n^{-1/60}\ln^{-1} n \quad \text{and} \quad \delta_n = n^{-1/60}.\end{gathered} \tag{4.3}$$

LEMMA 4.2. *The limiting distribution function* $\nu(x, z)$ *satisfies*

$$|\nu(x+u, z) - \nu(x, z)| \le \pi^{-1}\sqrt{2}\max\left\{\sqrt{|u|}, |u|\right\} \quad \text{for all } z. \tag{4.4}$$

Also, the limiting distribution function $\nu(x, z)$ *is supported by the interval* (x_1, x_2) *when* $|z| > 1$ *and by* $(0, x_2)$ *when* $|z| \le 1$, *where*

$$\begin{aligned} x_1 &= \frac{1}{8|z|^2}\left[-1 + 20|z|^2 + 8|z|^4 - \sqrt{(1+8|z|^2)^3}\right],\\ x_2 &= \frac{1}{8|z|^2}\left[\sqrt{(1+8|z|^2)^3} - 1 + 20|z|^2 + 8|z|^4\right] \quad \text{when } z \ne 0,\\ &= 4 \quad \text{when } z = 0.\end{aligned}$$

LEMMA 4.3. *Let $m_2(\alpha)$ and $m_3(\alpha)$ denote the two solutions of* (4.2) *other than* $\Delta(\alpha)$. *For any given constants $N > 0$, $A > 0$ and $\varepsilon \in (0,1)$, there exist positive constants ε_0 and ε_1 such that for all large n, $|\alpha| \le N$, $y \ge 0$ and $z \in T$, we have the following*:

(i)

$$\max_{j=2,3} |\Delta(\alpha) - m_j(\alpha)| \ge \varepsilon_0, \tag{4.5}$$

(ii) *for* $|\alpha - x_2| \ge \varepsilon_1$, (*and* $|\alpha - x_1| \ge \varepsilon_1$, *if* $|z| > 1$),

$$\min_{j=2,3} |\Delta(\alpha) - m_j(\alpha)| \ge \varepsilon_0, \tag{4.6}$$

(iii) *for* $|\alpha - x_2| < \varepsilon_1$,

$$\min_{j=2,3} |\Delta(\alpha) - m_j(\alpha)| \ge \varepsilon_0 \sqrt{|\alpha - x_2|}, \tag{4.7}$$

(iv) *for* $|z| > 1 + \varepsilon$, *and* $|\alpha - x_1| < \varepsilon_1$,

$$|\Delta(\alpha) - m_j(\alpha)| \ge \varepsilon_0 \sqrt{|\alpha - x_1|}. \tag{4.8}$$

REMARK 4.1. This lemma basically says that the Stieltjes transform of the limiting spectral distribution $\nu(\cdot, z)$ is distinguishable from the other two solutions of the equation (4.2). Here, we give a more explicit estimate of the distance of $\Delta(\alpha)$ from the other two solutions. This lemma simply implies that the limiting spectral distribution of the sequence of matrices $\mathbf{H}_n$ is unique and nonrandom since the variation from ν_n to ν_{n+1} is of order $O(1/n)$ and hence the variation from $\Delta_n(\alpha)$ to $\Delta_{n+1}(\alpha)$ is $O(1/ny)$.

LEMMA 4.4. *We have*

$$\frac{\partial}{\partial s} \int_0^\infty \ln x \nu(dx, z) = g(s, t). \tag{4.9}$$

LEMMA 4.5. *Under the conditions of Theorem* 1.1, *for any* $M_2 > M_1 > 0$,

$$\begin{aligned} \sup_{M_1 \le |z| \le M_2} \|\nu_n(\cdot, z) - \nu(\cdot, z)\| &:= \sup_{x,\, M_1 \le |z| \le M_2} |\nu_n(x, z) - \nu(x, z)| \\ &= o(n^{-1/120}). \end{aligned} \tag{4.10}$$

REMARK 4.2. Lemma 4.5 is used only in proving (1.3) for a suitably chosen ε_n. From the proof of the lemma and comparing with the results in Bai (1993a, b) one can see that a better rate of convergence can be obtained by considering more terms in the expansion. As the rate given in (4.10) is enough for our purposes, we restrict ourselves to the weaker result (4.10) by a simpler proof, rather than trying to get a better rate by long and tedious arguments.

PROOF OF LEMMA 4.1. This lemma plays a key role in establishing a convergence rate of the empirical spectral distribution $\nu_n(\cdot, z)$ of $\mathbf{H}$. The approach used in the proof of this lemma is in a manner typical in the application of Stieltjes transforms to the spectral analysis of large-dimensional random matrices. The basic idea of the proof relies on the following two facts: (1) the n diagonal elements of $(\mathbf{H} - \alpha\mathbf{I})^{-1}$ are identically distributed and asymptotically the same as their average, the Stieltjes transform of the empirical spectral distribution of $\mathbf{H}$; (2) for all $k \le n$, $(1/n)\operatorname{tr}((\tilde{\mathbf{H}}_k - \alpha\mathbf{I}_{n-1})^{-1})$ are identically distributed and asymptotically equivalent to $(1/n)\operatorname{tr}((\mathbf{H} - \alpha\mathbf{I}_n)^{-1})$, where the matrix $\tilde{\mathbf{H}}_k$ is defined similarly as $\mathbf{H}$ by $\boldsymbol{\Xi}$ with the kth column and row removed. By certain steps of expansion, one can obtain the equation (4.1) which determines the Stieltjes transform $\Delta_n(\alpha)$ of $\mathbf{H}$.

Since $\Delta(\alpha)$ is the limit of some convergent subsequence of $\Delta_n(\alpha)$ and hence (4.2) is a consequence of (4.3), only (4.3) need be shown.

To begin, we need to reduce the uniform convergence of (4.3) over an uncountable set to that over a finite set. Yin, Bai and Krishnaiah (1988), proved that $\|\boldsymbol{\Xi}_n\| \to 2$, a.s., where $\|\boldsymbol{\Xi}_n\|$ denotes the operator norm, that is, the largest singular value, of the matrix $\boldsymbol{\Xi}_n$ when the entries of $\mathbf{X}$ are all real. Their proofs can be translated to the complex case word for word so that the above result is still true when the entries are complex. Therefore, with probability 1, when n is large enough, for all $|z| \le M$,

$$\Lambda_{\max}(\mathbf{H}_n) \le \left(\|\boldsymbol{\Xi}_n\| + |z|\right)^2 \le (3 + M)^2. \tag{4.11}$$

Hence, when $|\alpha| \ge n^{1/60} \ln n$ and (4.11) is true, we have for all large n

$$|\Delta_n(\alpha)| \le \frac{1}{|\alpha| - (3 + M)^2}$$

and consequently,

$$\begin{aligned} |r_n| &= \left| \Delta_n^2 + 2\Delta_n^2 + \frac{\alpha + 1 - |z|^2}{\alpha}\Delta_n + \frac{1}{\alpha} \right| \\ &\le 4Mn^{-1/60} \ln^{-1} n = o(\delta_n). \end{aligned} \tag{4.12}$$

If $\max(|\alpha|, |\alpha'|) < n^{1/60} \ln n$ and $|\alpha - \alpha'| \le n^{-1/7}$, then

$$|\Delta_n(\alpha) - \Delta_n(\alpha')| \le \left[\min(y, y')\right]^{-2} |\alpha - \alpha'| \le y_n^{-2} n^{-1/7},$$

which implies that

$$|r_n(\alpha) - r_n(\alpha')| \le M y_n^{-4} n^{-1/7} \le M n^{-1/14} \tag{4.13}$$

for some positive constant M.

Suppose that $|z - z'| \le n^{-1/4}$. Let $\Lambda_k(z)$ and $\Lambda_k(z')$ (arranged in increasing order) be eigenvalues of the matrices $\mathbf{H}(z) = (\boldsymbol{\Xi}_n - z\mathbf{I})^*(\boldsymbol{\Xi}_n - z\mathbf{I})$ and

$\mathbf{H}(z') = (\boldsymbol{\Xi}_n - z'\mathbf{I})^*(\boldsymbol{\Xi}_n - z'\mathbf{I})$, respectively. Then for any fixed α, by Lemma A.5, we have

$$
\begin{aligned}
&|\Delta_n(\alpha, z) - \Delta_n(\alpha, z')| \\
&\quad \le \frac{1}{n}\sum_{k=1}^{n} \frac{|\Lambda_k(z) - \Lambda_k(z')|}{|\Lambda_k(z) - \alpha||\Lambda_k(z') - \alpha|} \\
&\quad \le y^{-2}|z - z'|\left(\frac{1}{n}\operatorname{tr}\big(2\boldsymbol{\Xi}_n - (z + z')\mathbf{I}\big)^*\big(2\boldsymbol{\Xi}_n - (z + z')\mathbf{I}\big)\right)^{1/2} \\
&\quad \le y^{-2}|z - z'|(3 + 2M) \le Mn^{-1/6}.
\end{aligned}
\tag{4.14}
$$

This, together with (4.12) and (4.13), shows that to finish the proof of (4.3), it is sufficient to show that

$$
\max_{l, j \le n}\big\{|r_n(\alpha_l, z_j)|\big\} = o(\delta_n),
\tag{4.15}
$$

where $\alpha_l = x(l) + iy(l)$, $l = 1, 2, \ldots, [n^{1/6}]$ and z_j, $j = 1, 2, \ldots, [n^{1/3}]$ are selected so that $|x(l)| \le n^{1/60} \ln n$, $y_n \le y(l) \le n^{1/60} \ln n$ and for each $|\alpha| \le n^{1/60} \ln n$ with $y \ge y_n$, there is an l such that $|\alpha - \alpha_l| < n^{-1/7}$; and for each $|z| \le M$, there is a j such that $|z - z_j| \le n^{-1/4}$.

In the rest of the proof of this lemma, we shall suppress the indices l and j from the variables α_l and z_j. The reader should remember that we shall only consider those α_l and z_j which are selected to satisfy the properties described in the last paragraph.

Let $\mathbf{R} = \mathbf{R}_n(z) = (r_{kj})$, where $r_{kj} = \xi_{kj}$ for $j \ne k$ and $r_{kk} = \xi_{kk} - z$. Then $\mathbf{H} = \mathbf{R}^*\mathbf{R}$. We have

$$
\begin{aligned}
\Delta_n(\alpha) &= \frac{1}{n}\operatorname{tr}(\mathbf{H} - \alpha\mathbf{I})^{-1} \\
&= \frac{1}{n}\sum_{k=1}^{n}\frac{1}{|\mathbf{r}_k|^2 - \alpha - \mathbf{r}_k^*\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\mathbf{r}_k},
\end{aligned}
\tag{4.16}
$$

where $\mathbf{r}_k$ denotes the kth column vector of $\mathbf{R}$, $\mathbf{R}_k$ consists of the remaining $n - 1$ columns of $\mathbf{R}$ when $\mathbf{r}_k$ is removed and $\mathbf{H}_k = \mathbf{R}_k^*\mathbf{R}_k$.

First, notice that

$$
\begin{aligned}
&\Big||\mathbf{r}_k|^2 - \alpha - \mathbf{r}_k^*\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\mathbf{r}_k\Big| \\
&\qquad \ge \Big|\operatorname{Im}\Big(|\mathbf{r}_k|^2 - \alpha - \mathbf{r}_k^*\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\mathbf{r}_k|\Big)\Big| \ge y.
\end{aligned}
\tag{4.17}
$$

By Lemma A.4, we conclude that

$$
\max_{j, l, k \le n}\Big||\mathbf{r}_k|^2 - \big(1 + |z|^2\big)\Big| = o(n^{-5/36}\ln^2 n).
\tag{4.18}
$$

As mentioned earlier, with probability 1 for all large n, the norm of $\mathbf{R}$ is not greater than $3 + M$. We conclude that with probability 1, for all large n, the eigenvalues and hence the entries of $\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*$ are bounded by $(3 + M)^2/y \le (3 + M)^2/y_n$. Therefore, the sum of squares of absolute values of entries of any row of $\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*$ is not greater than $(3+M)^4/y^2 \le (3+M)^4/y_n^2$. By applying Lemma A.4 and noticing that $\mathbf{r}_k = (1/\sqrt{n})\mathbf{x}_k - z\mathbf{e}_k$, where $\mathbf{e}_k$ is the vector whose elements are all zero except the kth element which is 1, we obtain

$$\begin{aligned}\max_{j,l,k\le n}\Big|\mathbf{r}_k^*\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\mathbf{r}_k \\ -\Big(\frac{1}{n}\operatorname{tr}\Big(\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big) \\ +|z|^2\Big[\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big]_{kk}\Big)\Big| \\ = O\big(y_n^{-1}n^{-5/36}\ln^2 n\big),\end{aligned} \tag{4.19}$$

where $[\mathbf{A}]_{kk}$ denote the (k,k)th element of the matrix $\mathbf{A}$.

Now, denote by $\Lambda_1 \le \cdots \le \Lambda_n$ and $\Lambda_{k,1} \le \cdots \le \Lambda_{k(n-1)}$ the eigenvalues of $\mathbf{H}$ and those of $\mathbf{H}_k$, respectively. Then by the relation $0 \le \Lambda_l - \Lambda_{k,l-1} \le \Lambda_l - \Lambda_{l-1}$, and by the fact that with probability 1

$$\Lambda_n \le (2+|z|)^2 + \varepsilon \quad \text{for all large } n,$$

$$\frac{1}{n}\operatorname{tr}\Big(\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big) = 1 - \frac{1}{n} + \frac{\alpha}{n}\operatorname{tr}\Big((\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\Big),$$

and

$$\begin{aligned}\Big|\operatorname{tr}\Big((\mathbf{H} - \alpha\mathbf{I})^{-1}\Big) - \operatorname{tr}\Big((\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\Big)\Big| \\ = \left|\sum_{l=2}^{n}\frac{\Lambda_l - \lambda_{k,l-1}}{(\Lambda_l - \alpha)(\Lambda_{k,l-1} - \alpha)} + \frac{1}{\Lambda_1 - \alpha}\right| \\ \le \Lambda_n/y^2 + 1/y,\end{aligned} \tag{4.20}$$

we conclude that

$$\begin{aligned}\max_{j,l,k\le n}\left|\frac{1}{n}\operatorname{tr}\Big(\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big) - 1 - \alpha\Delta_n(\alpha)\right| \\ \le \frac{1}{n} + \frac{|\alpha|}{n}\big(\Lambda_n/y^2 + 1/y\big) = o(n^{-4/5}).\end{aligned} \tag{4.21}$$

We now estimate $[\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*]_{kk}$. Let β_k' denote the kth row of $\mathbf{R}_k$, and $\tilde{\mathbf{R}}_k$ denote the matrix of the remaining $n-1$ rows of $\mathbf{R}_k$ when β_k' is removed. Also, write $\tilde{\mathbf{H}}_k = \tilde{\mathbf{R}}_k^*\tilde{\mathbf{R}}_k$. Note that β_k' is just the kth row of $\boldsymbol{\Xi}_n$ with the kth element removed. Then we have

$$\begin{aligned}\Big[\mathbf{R}_k(\mathbf{H}_k - \alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big]_{kk} &= \beta_k'\Big(\tilde{\mathbf{H}}_k - \alpha\mathbf{I}_{n-1} + \bar{\beta}_k\beta_k'\Big)^{-1}\bar{\beta}_k \\ &= \frac{\beta_k'\big(\tilde{\mathbf{H}}_k - \alpha\mathbf{I}_{n-1}\big)^{-1}\bar{\beta}_k}{1 + \beta_k'\big(\tilde{\mathbf{H}}_k - \alpha\mathbf{I}_{n-1}\big)^{-1}\bar{\beta}_k}.\end{aligned} \tag{4.22}$$

Applying Lemma A.4 with $K_a = y_n^{-1}$, we obtain

$$(4.23)\qquad \begin{aligned}\max_{j,l,k\le n}&\left|\beta_k'\big(\tilde{\mathbf{H}}_k-\alpha\mathbf{I}_{n-1}\big)^{-1}\bar{\beta}_k-\frac{1}{n}\operatorname{tr}\Big(\big(\tilde{\mathbf{H}}_k-\alpha\mathbf{I}_{n-1}\big)^{-1}\Big)\right|\\ &=O\big(y_n^{-1}n^{-5/36}\ln^2 n\big).\end{aligned}$$

By elementary knowledge of matrix theory, $\Lambda_j(\mathbf{H})) \le \Lambda_j(\tilde{\mathbf{H}}_k) \le \Lambda_{j+2}(\mathbf{H})$, and we have

$$(4.24)\qquad \left|\frac{1}{n}\operatorname{tr}\Big(\big(\tilde{\mathbf{H}}_k-\alpha\mathbf{I}_{n-1}\big)^{-1}\Big)-\Delta_n(\alpha)\right|\le 4y_n^{-2}n^{-1}=o\big(n^{-9/10}\big).$$

By noticing

$$\big|\alpha(1+\Delta_n(\alpha))\big|\ge\big|\operatorname{Im}\big(\alpha(1+\Delta_n(\alpha))\big)\big|\ge y,$$

$$\left|\alpha\Big(1+\beta_k'\big(\tilde{\mathbf{H}}_k-\alpha\mathbf{I}_{n-1}\big)^{-1}\tilde{\beta}_k\Big)\right|\ge y$$

and by (4.23) and (4.24), one obtains

$$(4.25)\qquad \begin{aligned}\max_{j,l,k\le n}&\left|\Big[\mathbf{R}_k(\mathbf{H}_k-\alpha\mathbf{I}_{n-1})^{-1}\mathbf{R}_k^*\Big]_{kk}-\frac{\Delta_n(\alpha)}{1+\Delta_n(\alpha)}\right|\\ &\le\max_{j,l,k\le n}|\alpha|^2y^{-2}\left|\beta_k'\big(\tilde{\mathbf{H}}_k-\alpha\mathbf{I}_{n-1}\big)^{-1}\bar{\beta}_k-\Delta_n(\alpha)\right|\\ &=o\big(\alpha^2y_n^{-3}n^{-5/36}\ln n\big).\end{aligned}$$

Combining estimates (4.16)–(4.25), we conclude that

$$(4.26)\qquad \max_{j,l\le n}\left|\Delta_n(\alpha)-\frac{1+\Delta_n(\alpha)}{|z|^2-\alpha\big(1+\Delta_n(\alpha)\big)^2}\right|=o\big(\alpha^2y_n^{-3}n^{-5/36}\ln^2 n\big).$$

From this, one can see that r_n is controlled by $o(\alpha^2y_n^{-5}n^{-5/36}\ln^2 n)) = o(\delta_n)$ and thus the error estimate (4.1) follows. The proof of Lemma 4.1 is complete. □

PROOF OF LEMMA 4.2. Note that the Stieltjes transform $\Delta(\alpha)$ of the limiting spectral distribution $\nu(\cdot, z)$ is an analytic solution in α on the upper half plane $y > 0$ to the equation (4.2). It can be continuously extended to the "closed" upper plane $y \ge 0$ (but $\alpha \ne 0$). By way of the Stieltjes transform [see Bai (1993a) or Silverstein and Choi (1995)], it can be shown that $\nu(\cdot, z)$ has a continuous density (probably excluding $x = 0$ when $|z| \le 1$), say $p(\cdot, z)$, such that

$$\nu(x,z)=\int_0^x p(u,z)\,du$$

and $p(x, z) = \pi^{-1}\operatorname{Im}(\Delta(x))$. Since $p(x, z)$ is the density of the limiting spectral distribution $\nu(\cdot, z)$, $p(x, z) = 0$ for all $x < 0$ and $x > (2 + |z|)^2$. Let $x > 0$ be an inner point of the support of $\nu(\cdot, z)$. Write $\Delta(x) = g(x) + ih(x)$. Then, to prove (4.4), it suffices to show

$$h(x)\le\max\Big\{\sqrt{2/x},1\Big\}.$$

Rewrite (4.2) for $\alpha = x$ as

$$\Delta^2 + 2\Delta + 1 + \frac{1 - |z|^2}{x} + \frac{1}{x\Delta} = 0.$$

Comparing the imaginary and real parts of both sides of the above equation, we obtain

$$2(g(x) + 1) = \frac{1}{x(g^2(x) + h^2(x))}. \tag{4.27}$$

and

$$\begin{aligned} h^2(x) &= \frac{1 - |z|^2}{x} + (g(x) + 1)^2 + \frac{g(x)}{x(g^2(x) + h^2(x))} \\ &\le \frac{1}{x} + \frac{1}{4x^2(g^2(x) + h^2(x))^2} + \frac{1}{2xh(x)} \\ &\le \frac{1}{x} + \frac{1}{4x^2h^4(x)} + \frac{1}{2xh(x)}. \end{aligned} \tag{4.28}$$

This implies $h(x) \le \max\{\sqrt{2/x}\ 1\}$, because substituting the reverse inequality $h(x) > \sqrt{2/x}$ (or $h(x) > 1$) will lead to a contradiction if $0 < x < 2$ (or $x \ge 2$, correspondingly). Thus, (4.4) is established.

Now, we proceed to find the boundary of the support of $\nu(\cdot, z)$. Since $\nu(\cdot, z)$ has no mass on the negative half line, we need only consider $x > 0$. Suppose $h(x) > 0$. Comparing the real and imaginary parts for both sides of (4.2) and then making x approach the boundary [namely, $h(x) \to 0$], we obtain

$$x(g^3 + 2g^2 + g) + (1 - |z|^2)g + 1 = 0$$

and

$$x(3g^2 + 4g + 1) + 1 - |z|^2 = 0. \tag{4.29}$$

Thus,

$$\left[(1 - |z|^2)g + 1\right](3g + 1) = (1 - |z|^2)g(g + 1).$$

For $|z| \ne 1$, the solution to this quadratic equation in g is

$$g = \frac{-3 \pm \sqrt{1 + 8|z|^2}}{4 - 4|z|^2} \quad \left(g = -\frac{1}{3} \text{ if } |z| = 1\right) \tag{4.30}$$

which, together with (4.29), implies that, for $|z| \ne 1$,

$$\begin{aligned} x_{1,2} &= -\frac{1 - |z|^2}{(g + 1)(3g + 1)} \\ &= \begin{cases} -\dfrac{1}{8|z|^2}\left\{1 - 20|z|^2 - 8|z|^4 \pm \sqrt{(1 + 8|z|^2)^3}\right\}, & \text{if } z \ne 0, \\ x_1 = -\infty \text{ and } x_2 = 4, & \text{if } z = 0. \end{cases} \end{aligned} \tag{4.31}$$

Note that $0 < x_1 < x_2$ when $|z| > 1$. Hence, the interval (x_1, x_2) is the support of $\nu(\cdot, z)$ since $p(x, z) = 0$ when x is very large. When $|z| < 1$, $x_1 < 0 < x_2$. Note that for the case $|z| < 1$, $g(x_1) < 0$ which contradicts the fact that $\Delta(x) > 0$ for all $x < 0$ and hence x_1 is not a solution of the boundary. Thus, the support of $\nu(\cdot, z)$ is the interval $(0, x_2)$. For $|z| = 1$, there is only one solution $x_2 = -1/[g(g+1)^2] = 27/4$, which can also be expressed by (4.31). In this case, the support of $\nu(\cdot, z)$ is $(0, x_2)$. The proof of Lemma 4.2 is complete. □

PROOF OF LEMMA 4.3. We first prove that $\Delta(\alpha)$ does not coincide with other roots of the equation (4.2) for $y \geq 0$ and $\alpha \neq x_{1,2}$. Otherwise, if for some α, $\Delta(\alpha)$ is a multiple root of (4.2), then it must be also a root of the derivative of the equation (4.2), that is,

$$3\Delta^2 + 4\Delta + \frac{\alpha + 1 - |z|^2}{\alpha} = 0. \tag{4.32}$$

Similar to the proof of Lemma 4.2, solving equations (4.2) and (4.32), one obtains $\alpha = x_1$ or x_2 and Δ is the same as g given in (4.30). Our assertion is proved.

We now prove (4.7). Let $\Delta + \rho$ be either m_2 or m_3. Since both Δ and $\Delta + \rho$ satisfy (4.2), we obtain

$$\rho = -\frac{3\Delta^2(\alpha) + 4\Delta(\alpha) + 1 + (1 - |z|^2)/\alpha}{3\Delta(\alpha) + 2 + \rho}. \tag{4.33}$$

Write $\hat{\rho} = \Delta(\alpha) - \Delta(x_2)$. By (4.29), we have

$$\begin{aligned} 3\Delta^2(\alpha) &+ 4\Delta(\alpha) + 1 + (1 - |z|^2)/\alpha \\ &= 3\Delta^2(\alpha) + 4\Delta(\alpha) + 1 + (1 - |z|^2)/\alpha \\ &\quad - \left[3\Delta^2(x_2) + 4\Delta(x_2) + 1 + (1 - |z|^2)/x_2\right] \\ &= \hat{\rho}\left[6\Delta(x_2) + 4 + 3\hat{\rho}\right] + \frac{(1 - |z|^2)(x_2 - \alpha)}{x_2 \alpha}. \end{aligned} \tag{4.34}$$

From (4.2) and (4.29), it follows that

$$\begin{aligned} 0 &= \left[3\Delta^2(x_2) + 4\Delta(x_2) + 1 + (1 - |z|^2)/\alpha\right]\hat{\rho} \\ &\quad + \left[3\Delta(x_2) + 2\right]\hat{\rho}^2 + \hat{\rho}^3 \\ &\quad + \frac{(x_2 - \alpha)\left(\Delta(x_2)(1 - |z|^2) + 1\right)}{x_2 \alpha} \\ &= \left[3\Delta(x_2) + 2 + \hat{\rho}\right]\hat{\rho}^2 \\ &\quad + \frac{(x_2 - \alpha)\left[(\Delta(x_2) + \hat{\rho})(1 - |z|^2) + 1\right]}{x_2 \alpha}. \end{aligned} \tag{4.35}$$

Note that $\Delta(x_2)(1-|z|^2)+1/4=1/4(1+\sqrt{1+8|z|^2})\geq 1/2$. Equation (4.35) implies that

$$|\hat{\rho}|\geq \min\left(\frac{1}{4M^2},\sqrt{\frac{|x_2-\alpha|}{6(|\Delta(x_2)|+1)}}\right)\geq c_1\sqrt{|x_2-\alpha|}, \tag{4.36}$$

for some positive constant c_1. Note that Δ is continuous in the rectangle $\{(\alpha,z);\ z\in T,\ x_{2,\min}-\varepsilon_1\leq x\leq x_{2,\max},\ 0\leq y\leq N\}$, where $x_{2,\min}=4$ (corresponding to $z=0$) and $x_{2,\max}=(1/8M^2)[(1+8M^2)^{3/2}-1+20M^2+8M^4]$ (corresponding to $|z|=M=\sqrt{A^2+A^4}$). Therefore, we may select a positive constant ε_1 such that for all $|z|\leq M$ and $|\alpha-x_2|\leq\varepsilon_1$, $|\hat{\rho}|\leq\min(\frac{1}{8},c_1^2/M^4)$. Then, from (4.33) and (4.34) and the fact that when $|\rho(\alpha)|\leq\frac{1}{8}$,

$$|3\Delta(\alpha)+2+\rho(\alpha)|\leq 4,$$

we conclude that

$$\begin{aligned}|\rho(\alpha)|&\geq\min\left(\frac{1}{8},\frac{1}{4}\left|\hat{\rho}\left[6\Delta(x_2)+4+3\hat{\rho}\right]+\frac{(1-|z|^2)(x_2-\alpha)}{x_2\alpha}\right|\right)\\&\geq\min\left(\frac{1}{8},\frac{1}{8}c_1\sqrt{|x_2-\alpha|}-\frac{1}{36}M^2|x_2-\alpha|\right)\\&\geq c_2\sqrt{|x_2-\alpha|}.\end{aligned}\tag{4.37}$$

This concludes the proof of (4.7).

The proof of (4.8) is similar to that of (4.7). Checking the proof of (4.7), one finds that equations (4.33)–(4.35) are still true if x_2 is replaced by x_1. The rest of the proof depends on the fact that for all $z\in T$, $|z|\geq 1+\varepsilon$ and $|\alpha-x_1|\leq\varepsilon_1$, $|3\Delta(\alpha)+2+\rho(\alpha)|$ has a uniform upper bound and $\hat{\rho}$ can be made as small as desired provided ε_1 is small enough. Indeed, this can be done because x_1 has a strictly positive minimum $x_{1,\min}$ at $|z|=1+\varepsilon$, and hence, $\Delta(\alpha)$ is uniformly continuous in the rectangle $\{(\alpha,z);\ z\in T,\ x_{1,\min}-\varepsilon_1\leq x\leq x_{1,\max},\ 0\leq y\leq N\}$, provided ε_1 is chosen so that $x_{1,\min}-\varepsilon_1>0$.

We claim that (4.6) is true. If not, then for each k, there exist α_k and z_k with $z_k\in T$ and $|\alpha_k-x_2|\geq\varepsilon_1$ (and $|\alpha_k-x_1|\geq\varepsilon_1$ if $|z_k|\geq 1+\varepsilon$), such that

$$\min_{j=2,3}|\Delta(\alpha_k)-m_j(\alpha_k)|<\frac{1}{k}.$$

Then, we may select a subsequence $\{k'\}$ such that $\alpha_{k'}\to\alpha_0$ and $z_{k'}\to z_0\in T$ and $|\alpha_0-x_2|\geq\varepsilon_1$. If $|z_0|\geq 1+\varepsilon$, we also have $|\alpha_0-x_1|\geq\varepsilon_1$. For at least one of $j=2$ or 3, say $j=2$,

$$|\Delta(\alpha_{k'})-m_2(\alpha_{k'})|<\frac{1}{k'}.$$

If $\alpha_0\neq 0$, by continuity of $\Delta(\alpha)$ and $m_2(\alpha)$, we shall have $\Delta(\alpha_0)=m_2(\alpha_0)$ which contradicts the fact that $\Delta(\alpha)$ does not coincide with $m_2(\alpha)$ except $\alpha=x_2$ or $\alpha=x_1$ when $|z|>1$. It is impossible that $\alpha_0=0$ and $|z_0|\geq 1+\varepsilon$, since $\Delta(\alpha_{k'})\to 1/(|z_0|^2-1)$ while $\min_{j=2,3}|m_j(\alpha_{k'})|\to\infty$. It is also impossible that $\alpha_0 = 0$ and $|z_0| \le 1 - \varepsilon$, since in this case, we should have $\mathrm{Re}(\Delta(\alpha_{k'})) \to +\infty$, $m_2(\alpha_{k'}) \to 1/(|z_0|^2 - 1)$ and $\mathrm{Re}(m_3(\alpha_{k'})) \to -\infty$ which follows from $\Delta(\alpha_{k'}) + m_2(\alpha_{k'}) + m_3(\alpha_{k'}) = -2$. This concludes the proof of (4.6).

The assertion (4.5) follows from the fact that equation (4.2) has no three identical roots for any α and z, since the second derivative of (4.2) gives $\Delta(\alpha) = -2/3$ equals neither $\Delta(x_2)$ nor $\Delta(x_1)$. The proof of Lemma 4.3 is then complete. □

PROOF OF LEMMA 4.4. For $x < 0$, we have: (1) $\Delta(x) > 0$ (real); (2) $\Delta(x) \to 0$ as $x \to -\infty$ and (3) from (4.2), as $x \uparrow 0$,

$$\begin{cases} \sqrt{|x|}\,\Delta(x) \uparrow \sqrt{1 - |z|^2}, & \text{if } |z| < 1, \\ \sqrt[3]{|x|}\,\Delta(x) \uparrow 1, & \text{if } |z| = 1, \\ \Delta(x) \uparrow \left(|z|^2 - 1\right)^{-1}, & \text{if } |z| > 1. \end{cases} \tag{4.38}$$

Thus, for any $C > 0$, the integral $\int_{-C}^{0} \Delta(x)\, dx$ exists. We have by exchanging the integration order,

$$\begin{aligned} \int_{-C}^{0} \Delta(x)\, dx &= \int_0^C \Delta(-x)\, dx = \int_0^C \int_0^\infty \frac{1}{u + x} \nu(du, z)\, dx \\ &= \int_0^\infty \left[\ln(C + u) - \ln u\right] \nu(du, z) \\ &= \ln C + \int_0^\infty \ln(1 + u/C)\nu(du, z) - \int_0^\infty \ln u\,\nu(du, z). \end{aligned} \tag{4.39}$$

Differentiating both sides with respect to s, we get

$$\begin{aligned} \frac{\partial}{\partial s} \int_0^\infty \ln u\,\nu(du, z) = \frac{\partial}{\partial s} \int_0^\infty & \ln(1 + u/C)\nu(du, z) \\ & - \int_{-C}^{0} \frac{\partial}{\partial s} \Delta(x)\, dx. \end{aligned} \tag{4.40}$$

[The reasons for the exchangability of the order of the integral and derivative are given after (4.47).]

Differentiating both sides of (4.2) with respect to s and x, we obtain

$$\frac{\partial}{\partial s} \Delta(x) \left[3\Delta^2(x) + 4\Delta(x) + \frac{x + 1 - |z|^2}{x}\right] = \frac{2s\Delta(x)}{x}, \tag{4.41}$$

and

$$\frac{\partial}{\partial x} \Delta(x) \left[3\Delta^2(x) + 4\Delta(x) + \frac{x + 1 - |z|^2}{x}\right] = \frac{\Delta(x)(1 - |z|^2) + 1}{x^2}.$$

Comparing the two equations, we get

$$\frac{\partial}{\partial s} \Delta(x) = \frac{2sx\Delta(x)}{1 + \Delta(x)(1 - |z|^2)} \frac{\partial}{\partial x} \Delta(x) = -\frac{2s}{(1 + \Delta(x))^2} \frac{\partial}{\partial x} \Delta(x), \tag{4.42}$$

where the last equality follows from the fact that

$$x = \frac{|z|^2}{(1+\Delta(x))^2} - \frac{1}{\Delta(x)(1+\Delta(x))} = -\frac{1+\Delta(x)(1-|z|^2)}{\Delta(x)(1+\Delta(x))^2}, \tag{4.43}$$

which is a solution of (4.2).

By (4.42), we obtain

$$\begin{aligned} \int_{-C}^{0} \frac{\partial}{\partial s}\Delta(x)\,dx &= -\int_{-C}^{0} \frac{\partial}{\partial x}\Delta(x)\frac{2s}{(1+\Delta(x))^2}\,dx \\ &= -2s\int_{\Delta(-C)}^{\Delta(0_-)} \frac{1}{(1+\Delta)^2}\,d\Delta \\ &= \frac{2s}{1+\Delta(0_-)} - \frac{2s}{1+\Delta(-C)}. \end{aligned} \tag{4.44}$$

Letting $x \uparrow 0$ in (4.2), we get

$$\Delta(0_-) = \begin{cases} \infty, & \text{if } |z|^2 \le 1, \\ \dfrac{1}{|z|^2-1}, & \text{if } |z|^2 > 1. \end{cases} \tag{4.45}$$

We also have $\Delta(-C) \to 0$ as $C \to \infty$. Thus, we get

$$\int_{-C}^{0} \frac{\partial}{\partial s}\Delta(x)\,dx \to -g(s,t). \tag{4.46}$$

Note that (4.42) is still true for $x > 0$. Therefore, by noticing

$$\nu(dx,z)/dx = \pi^{-1}\,\mathrm{Im}(\Delta(x)) = p(x,z),$$

we have

$$\begin{aligned} &\left|\frac{\partial}{\partial s}\int_0^\infty \ln(1+u/C)\,\nu(du,z)\right| \\ &\quad = \left|\frac{1}{\pi}\mathrm{Im}\left(\int_0^\infty \ln(1+u/C)\frac{\partial}{\partial s}(\Delta(u))\,du\right)\right| \\ &\quad = \left|\frac{1}{\pi}\mathrm{Im}\left(\int_0^{(2+|z|)^2} \ln(1+u/C)\frac{2s}{(1+\Delta(u))^2}\frac{\partial}{\partial u}\Delta(u)\,du\right)\right| \\ &\quad \le \left|\frac{2|s|(2+|z|)^2}{C\pi}\mathrm{Im}\left(\int_0^{(2+|z|)^2} \frac{1}{(1+\Delta(u))^2}\frac{\partial}{\partial u}\Delta(u)\,du\right)\right| \\ &\quad \le \frac{2|s|(2+|z|)^2}{C\pi}\int_0^\infty \frac{1}{(1+\Delta)^2}\,d\Delta \\ &\quad \le \frac{|s|(2+|z|)^2}{C} \to 0 \quad \text{as } C \to \infty. \end{aligned} \tag{4.47}$$

In the first equality above and in (4.40), the justification of the exchangability of the order of the integral and derivative follows from the dominated convergence theorem and the following facts:

(i) When $|z| > 1$, $\operatorname{Im}((\partial/\partial s)(\Delta(u))$ is continuous in u and vanishes when $u > x_2$ and $u < x_1$.

(ii) When $|z| < 1$, for $u > 0$ $\operatorname{Im}((\partial/\partial s)(\Delta(u))$ is continuous in u and vanishes when $u > x_2$, and for small u, by $u\Delta^2(u) \to -1 + |z|^2$ [see (4.2) and (4.41)],

$$\left|\operatorname{Im}\left(\frac{\partial}{\partial s}(\Delta(u))\right)\right| \le \frac{2|s\Delta(u)|}{|3u\Delta^2(u) + 4u\Delta(u) + u + 1 - |z|^2|} \le \frac{4|s|}{\sqrt{xu(1-|z|^2)^3}}$$

which is integrable w.r.t. u.

(iii) When $|z| = 1$ and u small, by $u\Delta^3(u) \to -1$,

$$\left|\operatorname{Im}\left(\frac{\partial}{\partial s}(\Delta(u))\right)\right| \le 4|s|u^{-2/3}$$

which is also integrable w.r.t. u.

The assertion (4.9) then follows from (4.40), (4.46) and (4.47) and Lemma 4.4 is proved. □

PROOF OF LEMMA 4.5. We shall prove (4.10) by employing Corollary 2.3 of Bai (1993a). For all $z \in T$, the supports of $\nu(\cdot, z)$ are commonly bounded. Therefore, we may select a constant N such that, for some absolute constant C,

$$\begin{aligned}
&\|\nu_n(\cdot, z) - \nu\cdot, z)\| \\
&\quad\le C\left(\int_{|x|\le N} |\Delta_n(\alpha) - \Delta(\alpha)|\,dx\right. \\
&\qquad\left. + y_n^{-1}\sup_x \int_{|y|\le 2y_n} |\nu(x+y, z) - \nu(x, z)|\,dy\right) \\
&\quad\le C\left(\int_{|x|\le N} |\Delta_n(\alpha) - \Delta(\alpha)|\,dx + y_n^{1/2}\right),
\end{aligned} \tag{4.48}$$

where the last step follows from (4.4).

Denoted by $m_1(\alpha) = \Delta(\alpha)$, $m_2(\alpha)$ and $m_3(\alpha)$ the three solutions of the equation (4.2). Note that $\Delta(\alpha)$ is analytic in α for $\operatorname{Im}(\alpha) > 0$. By a suitable selection, the three solutions are all analytic in α on the upper half complex plane.

By Lemma 4.3, there are constant ε_0 and ε_1 such that (4.5)–(4.8) hold. By Lemma 4.1, there is an n_0 such that for all $n \ge n_0$,

$$|(\Delta_n - m_1)(\Delta_n - m_2)(\Delta_n - m_3)| = o(\delta_n) \le \tfrac{4}{27}\varepsilon_0^3\delta_n. \tag{4.49}$$

Now, choose an $\alpha_0 = x_0 + iy_0$ with $|x_0| \le N$, $y_0 > 0$ and $\min_{k=1,2}(|x_0 - x_k|) \ge \varepsilon_1$. For a fixed $z \in T$, as argued earlier, $\Delta_n(\alpha_0)$ converges to $\Delta(\alpha_0)$ when n goes to infinity along some subsequence. Then, for infinitely many $n > n_0$, $|\Delta_n(\alpha_0) - \Delta(\alpha_0)| < \varepsilon_0/3$. Hence,

$$\min_{k=2,3}\left(|\Delta_n(\alpha_0) - m_k(\alpha_0)|\right)$$
$$\ge \min_{k=2,3}\left(|\Delta(\alpha_0) - m_k(\alpha_0)|\right) - |\Delta_n(\alpha_0) - \Delta(\alpha_0)|\big) > \tfrac{2}{3}\varepsilon_0.$$

This and (4.49) imply, for infinitely many n,

$$|\Delta_n(\alpha_0) - \Delta(\alpha_0)| = o(\delta_n) \le \tfrac{1}{3}\varepsilon_0\delta_n. \tag{4.50}$$

Let n_0 be also such that $2/(y_n^2 n_0) + \frac{1}{3}\varepsilon_0 n_0^{-1/120} < \varepsilon_0/3$. We claim that (4.50) is true for all $n \ge n_0$. In fact, if (4.50) is true for some $n > n_0$, then

$$|\Delta_{n-1}(\alpha_0) - \Delta(\alpha_0)| \le |\Delta_{n-1}(\alpha_0) - \Delta_n(\alpha_0)| + |\Delta_n(\alpha_0) - \Delta(\alpha_0)|$$
$$< 2/(y_n n) + \tfrac{1}{3}\varepsilon_0 n^{-1/120} < \varepsilon_0/3.$$

Here we have used the trivial fact that $\|\nu_n(\cdot, z) - \nu_{n-1}(\cdot, z)\| \le 2/n$ which implies $|\Delta_{n-1}(\alpha_0) - \Delta_n(\alpha_0)| \le 2/(y_n n)$. This shows that

$$\min_{k=2,3}\left(|\Delta_{n-1}(\alpha_0) - m_k(\alpha_0)|\right) > \tfrac{2}{3}\varepsilon_0,$$

which implies that (4.50) is true for $n-1$. This completes the proof of our assertion.

Now, we claim that (4.50) is true for all $n > n_0$ and $|\alpha| \le N$, $\min_{k=1,2}(|x - x_k|) \ge \varepsilon_1$, that is,

$$|\Delta_n(\alpha) - \Delta(\alpha)| \le o(\delta_n) \le \tfrac{1}{3}\varepsilon_0\delta_n. \tag{4.51}$$

By (4.6) and (4.49), we conclude that (4.51) is equivalent to

$$\min_{k=2,3}\left(|\Delta_n(\alpha) - m_k(\alpha)|\right) > \tfrac{2}{3}\varepsilon_0. \tag{4.52}$$

Note that both Δ_n and $m_j(\alpha)$, $j = 1, 2, 3$, are continuous functions in both α and z. Therefore, on the boundary of the set of points (α, z) at which (4.51) does not hold, we should have $|\Delta_n(\alpha) - \Delta(\alpha)| = \frac{1}{3}\varepsilon_0\delta_n$ and $\min_{k=2,3}(|\Delta_n(\alpha) - m_k(\alpha)|) = \frac{2}{3}\varepsilon_0$. This is impossible because these two equalities contradict (4.6).

For $|\alpha - x_k| \le \varepsilon_1$, $k = 1$ or 2, (4.5), (4.7) and (4.8) imply that

$$|\Delta_n(\alpha) - \Delta(\alpha)| \le o\left(\delta_n/\sqrt{|\alpha - x_k|}\right). \tag{4.53}$$

This, together with (4.48) and (4.51), implies (4.10). The proof of Lemma 4.5 is complete. □

5. Proof of (1.4). In this section, we shall show that probability 1,

$$\int_{z\in T}\left|\int_0^{\varepsilon_n} \ln x\,\nu_n(dx, z)\right| dt\,ds \to 0, \tag{5.1}$$

where $\varepsilon_n = \exp(-n^{1/120})$.

Denote by $\mathbf{Z}_1$ and $\mathbf{Z}$ the matrix of the first two columns of $\mathbf{R}$ and that formed by the last $n-2$ columns. Let $\Lambda_1 \le \cdots \le \Lambda_n$ denote the eigenvalues of the matrix $\mathbf{R}^*\mathbf{R}$ and let $\eta_1 \le \cdots \le \eta_{n-2}$ denote the eigenvalues of $\mathbf{Z}^*\mathbf{Z}$. Then, for any $k \le n-2$, we have $\Lambda_k \le \eta_k \le \Lambda_{k+2}$ and $\det(\mathbf{R}^*\mathbf{R}) = \det(\mathbf{Z}^*\mathbf{Z})\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)$, where $\mathbf{Q} = \mathbf{I} - \mathbf{Z}(\mathbf{Z}^*\mathbf{Z})^{-1}\mathbf{Z}^*$. This identity can be written as

$$\sum_{k=1}^{n} \ln(\Lambda_k) = \ln\bigl(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)\bigr) + \sum_{k=1}^{n-2} \ln(\eta_k).$$

If l is the smallest integer such that $\eta_l \ge \varepsilon_n$, then $\Lambda_{l-1} < \varepsilon_n$ and $\Lambda_{l+2} > \varepsilon_n$. Therefore, we have

$$\begin{aligned} 0 > \int_0^{\varepsilon_n} \ln x \nu_n(dx, z) &= \frac{1}{n} \sum_{\Lambda_k < \varepsilon_n} \ln \Lambda_k \\ &\ge \frac{1}{n} \min\bigl\{\ln\bigl(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)\bigr), 0\bigr\} + \frac{1}{n} \sum_{\eta_k < \varepsilon_n} \ln \eta_k - \frac{2}{n} \ln\bigl(\max(\Lambda_n, 1)\bigr). \end{aligned} \tag{5.2}$$

To prove (5.1), we first estimate the integral of $(1/n)\ln(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1))$ with respect to s and t. Note that with probability one, the rank of the matrix $\mathbf{Q}$ is 2. Hence, there are two orthogonal complex unit vectors γ_1 and γ_2 such that $\mathbf{Q} = \gamma_1\gamma_1^* + \gamma_2\gamma_2^*$. Denote the two column vectors of $\mathbf{Z}_1$ by $\mathbf{r}_1$ and $\mathbf{r}_2$. Then we have

$$\frac{1}{n} \ln\bigl(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)\bigr) = \frac{1}{n} \ln\bigl(|\gamma_1^*\mathbf{r}_1\gamma_2^*\mathbf{r}_2 - \gamma_2^*\mathbf{r}_1\gamma_1^*\mathbf{r}_2|^2\bigr).$$

Define the random sets

$$E = \bigl\{(s,t) \colon |\gamma_1^*\mathbf{r}_1\gamma_2^*\mathbf{r}_2 - \gamma_2^*\mathbf{r}_1\gamma_1^*\mathbf{r}_2| \ge n^{-14}, |\boldsymbol{\xi}_1| \le n, |\boldsymbol{\xi}_2| \le n\bigr\}$$

and

$$F = \bigl\{(s,t) \colon |\gamma_1^*\mathbf{r}_1\gamma_2^*\mathbf{r}_2 - \gamma_2^*\mathbf{r}_1\gamma_1^*\mathbf{r}_2| < n^{-14}, |\boldsymbol{\xi}_1| \le n, |\boldsymbol{\xi}_2| \le n\bigr\}.$$

It is trivial to see that

$$P\bigl(|\boldsymbol{\xi}_1| > n \quad \text{or} \quad |\boldsymbol{\xi}_2| > n\bigr) < 2n^{-2}. \tag{5.3}$$

When $|\boldsymbol{\xi}_1| \le n$ and $|\boldsymbol{\xi}_2| \le n$, we have $\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1) = |\gamma_1^*\mathbf{r}_1\gamma_2^*\mathbf{r}_2 - \gamma_2^*\mathbf{r}_1\gamma_1^*\mathbf{r}_2|^2 \le 4(n+M)^4$. Thus,

$$\frac{1}{n} \int_{z \in T} \bigl|I_E \ln\bigl(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)\bigr)\bigr| \, dt \, ds \le Cn^{-1} \ln n \to 0. \tag{5.4}$$

On the other hand, for any $\varepsilon > 0$, we have

$$\begin{aligned} &P\left(\frac{1}{n} \int_{z \in T} \bigl|I_F \ln\bigl(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1)\bigr)\bigr| \, dt \, ds \ge \varepsilon\right) \\ &\qquad \le \frac{1}{\varepsilon n} \int_{z \in T} E\Bigl| I_F \ln\bigl(|\gamma_1^*\mathbf{r}_1\gamma_2^*\mathbf{r}_2 - \gamma_2^*\mathbf{r}_1\gamma_1^*\mathbf{r}_2|^2\bigr)\Bigr| \, dt \, ds. \end{aligned} \tag{5.5}$$

Note that the elements of $\sqrt{n}\,\mathbf{r}_1$ and $\sqrt{n}\,\mathbf{r}_2$ are independent of each other and the joint densities of their real and imaginary parts have a common upper bound K_d. Also, they are independent of γ_1 and γ_2. Therefore, by Corollary A.2, the conditional joint density of the real and imaginary parts of $\sqrt{n}\,\gamma_1^*\mathbf{r}_1$, $\sqrt{n}\,\gamma_2^*\mathbf{r}_2$ $\sqrt{n}\,\gamma_2^*\mathbf{r}_1$ and $\sqrt{n}\,\gamma_1^*\mathbf{r}_2$, when γ_1 and γ_2 are given, is bounded by $(2K_d n)^4$. Hence, the conditional joint density of the real and imaginary parts of $\gamma_1^*\mathbf{r}_1$, $\gamma_2^*\mathbf{r}_2$, $\gamma_2^*\mathbf{r}_1$ and $\gamma_1^*\mathbf{r}_2$, when γ_1 and γ_2 are given, is bounded by $K_d^4 2^4 n^8$. Set $\mathbf{x} = (\gamma_1^*\mathbf{r}_1, \gamma_2^*\mathbf{r}_1)'$ and $\mathbf{y} = (\mathbf{r}_2^*\gamma_2, -\mathbf{r}_2^*\gamma_1)'$. Note that by Corollary A.2, the joint density of $\mathbf{x}$ and $\mathbf{y}$ is bounded by $K_d^4 2^r n^8$.

If $|\boldsymbol{\xi}_1| \le n$, $|\boldsymbol{\xi}_2| \le n$, then $\max(|\mathbf{x}|, |\mathbf{y}|) \le n + |z| \le n + M$. Applying Lemma A.3 with $f(t) = \ln t$, $M = \mu = 1$, we obtain

$$\text{(5.6)}\qquad \begin{aligned} &\frac{2}{n}\int_{z\in T}\left|\mathrm{E}\Big(I_{(|\mathbf{x}^*\mathbf{y}|<n^{-14},\,|\boldsymbol{\xi}_1|\le n,\,|\boldsymbol{\xi}_2|\le n)}\ln(|\mathbf{x}^*\mathbf{y}|)\Big|\gamma_1,\gamma_2\Big)\right|\,dt\,ds\\ &\qquad \le Cn^{12}n^{-14} \le Cn^{-2}, \end{aligned}$$

for some positive constant C.

From (5.3), (5.5) and (5.6), it follows that

$$\text{(5.7)}\qquad \frac{1}{n}\int_{z\in T}|I_F\ln(\det(\mathbf{Z}_1^*\mathbf{Q}\mathbf{Z}_1))|\,dt\,ds \to 0 \quad \text{a.s.}$$

Next, we estimate the second term in (5.2). We have

$$\text{(5.8)}\qquad \begin{aligned} \frac{1}{n}\left|\sum_{\eta_k<\varepsilon_n}\ln(\eta_k)\right| &\le n^{-119/120}\varepsilon_n\sum_{k=1}^{n-2}\frac{1}{\eta_k}\\ &= n^{-119/120}\varepsilon_n\,\mathrm{tr}\big((\mathbf{Z}^*\mathbf{Z})^{-1}\big)\\ &= n^{-119/120}\varepsilon_n\sum_{k=3}^{n}\frac{1}{\mathbf{r}_k^*\mathbf{Q}_k\mathbf{r}_k}\\ &= n^{-119/120}\varepsilon_n\sum_{k=3}^{n}\frac{1}{|\mathbf{r}_k^*\gamma_{k1}|^2+|\mathbf{r}_k^*\gamma_{k2}|^2+|\mathbf{r}_k^*\gamma_{k3}|^2}, \end{aligned}$$

where for each k, γ_{kj}, $j = 1, 2, 3$, are orthonormal complex vectors such that $\mathbf{Q}_k = \gamma_{k1}\gamma_{k1}^* + \gamma_{k2}\gamma_{k2}^* + \gamma_{k3}\gamma_{k3}^*$ which is the projection matrix onto the orthogonal complement of the space spanned by the $3, \dots, k-1, k+1, \dots, n$ columns of $\mathbf{R}(z)$.

As in the proof of (5.7), one can show that the conditional joint density of the real and imaginary parts of $\mathbf{r}_k^*\gamma_{k1}$, $\mathbf{r}_k^*\gamma_{k2}$ and $\mathbf{r}_k^*\gamma_{k3}$ when γ_{kj}, $j = 1, 2, 3$ are given, is bounded by $CK_d n^{12}$. Therefore, we have

$$\text{(5.9)}\qquad \begin{aligned} &n^{-119/120}\varepsilon_n\sum_{k=3}^{n}\int_{z\in T}\mathrm{E}\left(\frac{dt\,ds}{|\mathbf{r}_k^*\gamma_{k1}|^2+|\mathbf{r}_k^*\gamma_{k2}|^2+|\mathbf{r}_k^*\gamma_{k3}|^2}\right)\\ &\quad\le Cn^{-119/120}\varepsilon_n\left(K_d^3 n^{13}\int\cdots\int_{u_1^2+\cdots+u_6^2<1}\frac{du_1\cdots du_6}{u_1^2+\cdots+u_6^2}+n\right)\\ &\quad\le Cn^{13}\varepsilon_n \quad \text{by a polar transformation}\\ &\quad\le Cn^{-2}. \end{aligned}$$

Therefore, by the Borel–Cantelli lemma,

$$n^{-119/120}\varepsilon_n \sum_{k=3}^{n} \int_{z\in T} \left(\frac{dt\,ds}{|\mathbf{r}_k^*\gamma_{k1}|^2 + |\mathbf{r}_k^*\gamma_{k2}|^2 + |\mathbf{r}_k^*\gamma_{k3}|^2} \right) \to 0 \quad \text{a.s.}$$

and hence, with probability 1,

$$\int_{z\in T} \frac{1}{n} \sum_{\eta_k<\varepsilon_n} \ln(\eta_k)\, dt\,ds \le n^{-119/120}\varepsilon_n \sum_{k=1}^{n-2} \int_{z\in T} \frac{dt\,ds}{\eta_k} \to 0. \tag{5.10}$$

Finally, we estimate the integral of the third term in (5.2). By Yin, Bai and Krishnaiah (1988), we have $\Lambda_n \le (\|\boldsymbol{\Xi}_n\| + |z|)^2 \to (2+|z|)^2$, a.s. We conclude that

$$\frac{2}{n} \int_{z\in T} \ln\bigl(\max(\Lambda_n, 1)\bigr)\, dt\,ds \to 0 \quad \text{a.s.} \tag{5.11}$$

Hence, (5.1) follows from (5.7), (5.10) and (5.11).

6. Proof of Theorem 1.1. In Section 3, the problem is reduced to showing (3.10). Recalling the definitions of $g_n(s,t)$ and $g(s,t)$, we have by integration by parts,

$$\begin{aligned} &\left| \int_{z\in T} (g_n(s,t) - g(s,t)) \exp(ius + itv)\, dt\,ds \right| \\ &\quad = \Bigg| -\int_{z\in T} iu\tau(s,t)\, dt\,ds \\ &\qquad + \int_{|t|\le A^2} [\tau(A,t)\, dt - \tau(-A,t)]\, dt \\ &\qquad + \int_{|t|\le 1+\varepsilon} \Bigl[\tau\Bigl(\sqrt{(1+\varepsilon)^2 - t^2}, t\Bigr)\, dt \\ &\qquad\qquad - \tau\Bigl(-\sqrt{(1+\varepsilon)^2 - t^2}, t\Bigr)\Bigr]\, dt \\ &\qquad + \int_{|t|\le 1-\varepsilon} \Bigl[\tau\Bigl(\sqrt{(1-\varepsilon)^2 - t^2}, t\Bigr)\, dt \\ &\qquad\qquad - \tau\Bigl(-\sqrt{(1-\varepsilon)^2 - t^2 1}, t\Bigr)\Bigr]\, dt \Bigg|, \end{aligned} \tag{6.1}$$

where

$$\tau(s,t) = \exp(ius + itv) \int_0^\infty \ln x\bigl(\nu_n(dx, z) - \nu(dx, z)\bigr).$$

When A is large enough, with probability 1, for all large n, the support of $\nu_n(\cdot, \pm A + it)$ is uniformly bounded by $(A-3)^2 > 1$ from the left and by

$(A + A^2 + 3)^2$ from the right. By Lemma 4.5, we have

$$\left|\int_{|t|\le A^2}\tau(\pm A, t)\,dt\right|$$
$$\le \int_{|t|\le A^2}\left|\int_{(A-3)^2}^{(A+A^2+3)^2}\ln x\big(\nu_n(dx, \pm A + it) - \nu(dx, \pm A + it)\big)\right|dt$$
$$\to 0 \quad \text{a.s.}$$

Let $\varepsilon_n = \exp(-n^{1/120})$. In Section 5, we proved that

$$\int_{z\in T}\left|\int_0^{\varepsilon_n}\ln x\nu_n(dx, z)\right|dt\,ds \to 0, \quad \text{a.s.}$$

By (4.4), we have

$$\int_{z\in T}\left|\int_0^{\varepsilon_n}\ln x\nu(dx, z)\right|dt\,ds \to 0.$$

By Lemma 4.5, we have

$$\int_{z\in T}\left|\int_{\varepsilon_n}^{(A+A^2+3)^2}\ln x\big(\nu_n(dx, z) - \nu(dx, z)\big)\right|dt\,ds$$
$$\le 4CA^3|\ln(\varepsilon_n)|\max_{z\in T}\|\nu_n(\cdot, z) - \nu(\cdot, z)\| \to 0.$$

This proves that

$$iu\int_{z\in T}\tau(s, t)\,dt\,ds \to 0.$$

Similarly, we can prove that

$$\int_{|t|\le 1\pm\varepsilon}\left[\pm\tau\left(\sqrt{(1\pm\varepsilon)^2 - t^2}, t\right)\right]dt \to 0.$$

The proof of Theorem 1.1 is complete. □

7. Comments and extensions.

7.1. *Relaxation of conditions assumed in Theorem* 1.1.

7.1.1. *On the moment of the underlying distribution*. Reviewing the definition of ε_n and checking the proofs given in Sections 5 and 6, one finds that $|\ln(\varepsilon_n)| = \sqrt{1/\delta_n}$ and $\max_{z\in T}\|\nu_n(\cdot, z) - \nu(\cdot, z)\| = o(\sqrt{\delta_n})$. Hence, (1.3) is always true for any choice of $\delta_n(\to 0)$. The rate of ε_n is required to be $o(n^{-M})$ for some large M, for the proof of (5.9). Reexaming the proofs of Lemmas 4.1, 4.5 and A.4, one may find that if $\mathrm{E}|x_{11}|^{4+\varepsilon} < \infty$, then $\max_{z\in T}\|\nu_n(\cdot, z) - \nu(\cdot, z)\| = o(n^{-\beta})$ for some $\beta > 0$. Therefore, the circular law is true when the moment condition in Theorem 1.1 is reduced to the existence of the $4 + \varepsilon$th moment. The details of the proof are omitted.

7.1.2. *On the smoothness of the underlying distribution*. The purpose of this subsection is to consider the circular law for real random matrices whose entries have a bounded density. The circular law for this case does not follow from Theorem 1.1 since the joint distribution of the real and imaginary parts of the entries does not have a joint two-dimensional density. In the following, we shall consider a more general case where the conditional density of one linear combination of the real and imaginary parts of the entry when another is given is uniformly bounded. Without loss of generality, we assume that the two linear combinations are $\mathrm{Re}(x_{11})\cos(\theta) + \mathrm{Im}(x_{11})\sin(\theta)$ and $\mathrm{Re}(x_{11})\sin(\theta) - \mathrm{Im}(x_{11})\cos(\theta)$. Note that the proof of the circular law for the matrix $\mathbf{X}$ is equivalent to that for the matrix $e^{i\theta}\mathbf{X}$ under the condition that the conditional density of the real part when the imaginary part is given is uniformly bounded. We shall establish the following theorem.

THEOREM 7.1. *Assume that the conditional density of the real part of the entries of* $\mathbf{X}$ *when given the imaginary part is uniformly bounded and assume that the entries have finite* $4 + \varepsilon$ *moment. Then the circular law holds.*

SKETCH OF THE PROOF. A review of the proof of Theorem 1.1 reveals that it is sufficient to prove the inequalities (5.7) and (5.10) under the conditions of Theorem 7.1. We start the proof of (5.7) from (5.5). Rewrite

$$\ln\left(|\mathbf{y}^* \mathbf{x}|^2\right) = \ln\left(|\mathbf{y}|^2\right) + \ln\left(|\tilde{\mathbf{y}}^* \mathbf{x}|^2\right),$$

where $\tilde{\mathbf{y}} = \mathbf{y}/|\mathbf{y}|$.

Denote by $\mathbf{x}_{jr}$ and $\mathbf{x}_{ji}$ the real and imaginary parts of the vector $\mathbf{x}_j$. Without loss of generality, we assume that $|\gamma_{ir}| \geq 1/\sqrt{2}$. Then, we have

$$\begin{aligned}|\mathbf{y}|^2 &= \mathbf{r}_2^*(\gamma_1\gamma_1^* + \gamma_2\gamma_2^*)\mathbf{r}_2\\ &\geq (\gamma_{1r}'\mathbf{r}_{2r} + \gamma_{1i}'\mathbf{r}_{2i})^2.\end{aligned}$$

Applying Lemma A.1, we find that the conditional density of $\gamma_{1r}'\mathbf{r}_{2r} + \gamma_{1i}'\mathbf{r}_{2i}$ when γ_1, γ_2 and $\mathbf{r}_{2i}$ are given is bounded by $CK_d n$. Therefore, by Lemma A.3,

$$\begin{aligned}&\frac{1}{n}\int_{z\in T}\left|\mathrm{E}\left(I_{(|\mathbf{y}|^2 < n^{-14},\, |\boldsymbol{\xi}_2| \leq n)} \ln\left(|\mathbf{y}|^2\right)\right)\right| dt\, ds\\ &\quad\leq \frac{1}{n}\int_{z\in T}\mathrm{E}\left|\mathrm{E}\left(I_{(|\gamma_{1r}'\mathbf{r}_{2r}+\gamma_{1i}'\mathbf{r}_{2i}|^2 < n^{-14},\, |\boldsymbol{\xi}_2|\leq n)}\right.\right.\\ &\qquad\qquad \left.\left.\times\ln\left(|\gamma_{1r}'\mathbf{r}_{2r} + \gamma_{1i}'\mathbf{r}_{2i}|^2\right)\Big|\gamma_1, \gamma_2, \mathbf{r}_{2i}\right)\right| dt\, ds\\ &\quad\leq CK_d\int_0^{n^{-7}} \ln x\, dx \leq Cn^{-7}\ln n,\end{aligned} \tag{7.1}$$

for some positive constant C.

Rewrite

$$|\tilde{\mathbf{y}}^* \mathbf{x}|^2 = (\beta_1'\mathbf{r}_{1r} + \zeta_1'\mathbf{r}_{1i})^2 + (\beta_2'\mathbf{r}_{1r} + \zeta_2'\mathbf{r}_{1i})^2,$$

where

$$\beta_1 = (\tilde{\mathbf{y}}_r', \mathbf{y}_i')(\gamma_{1r}, \gamma_{2r}, -\gamma_{1i}, \gamma_{2i})', \qquad \beta_2 = (\tilde{\mathbf{y}}_r', \tilde{\mathbf{y}}_i')(-\gamma_{1i}, -\gamma_{2i}, \gamma_{1r}, -\gamma_{2r})'$$
$$\zeta_1 = (\tilde{\mathbf{y}}_r', \tilde{\mathbf{y}}_i')(\gamma_{1i}, \gamma_{2i}, \gamma_{1r}, -\gamma_{2r})', \qquad \zeta_2 = (\tilde{\mathbf{y}}_r', \tilde{\mathbf{y}}_i')(\gamma_{1r}, -\gamma_{2r}, -\gamma_{1i}, -\gamma_{2i})'.$$

It can be verified that $|\beta_1|^2 + |\beta_2|^2 = 1$. Thus, we may assume that $|\beta_1| \geq 1/\sqrt{2}$. By Lemma A.1, the conditional density of $\beta_1'\mathbf{r}_{1r} + \zeta_1'\mathbf{r}_{1i}$ when γ_1, γ_2, $\mathbf{y}$ and $\mathbf{r}_{1i}$ are given is bounded by $2K_d n$. Consequently, we can prove that

$$\begin{aligned}
&\frac{1}{n}\int_{z\in T} \mathrm{E}\Big|\mathrm{E}\Big(I_{(|\tilde{\mathbf{y}}^*\mathbf{x}|<n^{-7})}\ln\big(|\tilde{\mathbf{y}}^*\mathbf{x}|^2\big)\Big|\gamma_1,\gamma_2,\mathbf{y},\mathbf{r}_{1i}\Big)\Big|\,dt\,ds\\
&\quad\leq \frac{1}{n}\int_{z\in T}\mathrm{E}\Big|\mathrm{E}\Big(I_{(|\beta_1'\mathbf{r}_{1r}+\zeta_1' r_{1i}|^2|<n^{-7})}\ln\big(|\beta_1'\mathbf{r}_{1r}+\zeta_1'\mathbf{r}_{1i}|^2\big)\Big|\gamma_1,\gamma_2,\mathbf{y},\mathbf{r}_{1i}\Big)\Big|\,dt\,ds\\
&\quad\leq CK_d\int_0^{n^{-7}}\ln x\,dx \leq Cn^{-7}\ln n.
\end{aligned}$$

This, together with (7.1), completes the proof of (5.7).

Now, we prove (5.10). For each k, consider the $2n \times 6$ matrix $\mathbf{A}$ whose first three columns are $(\gamma_{jkr}', -\gamma_{jki}')'$, $j = 1, 2, 3$, and other three columns are $(\gamma_{jki}', \gamma_{jkr}')'$. Since γ_{kj} are orthonormal, we have $\mathbf{A}'\mathbf{A} = \mathbf{I}_6$. Using the same approach as the proof of Lemma A.1, one may select a 6×6 submatrix $\mathbf{A}_1$ of $\mathbf{A}$ such that $|\det(\mathbf{A}_1)| \geq n^{-3}$. Within the six rows of $\mathbf{A}_1$, either three rows come from the first n rows of $\mathbf{A}$ or three come from the last n rows. Without loss of generality, assume that $\mathbf{A}_1$ has three rows coming from the first n rows of $\mathbf{A}$. Then, consider the Laplace expansion of the determinant of $\mathbf{A}_1$ with respect to the first three rows. Within the 20 terms, we may select one whose absolute value is not less than $\frac{1}{20}n^{-3}$. This term is the product of a minor from the first three rows of $\mathbf{A}_1$ and its cofactor. Since the absolute value of the entries of $\mathbf{A}$ is not greater than 1, the absolute value of the cofactor is not greater than 6. Therefore, the absolute value of the minor is not less than $\frac{1}{20}n^{-3}$. Suppose the three columns of the minor come from the first, second and fourth columns of $\mathbf{A}$, that is, from γ_{1kr}, γ_{2kr} and γ_{1ki} (the proof of the other 19 cases is similar). Then, as in the proof of Lemma A.1, one can prove that the conditional joint density of $\gamma_{1kr}'\mathbf{r}_{kr}$, $\gamma_{2kr}'\mathbf{r}_{kr}$ and $\gamma_{1ki}'\mathbf{r}_{kr}$ when γ_{jk} and $\mathbf{r}_{ki}$ are given is uniformly bounded by $120K_d n^{4.5}$. Finally, from (5.8), we have

$$\begin{aligned}
&\varepsilon_n\sum_{k=3}^n \frac{1}{|\mathbf{r}_k^*\gamma_{k1}|^2 + |\mathbf{r}_k^*\gamma_{k2}|^2 + |\mathbf{r}_k^*\gamma_{k3}|^2}\\
&\quad\leq \varepsilon_n\sum_{k=3}^n\frac{1}{(\mathbf{r}_{kr}'\gamma_{k1r}+\mathbf{r}_{ki}'\gamma_{k1i})^2 + (\mathbf{r}_{kr}'\gamma_{k2r}+\mathbf{r}_{ki}'\gamma_{k2i})^2 + (\mathbf{r}_{kr}'\gamma_{k1i}-\mathbf{r}_{ki}'\gamma_{k1r})^2}.
\end{aligned}$$

Using this and the same approach as in Section 5, one may prove that the right-hand side of the above tends to zero almost surely. Thus, (5.10) is proved and consequently, Theorem 7.1 follows. □

7.1.3. *Extension to the nonidentical case.* Reviewing the proofs of Theorem 1.1, one finds that the moment condition and the distributional identity of the entries of the random matrix were used only in Lemma A.4, for establishing the uniform convergence rate of certain quadratic forms. One requirement for this purpose is that the variables can be truncated at $n^{1/3}$ (actually, $n^{1/2-\varepsilon}$ is good enough as discussed in subsection 7.1.1). Two other requirements are $\max_{j_1, j_2}|\mathrm{E}(X_{m, j_1, j_2})| = o(n^{-1})$ and $\max_{j_1, j_2}|\mathrm{E}(|X_{m, j_1, j_2}|^2) - 1| = o(1)$. Therefore, we have the following theorem.

THEOREM 7.2. *In additional to the smoothness condition assumed in Theorem* 1.1, *we further assume that*

$$\max_{j_1, j_2}|\mathrm{E}\big(X_{j,k}\big) I_{(|X_{j,k}| \le n^{1/2-\varepsilon})}| = o\big(n^{-1}\big),$$
$$\max_{j_1, j_2}|\mathrm{E}\Big(|X_{j_1, j_2}|^2 I_{(|X_{j,k}| \le n^{1/2-\varepsilon})}\Big) - 1| = o(1) \tag{7.2}$$

and

$$\limsup_{n\to\infty} \mathrm{P}\bigg(\bigcup_{k, j \le n}\big(|x_{nkj}| > n^{1/2-\varepsilon}\big)\bigg) = 0. \tag{7.3}$$

Then the circular law is true.

A sufficient condition for (7.2) and (7.3) is the following: in addition to $\mathrm{E}(X_{jk}) = 0$, $\mathrm{E}(|X_{jk}|^2) = 1$,

$$\max_{k, j} \mathrm{E}|x_{kj}|^{4+\varepsilon} < \infty \quad \text{if all } x_{kj} \text{ come from a double array,}$$

or

$$\max_{n, k, j} \mathrm{E}|x_{nkj}|^{6+\varepsilon} < \infty \quad \text{if } x_{nkj} \text{ depends on } n.$$

7.2. *Spectral radius.* As mentioned earlier, Bai and Yin (1986) and Geman (1986), proved that with probability 1, the upper limit of the spectral radius of Ξ_n is not greater than 1. Combining this result together with Theorem 1.1, it follows immediately that with probability 1, the spectral radius of Ξ_n converges to 1. In fact, we can get more, that is, under the conditions of Theorem 1.1, we have, with probability 1,

$$\lim_{n\to\infty} \inf_{a^2+b^2=1} \max_{k \le n}\big(a\,\mathrm{Re}(\lambda_k) + b\,\mathrm{Im}(\lambda_k)\big)$$
$$= \lim_{n\to\infty} \sup_{a^2+b^2=1} \max_{k \le n}\big(a\,\mathrm{Re}(\lambda_k) + b\,\mathrm{Im}(\lambda_k)\big) = 1.$$

APPENDIX

Elementary lemmas.

A1. Lemmas on densities or expectations of functions of random variables.

LEMMA A.1. *Let* $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ *be a* $p \times n$ *real random matrix of* n *independent column vectors whose probability densities have a common bound* K_d *and let* $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$, $(k < n)$ *be* k *orthogonal real unit* n*-vectors. Then, the joint density of the random* p*-vectors* $\mathbf{y}_j = \mathbf{X}\boldsymbol{\alpha}_j$, $j = 1, \ldots, k$, *is bounded by* $K_d^k n^{kp/2}$.

PROOF. Write $\mathbf{C} = (\boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_k)'$ and let $\mathbf{C}(j_1, \ldots, j_k)$ denote the $k \times k$ submatrix formed by the $j_1 \cdots j_k$th columns of C. By Bennett's formula, we have

$$\sum_{1 \le j_1 < \cdots < j_k \le n} \det{}^2\big(\mathbf{C}(j_1, \ldots, j_k)\big) = \det(\mathbf{C}'\mathbf{C}) = 1.$$

Thus, we may select $1 \le j_1 < \cdots < j_k \le n$, say, $1, 2, \ldots, k$ for simplicity, such that $|\det(\mathbf{C}(1, \ldots, k))| \ge n^{-k/2}$. Let $\mathbf{C}_1$ and $\mathbf{C}_2$ denote the submatrices of the first k and last $n - k$ columns of $\mathbf{C}$ and let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the submatrices of the first k and last $n - k$ columns of $\mathbf{X}$, respectively. Furthermore, denote by $\mathbf{c}_1, \ldots, \mathbf{c}_k$, the row vectors of the matrix $\mathbf{C}^{-1}(1, 2, \ldots, k)$. Then, the joint density of $\mathbf{y}_1, \ldots, \mathbf{y}_k$ is given by

$$p(\mathbf{y}_1, \ldots, \mathbf{y}_k) = |\det{}^{-p}(C(1, 2, \ldots, k))|\mathrm{E}\left(\prod_{i=1}^{k} f_i((\mathbf{y} - \mathbf{X}_2\mathbf{C}_2')\mathbf{c}_i)\right) \le K_d^k n^{kp/2},$$

where $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)'$. The proof of the lemma is complete. □

For the complex case, we have the following corollary.

COROLLARY A.2. *If the vectors and matrices in Lemma* A.1 *are assumed to be complex and the joint density of the real and imaginary parts of* $\mathbf{x}_j$ *are uniformly bounded by* K_d, *then the joint density of the real and imaginary parts of* $\mathbf{y}_1, \ldots, \mathbf{y}_k$ *is bounded by* $K_d^{2k} n^{2kp}$.

LEMMA A.3. *Suppose that* $f(t)$ *is a function such that* $\int_0^\delta t|f(t)|\,dt \le M\delta^\mu$, *for some* $\mu > 0$ *and all small* δ. *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two complex random* k*-vectors* $(k > 1)$ *whose joint density of the real and imaginary parts of* $\mathbf{x}$ *and* $\mathbf{y}$ *is bounded by* K_d. *Then,*

$$\mathrm{E}\big(f(|\mathbf{x}^*\mathbf{y}|)I(|\mathbf{x}^*\mathbf{y}| < \delta, |\mathbf{x}| \le K_e, |\mathbf{y}| \le K_e\big) \le C_k M\delta^\mu K_d K_e^{4k-4}, \tag{A.1}$$

where C_k *is a positive constant depending on* k *only.*

PROOF. Note that the measure of $\mathbf{x} = 0$ is zero. For each $\mathbf{x} \ne 0$, define a unitary $k \times k$ matrix U with $\mathbf{x}/|\mathbf{x}|$ as its first column. Now, make a change of variables $\mathbf{u} = \mathbf{x}$ and $\mathbf{v} = \mathbf{U}^* \mathbf{y}$. It is known that the Jacobian of this variable transformation is 1. This leads to $|\mathbf{x}^* \mathbf{y}| = |\mathbf{u}||v_1|$. Thus,

$$\begin{aligned}
&\mathrm{E}\big(f(|\mathbf{x}^* \mathbf{y}|)\, I(|\mathbf{x}^* \mathbf{y}| < \delta, |\mathbf{x}| \le K_e, |\mathbf{y}| \le K_e\big) \\
&\quad = \int \cdots \int f(|\mathbf{u}||v_1|)\, I(|\mathbf{u}||v_1| < \delta, |\mathbf{u}| \le K_e, |\mathbf{v}| \le K_e) \\
&\qquad \times p(\mathbf{u}, U\mathbf{v})\, d\mathbf{u}\, d\mathbf{v} \qquad \text{(A.2)}\\
&\quad \le K_d s_{2k} 2\pi (2K_e)^{2k-2} \int_0^{\sqrt{k}K_e} \rho_1^{2k-1}\, d\rho_1 \int_0^{\delta/\rho_1} \rho_2 f(\rho_1 \rho_2)\, d\rho_2 \\
&\quad \le K_d s_{2k} 2\pi (2K_e)^{2k-2} (2k-2)^{-1} \big(\sqrt{k}K_e\big)^{2k-2} \int_0^{\delta} t f(t)\, dt, \qquad \text{(A.3)}
\end{aligned}$$

where s_{2k} denotes the Euclidean area of the $2k$-dimensional unit sphere. Here, the inequality (A.2) follows from a polar transformation for the real and imaginary parts of $\mathbf{u}$ (dimension $= 2k$) and from a polar transformation for the real and imaginary parts (dimension $= 2$) of v_1. The lemma now follows from (A.3). □

LEMMA A.4. *Let $\{a_{nlkj_1, j_2}\}$, $l \le n^d$, $k, j_1, j_2 \le n$, be complex random variables satisfying $\max_{n,l,k,j_2} \sum_{j_1} |a_{nlkj_1, j_2}|^2 < K^2$, $\max_{n,l,k,j_1} \sum_{j_2} |a_{nlkj_1, j_2}|^2 < K^2$, and $|z_l| \le n^{1/36}$ are complex constants. Suppose that $\{X_{kj}, k, j = 1, 2, \dots\}$ is a double array of iid complex random variables with mean zero and finite sixth moment. Assume for each fixed k, $\{X_{kj}, j = 1, 2, \dots\}$ is independent of $\{a_{nlkj_1, j_2}\}$, $l \le n^d$, $j_1, j_2 \le n$. Then,*

$$\begin{aligned}
&\sup_{l \le n^d, k \le n} \Bigg| \sum_{j_1, j_2} a_{nlkj_1, j_2} \Bigg[\bigg(\frac{1}{\sqrt{n}} X_{kj_1} - z_l \delta_{k, j_1} \bigg) \bigg(\frac{1}{\sqrt{n}} \bar{X}_{kj_2} - \bar{z}_l \delta_{k, j_2} \bigg) \\
&\qquad\qquad - \delta_{j_1, j_2} \bigg(\frac{1}{n} \mathrm{E}|X_{11}|^2 + \delta_{k, j_1} |z_l|^2 \bigg) \Bigg] \Bigg| \qquad \text{(A.4)} \\
&\quad = o(n^{-5/36} K \ln^2 n),
\end{aligned}$$

where $d > 0$ is a positive constant and δ_{kj} is the Kronecker delta, that is, $= 1$ or 0 corresponding to $k = j$ or not.

PROOF. Without loss of generality, we may assume that $K = 1$, $\mathrm{E}(|X_{11}|^2) = 1$, and that a_{nlkj_1, j_2} are real nonrandom constants and $\{z_l\}$ and $X_{k,j}$ are real constants and random variables, respectively.

Now, let m be a positive integer. For k, j, defined $X_{mkj} = X_{kj}$ or zero according to $|X_{kj}| \le 2^{m/3}$ or not, respectively. Note that

$$\mathrm{P}\Bigg(\bigcup_{m=1}^{\infty} \bigcup_{k, j \le 2^m} (X_{mkj} \ne X_{kj}) \Bigg) \le \sum_{m=1}^{\infty} 2^{2m} \mathrm{P}\big(|X_{11}| > 2^{m/3}\big) < \infty$$

by the finiteness of the sixth moment of X_{11}. Therefore, by the Borel–Cantelli lemma, the variables X_{kj} in (A.4) can be replaced by X_{mkj}, for all $n \in (2^m, 2^{m+1}]$. In other words, we may assume that for each n, $|X_{mkj}| \le n^{1/3}$.

In the rest of the proof of this lemma, all probabilities and expectations are conditional probabilities and expectations for the a-variables given, namely, we treat the a-variables as nonrandom. For fixed $\varepsilon > 0$, by Bernstein's inequality, we have

$$\sum_n \sum_{l,k} \mathrm{P}\left(\frac{1}{n} \left| \sum_{j=1}^{n} a_{nlkjj}\left[X_{mkj}^2 - \mathrm{E}\, X_{m11}^2 \right] \right| \ge \varepsilon n^{-1/3} \ln^2 n \right)$$
$$\le M \sum_n n^{d+1} \exp\left(-\varepsilon^2 \ln^4 n \big/ \left[\mathrm{E} |X_{m11}|^4 + \varepsilon \ln^2 n \right] \right) < \infty.$$

which, together with Borel–Cantelli, implies that

$$\max_{l \le n^d,\, k \le n} \left| \frac{1}{n} \sum_{j=1}^{n} a_{nlkjj}\left[X_{mjj}^2 - \mathrm{E}\, X_{m11}^2 \right] \right| = o(n^{-1/3} \ln^2 n). \tag{A.5}$$

Because of the truncation, we have

$$\max_{l \le n^d,\, k \le n} \left| \frac{1}{\sqrt{n}} a_{nlkkk} z_l X_{m,kk} \right| < M n^{-5/36} = o(n^{-5/36} \ln^2 n). \tag{A.6}$$

By (A.5) and (A.6) and the fact that

$$\max_{l \le n^d,\, k \le n} \left| \frac{1}{n} \sum_{j=1}^{n} a_{nlkjj}\left[\mathrm{E}\left(X_{11}^2 \right) - \mathrm{E}\left(X_{m11}^2 \right) \right] \right| \le M \max_{l \le n^d,\, k \le n} n^{-7/3} \sum_{j=1}^{n} |a_{nlkjj}|$$
$$\le M n^{-4/3} = o(n^{-1/2} \ln^2 n),$$

to finish the proof of the lemma, one need only show that

$$\max_{l \le n^d,\, k \le n} \left| \frac{1}{n} \sum_{j_1 \ne j_2} a_{nlkj_1,j_2} X_{mkj_1} X_{mkj_2} \right| = o(n^{-5/36} \ln^2 n), \tag{A.7}$$

and

$$\max_{l \le n^d,\, k \le n} |z_l| \left| \frac{1}{\sqrt{n}} \sum_{j=1(j \ne k)}^{n} a_{nlkk,j} X_{mkj} \right| = o(n^{-5/36} \ln^2 n). \tag{A.8}$$

Note that (A.8) is implied by

$$\max_{l \le n^d,\, k \le n} \left| \frac{1}{\sqrt{n}} \sum_{j=1(j \ne k)}^{n} a_{nlkk,j} X_{mkj} \right| = o(n^{-5/36} \ln^2 n), \tag{A.9}$$

which can be proved by the same lines in the proof of (A.5).

In the proof of (A.7), for convenience of notation, we shall omit the subscripts n, k and l from a_{nlk,j_1,j_2} and rewrite j_1, j_2 as k and j. Also, we simplify X_{mkj} as X_j and assume that $a_{j_1 j_2} = 0$ if $j_1 \ge j_2$. We will finish the proof of (A.7) by establishing that the probability when the left-hand side of (A.7) is greater than $\varepsilon n^{-5/36}\ln^2 n$ can be smaller than any a fixed negative power of n.

Define $b_{nkj} = n^{-31/36}a_{kj}$, and for $2 \le h < n$ and $1 \le k < j \le h$, $b_{hkj} = b_{h+1,kj} + 2b_{h+1,j,h+1}b_{h+1,k,h+1}$. By induction and the condition that $\sum_{1\le k<j\le n}|b^2_{nkj}| \le n^{-13/18}$, one can prove that for any $2 \le h < n$ and $n > 60$,

$$
\begin{aligned}
\sum_{1\le k<j\le h} & b^2_{hkj} \\
&= \sum_{1\le k<j\le h}\Big[b^2_{h+1,k,j} + 4b_{h+1,k,j}b_{h+1,k,h+1}b_{h+1,j,h+1} \\
&\qquad\qquad + 4b^2_{h+1,k,h+1}b^2_{h+1,j,h+1}\Big] \\
&\le \sum_{1\le k<j\le h} b^2_{h+1,k,j} + 4\Bigg(\sum_{1\le k<j\le h} b^2_{h+1,k,j}\Bigg)^{1/2} \\
&\quad \times \sum_{k=1}^{h} b^2_{h+1,k,h+1} + 4\Bigg(\sum_{k=1}^{h} b^2_{h+1,k,h+1}\Bigg)^2 \\
&\le \sum_{1\le k<j\le h+1} b^2_{h+1,k,j} + \big(4n^{-13/36} + 4n^{-13/18}\big)\sum_{k=1}^{h} b^2_{h+1,k,h+1} \\
&\le \sum_{1\le k<j\le h+1} b^2_{h+1,k,j} \le n^{-13/18}.
\end{aligned}
\tag{A.10}
$$

Let $q = \sum_{h=2}^{n}\sum_{k=1}^{h-1} b^2_{hkh}$. Then, by definition, we have

$$
\begin{aligned}
q &= \sum_{k=1}^{n-1} n^{-31/18}a^2_{n-1,n} \\
&\quad + \sum_{h=2}^{n-1}\sum_{k=1}^{h-1}\big(b^2_{h+1,kh} + 4b_{h+1,kh}b_{h+1,k,h+1}b_{h+1,h,h+1} \\
&\qquad\qquad + 4b^2_{h+1,k,h+1}b^2_{h+1,h,h+1}\big) \\
&= \sum_{k<h} n^{-31/18}a^2_{k,h} \\
&\quad + 4\sum_{h=2}^{n-1}\sum_{k=1}^{h-1}\sum_{s=h+1}^{n}\big(b_{s,kh}b_{s,k,s}b_{s,h,s} + 4b^2_{s,k,s}b^2_{s,h,s}\big) \\
&\le n^{-13/18} + 4\sum_{s=3}^{n}\Bigg(\sum_{k<h<s} b^2_{s,kh}\Bigg)^{1/2}\Bigg(\sum_{k<s} b^2_{sks}\Bigg)^{1/2}\Bigg(\sum_{h<s} b^2_{shs}\Bigg)^{1/2} \\
&\quad + 4\sum_{s=3}^{n}\Bigg(\sum_{k<s} b^2_{sks}\Bigg)^2 \\
&\le n^{-13/18} + 4\big(n^{-13/36} + n^{-13/18}\big)q \quad [\text{by (A.10)}] \\
&\le 2n^{-13/18}.
\end{aligned}
\tag{A.11}
$$

Write $\tilde{b}_j = \sum_{k=j+1}^{n} b_{jkj}^2$. The estimate (A.11) implies that $\tilde{b}_j < 2n^{-13/18}$. Notice that

$$\sum_{j=1} \tilde{b}_j \mathrm{E}(X_j^2) = O(n^{-13/18}).$$

Then, by (A.11) and applying Bernstein's inequality, we obtain

$$\text{(A.12)} \qquad \begin{aligned} \mathrm{P}\left(\sum_{j=1} \tilde{b}_j X_j^2 \geq 2\right) &\leq \mathrm{P}\left(\sum_{j=1} \tilde{b}_j \left(X_j^2 - \mathrm{E}\left(X_j^2\right)\right) \geq 1\right) \\ &\leq \exp(-cn^{1/18}) \end{aligned}$$

for some positive constant c.

Now, we shall find the bound for the b_{khj}'s. By the definition of b_{hkj} and the estimation (A.11), we have

$$\text{(A.13)} \qquad \begin{aligned} |b_{khj}| &\leq |b_{h+1,\,kj}| + |b_{h+1,\,k,\,h+1}||b_{h+1,\,j,\,h+1}| \\ &\leq |b_{n,\,kj}| + \sum_{s=h+1}^{n} |b_{sks}||b_{sjs}| \\ &\leq n^{-31/36} + q \leq 3n^{-13/18}. \end{aligned}$$

Define

$$E_h = \left\{\left|\sum_{k=1}^{h-1} b_{hkh} X_k\right| < \tfrac{1}{2} n^{-13/36}\right\},$$

$$E_0 = \left\{\sum_{j=1} \tilde{b}_j X_j^2 < 2\right\} = \left\{\sum_{h=2}^{n} \sum_{j=1}^{h-1} b_{hjh}^2 X_j^2 < 2\right\}$$

and

$$F_h = \bigcap_{l=1}^{h} E_l.$$

Then, by (A.13) and applying Bernstein's inequality or Kolmogorov's inequality [Loève (1977), page 266], we have

$$\mathrm{P}(E_h^c) \leq C \exp(-cn^{1/36}),$$

for some positive constants C and c. Consequently,

$$\begin{aligned} &\mathrm{P}\left(n^{-1}\left|\sum_{k<j} a_{kj} X_k X_j\right| \geq \varepsilon n^{-5/36} \ln^2 n\right) \\ &\quad \leq \mathrm{P}(E_0) + \sum_{h=3}^{n} \mathrm{P}(E_m^c) \\ &\qquad + \exp(-\varepsilon \ln^2 n) \mathrm{E}\, I(F_n) I(E_0) \exp\left\{\sum_{j=2}^{n} \sum_{k=1}^{j-1} b_{nkj} X_k X_j\right\}. \end{aligned}$$

Thus, to complete the proof of the lemma, it suffices to show that

$$\text{(A.14)} \qquad \mathrm{E}\, I(F_n)\, I(E_0) \exp\left\{ \sum_{j=2}^{n} \sum_{k=1}^{j-1} b_{nkj} X_k X_j \right\} = O(1),$$

which is obviously [see (A.12)] implied by

$$\text{(A.15)} \quad \mathrm{E}\, I(F_n) \exp\left\{ -\sum_{h=2}^{n} \sum_{j=1}^{h-1} b_{hjh}^2 X_j^2 + \sum_{j=2}^{n} \sum_{k=1}^{j-1} b_{nkj} X_k X_j \right\} = O(1).$$

In fact, by induction, we have

$$\begin{aligned}
&\mathrm{E}\Big(I(F_n) \exp\left\{ -\sum_{h=2}^{n} \sum_{j=1}^{h-1} b_{hjh}^2 X_j^2 + \sum_{j=2}^{n} \sum_{k=1}^{j-1} b_{nkj} X_k X_j \right\} \\
&\quad \le \mathrm{E}\Big(I(F_n) \exp\left\{ -\sum_{h=2}^{n} \sum_{j=1}^{h-1} b_{hjh}^2 X_j^2 + \sum_{j=2}^{n-1} \sum_{k=1}^{j-1} b_{nkj} X_k X_j \right\} \\
&\qquad \times \left[1 + \sum_{k=1}^{n-1} b_{nkj} X_k X_n + \left(\sum_{k=1}^{n-1} b_{nkj} X_k X_n \right)^2 \right] \quad \left(\text{by} \left| \sum_{k=1}^{n-1} b_{nkj} X_k X_n \right| \le \tfrac{1}{2} \right) \\
&\quad \le \mathrm{E}\Big(I(F_n) \exp\left\{ \sum_{j=2}^{n-1} \sum_{k=1}^{j-1} b_{nkj} X_k X_j \right\} \left[1 + n^{-5/3} + \left(\sum_{k=1}^{n-1} b_{nkj} X_k \right)^2 \right] \\
&\qquad\qquad\qquad\qquad \left[X_n \text{ is independent of } F_n \text{ and } |\mathrm{E}(X_n)| \le Cn^{-5/3}, \right] \\
&\quad \le \mathrm{E}\Big(I(F_{n-1}) \exp\left\{ -\sum_{h=2}^{n-1} \sum_{j=1}^{h-1} b_{hjh}^2 X_j^2 + \sum_{j=2}^{n-1} \sum_{k=1}^{j-1} b_{n-1,kj} X_k X_j + n^{-5/3} \right\} \cdots \\
&\quad \le \mathrm{E} \exp\left\{ b_{212} X_1 X_2 + n^{-2/3} \right\} \le \exp\left\{ n^{-1/18} + n^{-2/3} \right\} \to 1.
\end{aligned}$$

This establishes (A.15) and consequently the proof of the lemma is complete. □

A2. A result known in the literature.

LEMMA A.5. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $m \times n$ *complex matrices and denote by* ν_a *and* ν_b *the empirical spectral distributions of* $\mathbf{A}^*\mathbf{A}$ *and* $\mathbf{B}^*\mathbf{B}$, *respectively. Then, we have*

$$\text{(A.16)} \qquad \begin{aligned} &\left(\int |\nu_a(u) - \nu_b(u)|\, du \right)^2 \\ &\qquad \le 2n^{-2} \operatorname{tr}(\mathbf{A}^*\mathbf{A} + \mathbf{B}^*\mathbf{B}) \operatorname{tr}((\mathbf{A} - \mathbf{B})^*(\mathbf{A} - \mathbf{B})). \end{aligned}$$

Similar versions of this lemma were used by Bai and Silverstein (1995), Wachter (1978) and Yin (1986). The exact version of this lemma (but for real

A and **B**) was used in Bai (1993b) but it was not stated as a lemma. An outline of the proof of the lemma is given below:

$$\begin{aligned}\left(\int|\nu_a(u)-\nu_b(u)|\,du\right)^2 &\le \left(\frac{1}{n}\sum_{k=1}^{n}|\lambda_k-\eta_k|\right)^2\\ &\le \left(\frac{1}{n}\sum_{k=1}^{n}\left|\sqrt{\lambda_k}+\sqrt{\eta_k}\right|^2\right)\left(\frac{1}{n}\sum_{k=1}^{n}\left|\sqrt{\lambda_k}-\sqrt{\eta_k}\right|^2\right)\\ &\le 2n^{-2}\operatorname{tr}(\mathbf{AA}^*+\mathbf{BB}^*)\operatorname{tr}((\mathbf{A}-\mathbf{B})^*(\mathbf{A}-\mathbf{B})).\end{aligned}$$

The last step follows from the von Neumann inequality.

Acknowledgments. The author thanks Professor J. W. Silverstein and the referee for their careful review of the paper and their helpful comments and suggestions in improving the paper.

REFERENCES

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

BAI, Z. D. (1993b). Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. *Ann. Probab.* **21** 649–672.

BAI, Z. D. and YIN, Y. Q. (1986). Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang. *Probab. Theory Related Fields* **73** 555–569.

BAI, Z. D. and YIN, Y. Q. (1988a). A convergence to the semicircle law. *Ann. Probab.* **16** 863–875.

BAI, Z. D. and YIN, Y. Q. (1988b). Necessary and sufficient conditions for the almost sure convergence of the largest eigenvalue of Wigner matrices. *Ann. Probab.* **16** 1729–1741.

BAI, Z. D. and YIN, Y. Q. (1993). Limit of the smallest eigenvalue of large dimensional covariance matrix. *Ann. Probab.* **21** 1275–1294.

EDELMAN, A. (1995). The circular law and the probability that a random matrix has *k* real eigenvalues. Unpublished manuscript.

GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261.

GEMAN, S. (1986). The spectral radius of large random matrices. *Ann. Probab.* **14** 1318–1328.

GINIBRE, J. (1965). Statistical ensembles of complex, quaterion and real matrices. *J. Math. Phys.* 440–449.

GIRKO, V. L. (1984a). Circle law. *Theory Probab. Appl.* 694–706.

GIRKO, V. L. (1984b). On the circle law. *Theory Probab. Math. Statist.* 15–23.

HWANG, C. R. (1986). A brief survey on the spectral radius and the spectral distribution of large dimensional random matrices with iid entries. In *Random Matrices and Their Applications* (M. L. Mehta, ed.) 145–152. Amer. Math. Soc., Providence, RI.

JONSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38.

LOÈVE, M. (1977). *Probability Theory*, 4th ed. Springer, New York.

MARČENKO, V. A. and PASTUR, L. A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sb.* **1** 457–483.

PASTUR, L. A. (1972). On the spectrum of random matrices. *Teoret. Mat. Fiz.* **10** 102–112. (English translation in *Theoret. and Math. Phys.* **10** 67–74.)

PASTUR, L. A. (1973). Spectra of random self-adjoint operators. *Uspekhi Mat. Nauk* **28** 4–63. (English translation in *Russian Math. Surveys* **28** 1–67.)

SILVERSTEIN, J. W. and BAI, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *J. Multivariate Anal.* **54** 175–192.

SILVERSTEIN, W. J. and CHOI, S. I. (1995). Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Multivariate Anal.* **54** 295–309.

WACHTER, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18.

WACHTER, K. W. (1980). The limiting empirical measure of multiple discriminant ratios. *Ann. Statist.* **8** 937–957.

YIN, Y. Q., BAI, Z. D. and KRISHNAIAH, P. R. (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** 509–521.

DEPARTMENT OF APPLIED MATHEMATICS
NATIONAL SUN YAT-SEN UNIVERSITY
KAOHSIUNG, TAIWAN 80424
E-MAIL: baiz@math.nsysu.edu.tw