

CONVERGENCE RATE OF EXPECTED SPECTRAL DISTRIBUTIONS OF LARGE RANDOM MATRICES. PART I. WIGNER MATRICES

BY Z. D. BAI

Temple University

In this paper, we shall develop certain inequalities to bound the difference between distributions in terms of their Stieltjes transforms. Using these inequalities, convergence rates of expected spectral distributions of large dimensional Wigner and sample covariance matrices are established. The paper is organized into two parts. This is the first part, which is devoted to establishing the basic inequalities and a convergence rate for Wigner matrices.

1. Introduction. Let W_n be an $n \times n$ symmetric matrix. Denote its eigenvalues by $\lambda_1 \leq \dots \leq \lambda_n$. Then its spectral distribution is defined by

$$F_n(x) = \frac{1}{n} \#\{i: \lambda_i \leq x\},$$

where $\#\{Q\}$ denotes the number of entries in the set Q . The interest in the spectral analysis of high dimensional random matrices is to investigate limiting theorems for spectral distributions of high-dimensional random matrices with nonrandom limiting spectral distributions.

Research on the limiting spectral analysis of high-dimensional random matrices dates back to Wigner's (1955, 1958) semicircular law for a Gaussian (or Wigner) matrix; he proved that the expected spectral distribution of a high-dimensional Wigner matrix tends to the so-called semicircular law. This work was generalized by Arnold (1967) and Grenander (1963) in various aspects. Bai and Yin (1988a) proved that the spectral distribution of a sample covariance matrix (suitably normalized) tends to the semicircular law when the dimension is relatively smaller than the sample size. Following the work by Pastur (1972, 1973), the asymptotic theory of spectral analysis of high-dimensional sample covariance matrices was developed by many researchers including Bai, Yin and Krishnaiah (1986), Grenander and Silverstein (1977), Jonsson (1982), Wachter (1978), Yin (1986) and Yin and Krishnaiah (1983). Also, Bai, Yin and Krishnaiah (1986, 1987), Silverstein (1985a), Wachter (1980), Yin (1986) and Yin and Krishnaiah (1983) investigated the limiting spectral distribution of the multivariate F matrix, or more generally, of products of random

* Received December 1990; revised January 1992.

AMS 1991 subject classifications. 60F15, 62F15.

Key words and phrases. Berry–Esseen inequality, convergence rate, large dimensional random matrix, Marchenko–Pastur distribution, sample covariance matrix, semicircular law, spectral analysis, Stieltjes transform, Wigner matrix.

matrices. In recent years, Voiculescu (1990, 1991) investigated the convergence to the semicircular law associated with free random variables.

In applications of the asymptotic theorems of spectral analysis of high-dimensional random matrices, two important problems arose after the limiting spectral distribution was found. The first is the bound on extreme eigenvalues; the second is the convergence rate of the spectral distribution, with respect to sample size. For the first problem, the literature is extensive. The first success was due to Geman (1980), who proved that the largest eigenvalue of a sample covariance matrix converges almost surely to a limit under a condition of existence of all moments of the underlying distribution. Yin, Bai and Krishnaiah (1988) proved the same result under existence of the fourth moment, and Bai, Silverstein and Yin (1988) proved that the existence of the fourth moment is also necessary for the existence of the limit. Bai and Yin (1988b) found necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. Bai and Yin (1990), Silverstein (1985b) and Yin, Bai and Krishnaiah (1983) considered the almost sure limit of the smallest eigenvalue of a covariance matrix. Some related works can be found in Geman (1986) and Bai and Yin (1986).

The second problem, the convergence rate of the spectral distributions of high-dimensional random matrices, is of practical interest, but has been open for decades. The principal approach to establishing limiting theorems for spectral analysis of high-dimensional random matrices is to show that each moment (with fixed order) of the spectral distribution tends to a nonrandom limit; this proves the existence of the limiting spectral distribution by the Carleman criterion. This method successfully established the limiting theorems for spectral distributions of high-dimensional Wigner matrices, sample covariance matrices and multivariate F matrices. However, this method cannot give a convergence rate.

This paper develops a new methodology to establish convergence rates of spectral distributions of high-dimensional random matrices. The paper is written in two parts: In Part I, we shall mainly consider the convergence rate of empirical spectral distributions of Wigner matrices. The convergence rate for sample covariance matrices will be discussed in Part II. The organization of Part I is as follows: In Section 2, basic concepts of Stieltjes transforms are introduced. Three inequalities to bound the difference between distribution functions in terms of their Stieltjes transforms are established. This paper involves a lot of computation of matrix algebra and complex-valued functions. For completeness, some necessary results in these areas are included in Section 3. Some lemmas are also presented in this section. Theorem 2.1 is used in Section 4 to establish a convergence rate for the expected spectral distribution of high-dimensional Wigner matrices.

The rate for Wigner matrices established in this part of the paper is $O(n^{-1/4})$. From the proof of the main theorem, one may find that the rate may be further improved to $O(n^{-1/3+\eta})$ by expanding more terms and assuming the existence of higher moments of the underlying distributions. However, it is

not known whether we can get improvements beyond the order of $O(n^{-1/3})$, say $O(n^{-1/2})$ or $O(n^{-1})$, as conjectured in Section 4.

2. Inequalities of distance between distributions in terms of their Stieltjes transforms. Suppose that F is a function of bounded variation. Then its Stieltjes transform is defined by

$$(2.1) \quad s(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dF(x),$$

where $z = u + iv$ is a complex variable. It is well known [see Girko (1989)] that the following inversion formula holds: For any continuity points $x_1 \leq x_2$ of F ,

$$(2.2) \quad F(x_2) - F(x_1) = \lim_{v \downarrow 0} \frac{1}{\pi} \int_{x_1}^{x_2} \text{Im}(s(u + iv)) du,$$

where $\text{Im}(\cdot)$ denotes the imaginary part of a complex number. From this, it is easy to show that if $\text{Im}(s(z))$ is continuous at $z = x + i0$, then F is differentiable at x and its derivative is given by

$$(2.3) \quad F'(x) = \frac{1}{\pi} \text{Im}(s(x + i0)).$$

This formula gives an easy way to extract the density function from its corresponding Stieltjes transform.

Also, one can easily verify the continuity theorem for Stieltjes transforms; that is, $F_n \rightarrow_w F$ if and only if $s_n(z) \rightarrow s(z)$ for all $z = u + iv$ with $v > 0$, where s_n and s are the Stieltjes transforms of the distributions F_n and F , respectively. Due to this fact, it is natural to ask whether we can establish a Berry–Esseen type inequality to evaluate the closeness between distributions in terms of their Stieltjes transforms. The first attempt was made by Girko (1989) who established an inequality by integrating both sides of Berry–Esseen’s basic inequality. Unfortunately, the justification of the exchange of integration signs in his proof is not obvious. More importantly, Girko’s inequality seems too complicated to apply. We establish the following basic inequality.

THEOREM 2.1. *Let F be a distribution function and let G be a function of bounded variation satisfying $\int |F(x) - G(x)| dx < \infty$. Denote their Stieltjes transforms by $f(z)$ and $g(z)$, respectively. Then we have*

$$(2.4) \quad \begin{aligned} \|F - G\| &:= \sup_x |F(x) - G(x)| \\ &\leq \frac{1}{\pi(2\gamma - 1)} \left[\int_{-\infty}^{\infty} |f(z) - g(z)| du \right. \\ &\quad \left. + \frac{1}{v} \sup_x \int_{|y| \leq 2va} |G(x + y) - G(x)| dy \right], \end{aligned}$$

where $z = u + iv$, $v > 0$, and a and γ are constants related to each other by

$$(2.5) \quad \gamma = \frac{1}{\pi} \int_{|u| < a} \frac{1}{u^2 + 1} du > \frac{1}{2}.$$

PROOF. Write $\Delta = \sup_x |F(x) - G(x)|$. Without loss of generality, we can assume that $\Delta > 0$. Then, there is a sequence $\{x_n\}$ such that $F(x_n) - G(x_n) \rightarrow \Delta$ or $-\Delta$.

We shall first consider the case that $F(x_n) - G(x_n) \rightarrow \Delta$. For each x , we have

$$(2.6) \quad \begin{aligned} \frac{1}{\pi} \int_{-\infty}^{\infty} |f(z) - g(z)| du &\geq \frac{1}{\pi} \int_{-\infty}^x \operatorname{Im}(f(z) - g(z)) du \\ &= \frac{1}{\pi} \int_{-\infty}^x \left[\int_{-\infty}^{\infty} \frac{v d(F(y) - G(y))}{(y - u)^2 + v^2} \right] du \\ &= \frac{1}{\pi} \int_{-\infty}^x \left[\int_{-\infty}^{\infty} \frac{2v(y - u)(F(y) - G(y)) dy}{((y - u)^2 + v^2)^2} \right] du \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} (F(y) - G(y)) \left[\int_{-\infty}^x \frac{2v(y - u) du}{((y - u)^2 + v^2)^2} \right] dy \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(F(x - vy) - G(x - vy)) dy}{y^2 + 1}. \end{aligned}$$

Here, the second equality follows from integration by parts while the third follows from Fubini's theorem due to the integrability of $|F(y) - G(y)|$. Since F is nondecreasing, we have

$$(2.7) \quad \begin{aligned} &\frac{1}{\pi} \int_{|y| < a} \frac{(F(x - vy) - G(x - vy)) dy}{y^2 + 1} \\ &\geq \gamma(F(x - va) - G(x - va)) \\ &\quad - \frac{1}{\pi} \int_{|y| < a} |G(x - vy) - G(x - va)| dy \\ &\geq \gamma(F(x - va) - G(x - va)) \\ &\quad - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x + y) - G(x)| dy. \end{aligned}$$

Take $x = x_n + va$. Then, (2.6) and (2.7) imply that

$$\begin{aligned}
 & \frac{1}{\pi} \int_{-\infty}^{\infty} |f(z) - g(z)| du \\
 & \geq \gamma(F(x_n) - G(x_n)) \\
 (2.8) \quad & - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x+y) - G(x)| dy - (1-\gamma)\Delta \\
 & \rightarrow (2\gamma - 1)\Delta - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x+y) - G(x)| dy,
 \end{aligned}$$

which implies (2.4).

Now we consider the case that $F(x_n) - G(x_n) \rightarrow -\Delta$. Similarly, we have, for each x , that

$$\begin{aligned}
 & \frac{1}{\pi} \int_{-\infty}^{\infty} |f(z) - g(z)| du \\
 (2.9) \quad & \geq \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(G(x-vy) - F(x-vy)) dy}{y^2 + 1} \\
 & \geq \gamma(G(x+va) - F(x+va)) \\
 & - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x+y) - G(x)| dy - (1-\gamma)\Delta.
 \end{aligned}$$

By taking $x = x_n - va$, we have

$$\begin{aligned}
 & \frac{1}{\pi} \int_{-\infty}^{\infty} |f(z) - g(z)| du \\
 (2.10) \quad & \geq \gamma(G(x_n) - F(x_n)) \\
 & - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x+y) - G(x)| dy - (1-\gamma)\Delta \\
 & \rightarrow (2\gamma - 1)\Delta - \frac{1}{\pi v} \sup_x \int_{|y| < 2va} |G(x+y) - G(x)| dy,
 \end{aligned}$$

which implies (2.4) for the latter case. This completes the proof of Theorem 2.1. \square

REMARK 2.1. In the proof of Theorem 2.1, one may find that the following version is stronger than Theorem 2.1:

$$\begin{aligned}
 (2.11) \quad \|F - G\| & \leq \frac{1}{\pi(2\gamma - 1)} \left[\int_{-\infty}^{\infty} |\operatorname{Im}(f(z) - g(z))| du \right. \\
 & \quad \left. + \frac{1}{v} \sup_x \int_{|y| \leq 2va} |G(x+y) - G(x)| dy \right].
 \end{aligned}$$

However, in application of the inequalities, we did not find any significant superiority of (2.11) over (2.4).

Sometimes the functions F and G may have light tails or both may even have bounded support. In such cases, we may establish a bound for $\|F - G\|$ by means of the integral of the absolute difference of their Stieltjes transforms on only a finite interval. We have the following theorem.

THEOREM 2.2. *Under the assumptions of Theorem 2.1, we have*

$$\begin{aligned}
 (2.12) \quad \|F - G\| \leq & \frac{1}{\pi(1 - \kappa)(2\gamma - 1)} \left[\int_{-A}^A |f(z) - g(z)| du \right. \\
 & + 2\pi v^{-1} \int_{|x|>B} |F(x) - G(x)| dx \\
 & \left. + v^{-1} \sup_x \int_{|y|\leq 2va} |G(x + y) - G(x)| dy \right],
 \end{aligned}$$

where A and B are positive constants such that $A > B$ and

$$(2.13) \quad \kappa = \frac{4B}{\pi(A - B)(2\gamma - 1)} < 1.$$

The following corollary is immediate.

COROLLARY 2.3. *In addition to the assumptions of Theorem 2.1, assume further that, for some constant $B > 0$, $F([-B, B]) = 1$ and $|G|((-\infty, -B)) = |G|((B, \infty)) = 0$, where $|G|((a, b))$ denotes the total variation of the signed measure G on the interval (a, b) . Then, we have*

$$\begin{aligned}
 (2.14) \quad \|F - G\| \leq & \frac{1}{\pi(1 - \kappa)(2\gamma - 1)} \left[\int_{-A}^A |f(z) - g(z)| du \right. \\
 & \left. + v^{-1} \sup_x \int_{|y|\leq 2va} |G(x + y) - G(x)| dy \right],
 \end{aligned}$$

where A , B and κ are defined in (2.13).

REMARK 2.2. The benefit of using Theorem 2.2 and Corollary 2.3 is that we need only estimate the difference of Stieltjes transforms of the two distributions of interest on a fixed interval. When Theorem 2.2 is applied to establish the convergence rate of the spectral distribution of a sample covariance matrix in Section 4, it is crucial to the proof of Theorem 4.1 that A is independent of the sample size n . It should also be noted that the integral limit A in Girko's (1989) inequality should tend to infinity with a rate of A^{-1} faster than the convergence rate to be established. Therefore, our Theorem 2.2 and Corollary 2.3 are much easier to use than Girko's inequality.

PROOF OF THEOREM 2.2. Using the notation given in the proof of Theorem 2.1, we have

$$\begin{aligned}
 & \int_A^\infty |f(z) - g(z)| \, du \\
 &= \int_A^\infty \left| \int_{-\infty}^\infty \frac{(F(x) - G(x)) \, dx}{(x - z)^2} \right| \, du \\
 (2.15) \quad &\leq \int_A^\infty \left| \int_{-B}^B \frac{(F(x) - G(x)) \, dx}{(x - z)^2} \right| \, du \\
 &\quad + \int_{-\infty}^\infty \left| \int_{|x|>B} \frac{(F(x) - G(x)) \, dx}{(x - z)^2} \right| \, du \\
 &\leq 2B\Delta \int_A^\infty \frac{du}{(u - B)^2 + v^2} + \pi v^{-1} \int_{|x|>B} |F(x) - G(x)| \, dx \\
 &\leq \frac{2B\Delta}{A - B} + \frac{\pi}{v} \int_{|x|>B} |F(x) - G(x)| \, dx.
 \end{aligned}$$

By symmetry, we get the same bound for $\int_{-\infty}^{-A} |f(z) - g(z)| \, du$. Substituting the above inequality into (2.4), we obtain (2.12) and the proof is complete. \square

3. Preliminaries.

3.1. *The notation \sqrt{z} .* We need first to clarify the notation \sqrt{z} , $z = u + iv$, ($v \neq 0, i = \sqrt{-1}$). Throughout this paper, \sqrt{z} denotes the square root of z with a positive imaginary part. In fact, we have the following expressions:

$$(3.1) \quad \sqrt{z} = \text{sign}(v) \frac{|z| + z}{\sqrt{2(|z| + u)}},$$

or

$$\text{Re}(\sqrt{z}) = \frac{1}{\sqrt{2}} \text{sign}(v) \sqrt{\sqrt{u^2 + v^2} + u} = \frac{v}{\sqrt{2(\sqrt{u^2 + v^2} - u)}}$$

and

$$\text{Im}(\sqrt{z}) = \frac{1}{\sqrt{2}} \sqrt{\sqrt{u^2 + v^2} - u} = \frac{|v|}{\sqrt{2(\sqrt{u^2 + v^2} + u)}},$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of a complex number indicated in the parentheses. If z is a real number, define $\sqrt{z} = \lim_{v \downarrow 0} \sqrt{z + iv}$. Then the definition agrees with the arithmetic square root of positive numbers. However, under this definition, the multiplication rule for square roots fails; that is, in general, $\sqrt{z_1 z_2} \neq \sqrt{z_1} \sqrt{z_2}$. The introduction of the definition (3.1) is merely for convenience and definiteness.

3.2. *Stieltjes transform of the semicircular law.* By definition,

$$\begin{aligned}
 s(z) &= \frac{1}{2\pi} \int_{-2}^2 \frac{\sqrt{4-x^2}}{x-z} dx \\
 &= \frac{1}{\pi} \int_0^\pi \frac{\sin^2 \theta}{\cos \theta - (1/2)z} d\theta \quad \text{by } x = 2 \cos \theta \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin^2 \theta}{\cos \theta - (1/2)z} d\theta \\
 &= -\frac{1}{4\pi i} \oint_{|\zeta|=1} \frac{(\zeta^2 - 1)^2}{\zeta^2(\zeta^2 - \zeta z + 1)} d\zeta \quad \text{by } \zeta = \exp\{i\theta\}.
 \end{aligned}$$

Now, we apply the residue theorem to evaluate the integral. First, we note that the function $(\zeta^2 - 1)^2/\zeta^2(\zeta^2 - \zeta z + 1)$ has three singular points, 0 and $\zeta_{1,2} = (1/2)(z \pm \sqrt{z^2 - 4})$, with residues z and $\pm \sqrt{z^2 - 4}$. Here $\zeta_{1,2}$ are in fact the roots of the quadratic equation $\zeta^2 - z\zeta + 1 = 0$. Thus, $\zeta_1\zeta_2 = 1$. Applying the formula (3.1) to the square root of $z^2 - 4 = (u^2 - v^2 - 4) + 2uvi$, one finds that the real parts of z and $\sqrt{z^2 - 4}$ have the same sign while their imaginary parts are positive. Hence, both the real and imaginary parts of ζ_1 have larger absolute values than those of ζ_2 . Therefore, ζ_1 is outside the unit circle while ζ_2 is inside. Hence, we obtain

$$(3.2) \quad s(z) = -\frac{1}{2}(z - \sqrt{z^2 - 4}).$$

Noting that $s(z) = -\zeta_2$, we have

$$(3.3) \quad |s(z)| < 1.$$

3.3. *Integrals of the square of the absolute value of Stieltjes transforms.*

LEMMA 3.1. *Suppose that $\phi(x)$ is a bounded probability density supported on a finite interval (A, B) . Then,*

$$\int_{-\infty}^{\infty} |s(z)|^2 du < 2\pi^2 M_\phi,$$

where $s(z)$ is the Stieltjes transform of ϕ and M_ϕ the upper bound of ϕ .

PROOF. We have

$$\begin{aligned}
 I &:= \int_{-\infty}^{\infty} |s(z)|^2 du \\
 &= \int_{-\infty}^{\infty} \int_A^B \int_A^B \frac{\phi(x)\phi(y)}{(x-z)(y-\bar{z})} dx dy du \\
 &= \int_A^B \int_A^B \phi(x)\phi(y) dx dy \int_{-\infty}^{\infty} \frac{1}{(x-z)(y-\bar{z})} du \quad (\text{Fubini's theorem}) \\
 &= \int_A^B \int_A^B \frac{2\pi i}{y-x+2vi} \phi(x)\phi(y) dx dy, \quad (\text{residue theorem}).
 \end{aligned}$$

Note that

$$\begin{aligned} & \int_A^B \int_A^B \operatorname{Re} \left(\frac{1}{y-x+2vi} \right) \phi(x) \phi(y) \, dx \, dy \\ &= \int_A^B \int_A^B \left(\frac{y-x}{(y-x)^2+4v^2} \right) \phi(x) \phi(y) \, dx \, dy = 0, \quad \text{by symmetry.} \end{aligned}$$

We finally obtain that

$$\begin{aligned} I &= -2\pi \int_A^B \int_A^B \operatorname{Im} \left(\frac{1}{y-x+2vi} \right) \phi(x) \phi(y) \, dx \, dy \\ &= 4\pi v \int_A^B \int_A^B \left(\frac{1}{(y-x)^2+4v^2} \right) \phi(x) \phi(y) \, dx \, dy \\ &\leq 4\pi v M_\phi \int_{-\infty}^{\infty} \int_A^B \phi(y) \left(\frac{1}{w^2+4v^2} \right) \, dw \, dy \quad \text{by setting } w = x-y \\ &= 2\pi^2 M_\phi. \end{aligned}$$

The proof is complete. \square

REMARK 3.1. The assumption that ϕ has finite support has been used in the verification of the conditions of Fubini's theorem.

Applying this lemma to the semicircular law, we get the following corollary.

COROLLARY 3.2. We have

$$(3.4) \quad \int |s(z)|^2 \, du \leq 2\pi.$$

3.4. *Some algebraic formulae used in this paper.* In this paper, certain algebraic formulae are used. Some of them are well known and will be listed only. For the others, brief proofs will be given. Most of the known results can be found in Xu (1982).

3.4.1. *Inverse matrix formula.* Let A be an $n \times n$ nonsingular matrix. Then

$$A^{-1} = \frac{1}{\det(A)} A^*,$$

where A^* is the adjoint matrix of A , that is, the transposed matrix of cofactors of order $n-1$ of A and $\det(A)$ denotes the determinant of the matrix A . By this formula, we have

$$(3.5) \quad \operatorname{tr}(A^{-1}) = \sum_{k=1}^n \det(A_k) / \det(A),$$

where A_k is the k th major submatrix of order $n - 1$ of the matrix A , that is, the matrix obtained from A by deleting the k th row and column.

3.4.2. If A is nonsingular, then

$$(3.6) \quad \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B),$$

which follows immediately from the fact that

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}.$$

3.4.3. If both A and A_k are nonsingular and if we write $A^{-1} = [\alpha^{kl}]$, then

$$(3.7) \quad \alpha^{kk} = \frac{1}{\alpha_{kk} - \alpha'_k A_k^{-1} \beta_k},$$

where α_{kk} is the k th diagonal entry of A , A_k the major submatrix of order $n - 1$ as defined in Section 3.4.1, α_k the vector obtained from the k th row of A by deleting the k th entry and β_k the vector from the k th column by deleting the k th entry. Then, (3.7) follows from (3.5) and (3.6).

If A is an $n \times n$ symmetric nonsingular matrix and all its major submatrices of order $(n - 1)$ are nonsingular, then from (3.5) and (3.7), it follows immediately that

$$(3.8) \quad \text{tr}(A^{-1}) = \sum_{k=1}^n \frac{1}{\alpha_{kk} - \alpha'_k A_k^{-1} \alpha_k}.$$

3.4.4. Use the notation of Section 3.4.3. If A and A_k are nonsingular symmetric matrices, then

$$(3.9) \quad \text{tr}(A^{-1}) - \text{tr}(A_k^{-1}) = \frac{1 + \alpha'_k A_k^{-2} \alpha_k}{\alpha_{kk} - \alpha'_k A_k^{-1} \alpha_k}.$$

This is a direct consequence of the following well-known formula for a nonsingular symmetric matrix:

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})^{-1} \Sigma_{21} \Sigma_{11}^{-1} & -\Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})^{-1} \\ -(\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})^{-1} \Sigma_{21} \Sigma_{11}^{-1} & (\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})^{-1} \end{bmatrix},$$

where $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ is a partition of the symmetric matrix Σ .

3.4.5. If real symmetric matrices A and B are commutative and such that $A^2 + B^2$ is nonsingular, then the complex matrix $A + iB$ is nonsingular and

$$(3.10) \quad (A + iB)^{-1} = (A - iB)(A^2 + B^2)^{-1}.$$

This can be directly verified.

3.4.6. Let $z = u + iv$, $v > 0$, and let A be an $n \times n$ real symmetric matrix. Then

$$(3.11) \quad |\operatorname{tr}(A - zI_n)^{-1} - \operatorname{tr}(A_k - zI_{n-1})^{-1}| \leq v^{-1}.$$

PROOF. By (3.9), we have

$$\operatorname{tr}(A - zI_n)^{-1} - \operatorname{tr}(A_k - zI_{n-1})^{-1} = \frac{1 + \alpha'_k(A_k - zI_{n-1})^{-2}\alpha_k}{\alpha_{kk} - z - \alpha'_k(A - zI_{n-1})^{-1}\alpha_k}.$$

If we denote $A_k = E' \operatorname{diag}[\lambda_1 \cdots \lambda_{n-1}]E$ and $\alpha'_k E' = (y_1, \dots, y_{n-1})$, where E is an $(n - 1) \times (n - 1)$ (real) orthogonal matrix, then we have

$$\begin{aligned} |1 + \alpha'_k(A_k - zI_{n-1})^{-2}\alpha_k| &= \left| 1 + \sum_{l=1}^{n-1} y_l^2 (\lambda_l - z)^{-2} \right| \\ &\leq 1 + \sum_{l=1}^{n-1} y_l^2 ((\lambda_l - u)^2 + v^2)^{-1} \\ &= 1 + \alpha'_k((A_k - uI_{n-1})^2 + v^2I_{n-1})^{-1}\alpha_k. \end{aligned}$$

On the other hand, by (3.10) we have

$$(3.12) \quad \begin{aligned} \operatorname{Im}(\alpha_{kk} - z - \alpha'_k(A - zI_{n-1})^{-1}\alpha_k) \\ = v \left(1 + \alpha'_k((A_k - uI_{n-1})^2 + v^2I_{n-1})^{-1}\alpha_k \right). \end{aligned}$$

From these estimates, (3.11) follows. \square

3.5. *A lemma on empirical spectral distributions.*

LEMMA 3.3. *Let W_n be an $n \times n$ symmetric matrix and W_{n-1} be an $(n - 1) \times (n - 1)$ major submatrix of W_n . Denote the spectral distributions of W_n and W_{n-1} by F_n and F_{n-1} , respectively. Then, we have*

$$\|nF_n - (n - 1)F_{n-1}\| \leq 1.$$

PROOF. Denote the eigenvalues of the matrices W_n and W_{n-1} by $\lambda_1 \leq \cdots \leq \lambda_n$ and $\mu_1 \leq \cdots \leq \mu_{n-1}$, respectively. Then, the lemma follows from the following well-known fact:

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \cdots \leq \mu_{n-1} \leq \lambda_n. \quad \square$$

4. Convergence rates of expected spectral distributions of Wigner matrices. In this section, we shall apply the inequality of Theorem 2.1 to establish a convergence rate of the spectral distributions of high dimensional Wigner matrices. A Wigner matrix $W_n = (x_{ij}(n))$, $i, j = 1, \dots, n$, is defined to be a symmetric matrix with independent entries on and above the diagonal. Throughout this section, we shall drop the index n from the entries of W_n and

assume that the following conditions hold:

$$\begin{aligned}
 (4.1) \quad & \text{(i)} \quad Ex_{ij} = 0, \quad \text{for all } 1 \leq i \leq j \leq n; \\
 & \text{(ii)} \quad Ex_{ij}^2 = 1, \quad \text{for all } 1 \leq i < j \leq n; \\
 & \quad \quad Ex_{ii}^2 = \sigma^2, \quad \text{for all } 1 \leq i \leq n; \\
 & \text{(iii)} \quad \sup_n \sup_{1 \leq i \leq j \leq n} Ex_{ij}^4 \leq M < \infty.
 \end{aligned}$$

Denote by F_n the empirical spectral distribution of $(1/\sqrt{n})W_n$. Under the conditions given in (4.1), it is well known that $F_n \rightarrow_w F$ in probability, where F is the limiting spectral distribution of F_n , known as Wigner’s semicircular law, that is,

$$(4.2) \quad F(x) = \frac{1}{2\pi} \int_{-\infty}^x \sqrt{4 - y^2} I_{[-2, 2]}(y) dy.$$

If W_n is the $n \times n$ submatrix of the upper-left corner of an infinite dimensional random matrix $[x_{ij}, i, j = 1, 2, \dots]$, then the convergence is almost sure (a.s.) [see Girko (1975) or Pastur (1972)].

In this section, we shall establish the following theorem.

THEOREM 4.1. *Under assumptions (4.1), we have*

$$(4.3) \quad \|EF_n - F\| = O(n^{-1/4}).$$

REMARK 4.1. In Section 3 of Girko (1989), an estimate of the difference between the expected Stieltjes transform of the spectral distribution F_n of Wigner matrices and that of the limiting spectral distribution F is established. In his proof, some arguments are not easily verifiable. If the proof is correct, then his result implies

$$\|EF_n - F\| = O(n^{-\gamma/14}), \quad \text{for some } 0 < \gamma < 1,$$

by applying Theorem 2.1. The result of Theorem 4.1 is stronger than that implied by Girko’s Theorem 3.1.

REMARK 4.2. It may be of greater interest to establish a convergence rate of $\|F_n - F\|$. This is under further investigation. In the proof of Theorem 4.1, one may find that the terms in the expansion of the Stieltjes transform of EF_n have a step-decreasing rate of n^{-1} if the estimation of the remainder term is not taken into account. Thus, we may conjecture that the ideal convergence rate of $\|EF_n - F\|$ is $O(n^{-1})$. Based on experience [say, for functions of sample means, the rate of expected bias is of $O(1/n)$, but $\sqrt{n}(f(\bar{X}_n) - f(\mu)) \rightarrow N(0, \sigma^2)$], one may conjecture that of $\|F_n - F\|$ is $O_p(n^{-1/2})$. But I was told through private communication that J. W. Silverstein conjectured that the rate for both cases is $O(n^{-1})$.

The proof of Theorem 4.1 is somewhat tedious. We first prove a preliminary result and then refine it.

PROPOSITION 4.2. *Under the assumptions of Theorem 4.1, we have*

$$(4.4) \quad \|EF_n - F\| = O(n^{-1/6}).$$

PROOF. It is shown in (3.2) that the Stieltjes transform of F is given by

$$(4.5) \quad s(z) = -\frac{1}{2}\{z - \sqrt{z^2 - 4}\}.$$

Let u and $v > 0$ be real numbers and let $z = u + iv$. Set

$$(4.6) \quad s_n(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dEF_n(x) = \frac{1}{n} E \operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n - zI_n \right)^{-1}.$$

Then, by the inverse matrix formula [see (3.8)], we have

$$(4.7) \quad \begin{aligned} s_n(z) &= \frac{1}{n} \sum_{k=1}^n E \frac{1}{(1/\sqrt{n})x_{kk} - z - (1/n)\alpha'(k)((1/\sqrt{n})W_n(k) - zI_{n-1})^{-1}\alpha(k)} \\ &= \frac{1}{n} \sum_{k=1}^n E \frac{1}{\varepsilon_k - z - s_n(z)} = -\frac{1}{z + s_n(z)} + \delta, \end{aligned}$$

where $\alpha'(k) = (x_{1k}, \dots, x_{k-1,k}, x_{k+1,k}, \dots, x_{nk})$, $W_n(k)$ is the matrix obtained from W_n by deleting the k th row and k th column,

$$(4.8) \quad \varepsilon_k = \frac{1}{\sqrt{n}}x_{kk} - \frac{1}{n}\alpha'(k) \left(\frac{1}{\sqrt{n}}W_n(k) - zI_{n-1} \right)^{-1} \alpha(k) + s_n(z)$$

and

$$(4.9) \quad \delta = \delta_n = -\frac{1}{n} \sum_{k=1}^n E \frac{\varepsilon_k}{(z + s_n(z))(z + s_n(z) - \varepsilon_k)}.$$

Solving (4.7), we obtain

$$(4.10) \quad s_{(1),(2)}(z) = -\frac{1}{2}\left(z - \delta \pm \sqrt{(z + \delta)^2 - 4}\right).$$

We claim that

$$(4.11) \quad s_n(z) = s_{(2)}(z) = -\frac{1}{2}\left(z - \delta - \sqrt{(z + \delta)^2 - 4}\right).$$

Note that

$$(4.12) \quad \operatorname{Im}(z + s_n(z)) = v \left(1 + E \int \frac{1}{(x - u)^2 + v^2} dF_n(x) \right) \geq v,$$

which immediately yields

$$(4.13) \quad |z + s_n(z)|^{-1} \leq v^{-1}.$$

By definition, it is obvious that

$$(4.14) \quad |s_n(z)| \leq v^{-1}.$$

Hence by (4.7),

$$(4.15) \quad |\delta| \leq 2/v.$$

We conclude that (4.11) is true for all $v > \sqrt{2}$ because, by (4.15), $\text{Im}(s_{(1)}(z)) \leq -(1/2)(v - |\delta|) \leq -(1/2)(v - 2/v) < 0$, which contradicts the fact that $\text{Im}(s_n(z)) > 0$. By definition, $s_n(z)$ is a continuous function of z on the upper half-plane $\{z = u + iv : v > 0\}$. By (4.13), $(z + s_n(z))^{-1}$ is also continuous. Hence, δ , and consequently, $s_{(1)}(z)$ and $s_{(2)}(z)$ are continuous on the upper half-plane. Therefore, to prove $s_n(z) \neq s_{(1)}(z)$, or equivalently the assertion (4.11), it is sufficient to show that the two continuous functions $s_{(1)}(z)$ and $s_{(2)}(z)$ cannot be equal at any point on the upper half-plane. If $s_{(1)}(z) = s_{(2)}(z)$ for some $z = u + iv$ with $v > 0$, then the square root in (4.11) should be zero, that is, $\delta = \pm 2 - z$. This implies that $s_n(z) = \pm 1 - z$, which contradicts the fact that $\text{Im}(s_n(z)) > 0$. This completes the proof of our assertion (4.11).

Comparing (4.5) and (4.11), we shall prove Proposition 4.2 by the following steps: Prove $|\delta|$ is “small” for both its absolute value and for the integral of its absolute value with respect to u . Then, find a bound of $s_n(z) - s(z)$ in terms of δ . First, let us begin to estimate $|\delta|$.

Applying (3.12), we have

$$(4.16) \quad |z + s_n(z) - \varepsilon_k|^{-1} \leq v^{-1}.$$

By (4.9), we have

$$(4.17) \quad |\delta| \leq \frac{1}{n} \sum_{k=1}^n \left[\frac{|E\varepsilon_k|}{|z + s_n(z)|^2} + E \left| \frac{\varepsilon_k^2}{(z + s_n(z))^2(z + s_n(z) - \varepsilon_k)} \right| \right]$$

$$\leq (z + s_n(z))^{-2} \left[\frac{1}{n} \sum_{k=1}^n (|E\varepsilon_k| + v^{-1}E|\varepsilon_k^2|) \right]$$

$$(4.18) \quad \leq \frac{1}{nv^2} \left(\sum_{k=1}^n |E\varepsilon_k| + v^{-1} \sum_{k=1}^n E|\varepsilon_k^2| \right).$$

Recalling the definition of ε_k in (4.8) and applying (3.11), we obtain

$$(4.19) \quad |E\varepsilon_k| = \frac{1}{n} \left| E \left[\text{tr} \left(\frac{1}{\sqrt{n}} W_n - zI_n \right)^{-1} - \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right] \right|$$

$$\leq \frac{1}{nv}.$$

Now, we begin to estimate $E|\varepsilon_k|^2$. By (4.8), we have

$$\begin{aligned}
 E|\varepsilon_k|^2 &= E|\varepsilon_k - E\varepsilon_k|^2 + |E(\varepsilon_k)|^2 \\
 &= \frac{\sigma^2}{n} + \frac{1}{n} E \left| \alpha'_k \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \alpha_k \right. \\
 &\quad \left. - \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right|^2 \\
 (4.20) \quad &+ E \left| \frac{1}{n} \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right. \\
 &\quad \left. - \frac{1}{n} E \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right|^2 \\
 &+ |E\varepsilon_k|^2.
 \end{aligned}$$

Let

$$(r_{ij}(k)) = \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1}.$$

Then, we have

$$\begin{aligned}
 &E \left| \frac{1}{n} \alpha'_k \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \alpha_k - \frac{1}{n} \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right|^2 \\
 &= E \left| \frac{1}{n} \sum_{jl} r_{jl}(k) (x_{jk} x_{lk} - E(x_{jk} x_{lk})) \right|^2 \\
 (4.21) \quad &\leq \frac{2M}{n^2} \sum_{i,j} E|r_{ij}^2(k)| \\
 &= \frac{2M}{n^2} E \text{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \left(\frac{1}{\sqrt{n}} W_n(k) - \bar{z}I_{n-1} \right)^{-1} \\
 &= \frac{2M}{n^2} E \text{tr} \left(\left(\frac{1}{\sqrt{n}} W_n(k) - uI_{n-1} \right)^2 + v^2 I_{n-1} \right)^{-1} \\
 &\leq \frac{2M}{nv^2}.
 \end{aligned}$$

Let $\gamma_k(k) = 0$ and for $d \neq k$, let

$$\begin{aligned} \gamma_d(k) &= E_{d-1} \operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} - E_d \operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \\ &= E_{d-1} \sigma_d(k) - E_d \sigma_d(k), \end{aligned}$$

where

$$(4.22) \quad \sigma_d(k) = \left[\operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} - \left(\frac{1}{\sqrt{n}} W(d, k) - zI_{n-2} \right)^{-1} \right],$$

E_d denotes the conditional expectation given $\{x_{ij}, d + 1 \leq i \leq j \leq n\}$, $\alpha(d, k)$ is the vector obtained from the d th column of W_n by deleting the d th and k th entries and $W_n(d, k)$ the matrix obtained from W_n by deleting the d th and k th columns and rows. By (3.11), we have

$$(4.23) \quad |\sigma_d(k)| \leq v^{-1},$$

which implies that

$$(4.24) \quad \begin{aligned} & E \left| \frac{1}{n} \operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} - \frac{1}{n} E \operatorname{tr} \left(\frac{1}{\sqrt{n}} W_n(k) - zI_{n-1} \right)^{-1} \right|^2 \\ & \leq n^{-2} \sum_{d=1}^n E |\gamma_d^2(k)| \leq n^{-1} v^{-2}. \end{aligned}$$

By (4.8), (4.19)–(4.21) and (4.24), we obtain, for all large n ,

$$(4.25) \quad E |\varepsilon_k|^2 \leq \frac{\sigma^2}{n} + \frac{2M}{nv^2} + \frac{1}{nv^2} + \frac{1}{n^2v^2} < \frac{2M + 5}{nv^2},$$

where M is the upper bound of the fourth moments of the entries of W_n in (4.1).

Take $v = ((2M + 6)/n)^{1/6}$ and assume $n > 2M + 6$. From (4.18), (4.19) and (4.24)–(4.25), we conclude that

$$(4.26) \quad |\delta| < \frac{2M + 6}{nv^5} = v.$$

By (4.7), (4.17), (4.19) and (4.25), for large n so that $v < 1/3$, we have

$$\begin{aligned} \int_{-\infty}^{\infty} |\delta| du &\leq \frac{2M + 6}{nv^3} \left[\int_{-\infty}^{\infty} |z + s_n(z)|^{-2} du \right] \\ &\leq \frac{4M + 12}{nv^3} \left[\int_{-\infty}^{\infty} |s_n(z)|^2 du + \int_{-\infty}^{\infty} |\delta|^2 du \right] \\ &\leq \frac{4M + 12}{nv^3} \left[\int_{-\infty}^{\infty} |s_n(z)|^2 du + v \int_{-\infty}^{\infty} |\delta| du \right]. \end{aligned}$$

By the simple fact that $y \leq ax + by$ implies $ax + by \leq (a/(1 - b))x$ for positive $a, b < 1, x$ and y , we get, for large n so that $nv^2 > (2M + 7)(4M + 12)$,

$$\begin{aligned}
 \int_{-\infty}^{\infty} |\delta| du &\leq \frac{4M + 12}{nv^3(1 - (4M + 12)/(nv^2))} \int_{-\infty}^{\infty} |s_n(z)|^2 du \\
 (4.27) \quad &\leq \frac{4M + 14}{nv^3} E \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{(x - u)^2 + v^2} du dF_n(x) \\
 &\leq \frac{4M + 14}{nv^4} \leq 2v^2.
 \end{aligned}$$

Now, we proceed to estimate $|s_n(z) - s(z)|$. By (4.5) and (4.11),

$$(4.28) \quad |s_n(z) - s(z)| \leq \frac{1}{2} |\delta| \left[1 + \frac{|2z + \delta|}{\left| \sqrt{z^2 - 4} + \sqrt{(z + \delta)^2 - 4} \right|} \right].$$

Since the signs of the real parts of $\sqrt{z^2 - 4}$ and $\sqrt{(z + \delta)^2 - 4}$ [see (3.1)] are $\text{sign}(uv) = \text{sign}(u)$ and $\text{sign}((u + \text{Re}(\delta))(v + \text{Im}(\delta))) = \text{sign}(u + \text{Re}(\delta))$, we conclude that for $|u| > v$ the real and imaginary parts of $\sqrt{z^2 - 4}$ and $\sqrt{(z + \delta)^2 - 4}$ have the same signs. Hence, by (4.28) we have

$$(4.29) \quad |s_n(z) - s(z)| \leq \frac{1}{2} |\delta| \left(1 + \frac{2|u| + 3v}{\sqrt{|u^2 - v^2 - 4|}} \right).$$

For $|u| > 4$ and n such that $v < 1/3$, we have

$$\frac{2|u| + 3v}{\sqrt{|u^2 - v^2 - 4|}} \leq \frac{8 + 3v}{\sqrt{12 - v^2}} < 3,$$

which, together with (4.29), implies that

$$(4.30) \quad |s_n(z) - s(z)| \leq 2|\delta|.$$

For $|u| \leq v$, by (4.26) we have $|\sqrt{(z + \delta)^2 - 4} - 2i| \leq (9/2)v^2$. Similarly we have $|\sqrt{z^2 - 4} - 2i| \leq 2v^2$. Therefore, we have for all n such that $v < 1/3$,

$$(4.31) \quad |s_n(z) - s(z)| \leq \frac{1}{2} |\delta| \left(1 + \frac{5v}{4 - 7v^2} \right) \leq 2|\delta|.$$

Summing up (4.29)–(4.31), we get that for n so large that $v < 1/3$,

$$(4.32) \quad |s_n(z) - s(z)| \leq \begin{cases} \frac{1}{2}|\delta| \left[1 + \frac{2|u| + 3v}{\sqrt{|u^2 - v^2 - 4|}} \right], & \text{if } v < |u| \leq 4, \\ 2|\delta|, & \text{otherwise.} \end{cases}$$

Finally, by (4.28), (4.27) and (4.32),

$$(4.33) \quad \begin{aligned} & \int_{-\infty}^{\infty} |s_n(z) - s(z)| du \\ &= \left\{ \int_{[|u| \leq v]} + \int_{[v < |u| \leq 4]} + \int_{[|u| > 4]} \right\} |s_n(z) - s(z)| du \\ &\leq 2 \int_{-\infty}^{\infty} |\delta| du + 4 \int_{-4}^4 |\delta| \frac{1}{\sqrt{|u^2 - v^2 - 4|}} du \\ &\leq 4v^2 + \eta v, \end{aligned}$$

where

$$\eta = 4 \sup_{v < 1} \int_{-4}^4 \frac{du}{\sqrt{|u^2 - v^2 - 4|}}.$$

Note that the density function of the semicircular law is bounded by $1/\pi$. An application of Theorem 2.1 completes the proof of Proposition 4.2. \square

Now, we are in position to prove Theorem 4.1. The basic approach to prove Theorem 4.1 is similar to that in the proof of Proposition 4.2. The only work needed to do is to refine the estimates of $E|\varepsilon_k^2|$ and the integral of $|\delta|$ by using the preliminary result of Proposition 4.2.

PROOF OF THEOREM 4.1. Denote by Δ_1 the initial estimate of the convergence rate of $\|EF_n - F\|$. By Proposition 4.2, we may choose $\Delta_1 = C_0 n^{-1/6}$ for some positive constant $C_0 \geq 1$.

Choose $v = Dn^{-1/4}$, where D is a positive constant to be specified later. Suppose that n is so large that $v < \Delta_1$.

For later use, let us derive an estimate of $|z + s_n(z)|^{-2}$.

For any two Stieltjes transforms $s_1(z)$ and $s_2(z)$ with their corresponding distributions F_1 and F_2 , integration by parts yields

$$(4.34) \quad \begin{aligned} |s_1(z) - s_2(z)| &= \left| \int_{-\infty}^{\infty} \frac{d(F_1(x) - F_2(x))}{x - z} \right| \\ &= \left| \int_{-\infty}^{\infty} \frac{(F_1(x) - F_2(x)) dx}{(x - z)^2} \right| \leq \frac{\pi \|F_1 - F_2\|}{v}. \end{aligned}$$

Then, by (4.7), (3.3) and (4.34), it follows that

$$(4.35) \quad \begin{aligned} |z + s_n(z)|^{-2} &\leq 3|s(z)|^2 + 3|\delta|^2 + 3|s_n(z) - s(z)|^2 \\ &\leq 3|\delta|^2 + 4\pi^2 v^{-2} \Delta_1^2. \end{aligned}$$

Now, we begin to refine the estimate of $(1/n)\sum_{i,j} E|r_{ij}^2(k)|$. By (4.21) and Lemma 3.3, we obtain that, for large n ,

$$(4.36) \quad \begin{aligned} \frac{1}{n} \sum_{i,j} E|r_{ij}^2(k)| &= \frac{n-1}{n} \int_{-\infty}^{\infty} \frac{1}{(x-u)^2 + v^2} dE F_{n-1}^{(k)}(x) \\ &\leq \left[\int_{-\infty}^{\infty} \frac{1}{(x-u)^2 + v^2} dF(x) \right. \\ &\quad \left. + \left| \int_{-\infty}^{\infty} \frac{2(x-u)(EF_n(x) - F(x))}{((x-u)^2 + v^2)^2} dx \right| \right. \\ &\quad \left. + \left| \int_{-\infty}^{\infty} \frac{2(x-u)E(((n-1)/n)F_{n-1}^{(k)}(x) - F_n(x))}{((x-u)^2 + v^2)^2} dx \right| \right] \\ &\leq \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{(x-u)^2 + v^2} dx + \left(\Delta_1 + \frac{1}{n} \right) \int_{-\infty}^{\infty} \frac{2|x-u| dx}{((x-u)^2 + v^2)^2} \\ &= v^{-1} + 2 \left(\Delta_1 + \frac{1}{n} \right) v^{-2} < \frac{4\Delta_1}{v^2}, \end{aligned}$$

where $F_{n-1}^{(k)}$ denotes the spectral distribution of the matrix $(1/\sqrt{n})W_n(k)$.

To get a refinement of the estimate of $E|\gamma_d^2(k)|$, we introduce the following notation. Set

$$\begin{aligned} R_d(k) &= \frac{1}{\sqrt{n}} W_n(d, k) - zI_{n-2}, \\ s_{nd(k)}(z) &= \frac{1}{n} E(\text{tr}(R_d^{-1}(k))), \\ \varepsilon_d(k) &= \frac{1}{\sqrt{n}} x_{dd} - \alpha'(d, k) R_d^{-1}(k) \alpha(d, k) + s_{nd}(z) \end{aligned}$$

and

$$R_d^{-2}(k) = (\rho_{ij}(d, k)),$$

where $W_n(d, k)$ and $\alpha(d, k)$ are defined below (4.22).

By (4.22), we have

$$\sigma_d(k) = - \frac{1 + (1/n)\alpha'(d, k)R_d^2(k)\alpha(d, k)}{z + s_{nd(k)}(z) - \varepsilon_d(k)}.$$

Then, rewrite $\sigma_d(k)$ [see (4.22)] in the following form:

$$\begin{aligned} \sigma_d(k) = & - \frac{1 + n^{-1} \operatorname{tr}(R_d^{-2}(k))}{z + s_{nd(k)}(z)} \\ (4.37) \quad & - \frac{\varepsilon_d(k)(1 + (1/n)\alpha'(d, k)R_d^{-2}(k)\alpha(d, k))}{(z + s_{nd(k)}(z))(z + s_{nd(k)}(z) - \varepsilon_d(k))} \\ & - \frac{(1/n)\alpha'(d, k)R_d^{-2}(k)\alpha(d, k) - (1/n)\operatorname{tr}(R_d^{-2}(k))}{z + s_{nd(k)}(z)}. \end{aligned}$$

Note that

$$(4.38) \quad E_{d-1} \frac{1 + n^{-1} \operatorname{tr}(R_d^{-2}(k))}{z + s_{nd(k)}(z)} - E_d \frac{1 + n^{-1} \operatorname{tr}(R_d^{-2}(k))}{z + s_{nd(k)}(z)} = 0$$

and by (4.23),

$$(4.39) \quad \left| \frac{1 + (1/n)\alpha'(d, k)R_d^{-2}(k)\alpha(d, k)}{z + s_{nd(k)}(z) - \varepsilon_d(k)} \right| = |\sigma_d(k)| \leq \frac{1}{v}.$$

We have

$$(4.40) \quad E|\gamma_d^2(k)| \leq \frac{2E|\varepsilon_d(k)|^2}{v^2|z + s_{nd(k)}(z)|^2} + \frac{4M \sum_{i,j} E|\rho_{ij}(d, k)|^2}{n^2|z + s_{nd(k)}(z)|^2}.$$

Similarly to estimating $n^{-1} \sum |r_{ij}^2(k)|$ in (4.21), one may obtain

$$(4.41) \quad \frac{1}{n} \sum_{i,j} |\rho_{ij}^2(d, k)| \leq v^{-4}.$$

Let $F_{n-2,d,k}$ denote the spectral distribution of the matrix $(1/\sqrt{n})W_n(d, k)$. Then, by Lemma 3.3, we have $\|((n-2)/n)F_{n-2,d,k} - F_n\| \leq 2/n$. Therefore, for all n ,

$$\begin{aligned} |s_{nd(k)}(z) - s_n(z)| &= \left| \int_{-\infty}^{\infty} \frac{(EF_n(x) - ((n-2)/n)EF_{n-2,d,k}(x)) dx}{(x-z)^2} \right| \\ (4.42) \quad &\leq \frac{2\pi}{nv}. \end{aligned}$$

Thus, by (3.3), (4.7), (4.13), (4.34) and (4.42), we have for all large n ,

$$\begin{aligned}
 |z + s_{nd(k)}(z)|^{-2} &\leq 4 \left\{ \left| \frac{s_n(z) - s_{nd(k)}(z)}{(z + s_n(z))(z + s_{nd(k)}(z))} \right|^2 + \left| \frac{1}{z + s_n(z)} + s_n(z) \right|^2 \right. \\
 &\qquad \qquad \qquad \left. + |s_n(z) - s(z)|^2 + |s(z)|^2 \right\} \\
 (4.43) \qquad &\leq 4 \left\{ \frac{4\pi^2}{n^2v^4} |z + s_{nd(k)}(z)|^{-2} + |\delta|^2 + \frac{\pi^2\Delta_1^2}{v^2} + 1 \right\} \\
 &\leq 5 \left(|\delta|^2 + \frac{\pi^2\Delta_1^2}{v^2} \right).
 \end{aligned}$$

Similarly to estimating $E|\varepsilon_k|^2$ in (4.25), one may obtain for large n that

$$(4.44) \qquad E|\varepsilon_d^2(k)| \leq \frac{2M + 5}{nv^2}.$$

Hence, by (4.40)–(4.44), we have

$$(4.45) \qquad E|\gamma_d^2(k)| \leq \frac{40M + 50}{nv^4} \left(|\delta|^2 + \frac{\pi^2\Delta_1^2}{v^2} \right).$$

Substituting (4.19), (4.36) and (4.45) into (4.25), we obtain, for all large n ,

$$\begin{aligned}
 (4.46) \qquad E|\varepsilon_k|^2 &\leq \frac{\sigma^2}{n} + \frac{8M\Delta_1}{nv^2} + \frac{40M + 50}{n^2v^4} \left(|\delta|^2 + \frac{\pi^2\Delta_1^2}{v^2} \right) + \frac{1}{n^2v^2} \\
 &\leq \frac{40M + 60}{n^2v^4} |\delta|^2 + \frac{(8M + 3)\Delta_1}{nv^2}.
 \end{aligned}$$

Consequently, by (4.17), (4.19), (4.35) and (4.46),

$$\begin{aligned}
 (4.47) \qquad |\delta| &\leq |z + s_n(z)|^{-2} \left(\frac{1}{n} \sum_{k=1}^n [|E\varepsilon_k| + v^{-1}E|\varepsilon_k^2|] \right) \\
 &\leq |z + s_n(z)|^{-2} \left[\frac{40M + 50}{n^2v^5} |\delta|^2 + \frac{(8M + 4)\Delta_1}{nv^3} \right] \\
 &\leq \left[\frac{40M + 50}{n^2v^5} |\delta|^2 + \frac{(8M + 4)\Delta_1}{nv^3} \right] \left(3|\delta|^2 + \frac{4\pi^2\Delta_1^2}{v^2} \right).
 \end{aligned}$$

From (4.15) and (4.47), it follows that

$$\begin{aligned}
 (4.48) \quad |\delta| &\leq \left[\frac{250(4M + 5)}{D^8} + \frac{24(2M + 1)\Delta_1}{D^4} + \frac{80(4M + 5)\pi^2\Delta_1^2}{n^2v^6} \right] |\delta| \\
 &\quad + \frac{6\pi^2(2M + 1)\Delta_1^3}{nv^5} \\
 &\leq \frac{300(4M + 5)}{D^8} |\delta| + \frac{6\pi^2(2M + 1)\Delta_1^3}{nv^5}.
 \end{aligned}$$

If D is chosen so large that $300(4M + 5)D^{-8} \leq 1/2$, again using the fact used in the proof of (4.27), we obtain

$$(4.49) \quad |\delta| \leq \frac{12\pi^2(2M + 1)\Delta_1^3}{nv^5} < v.$$

By (4.5), (4.11) and (4.48) we have, for large n ,

$$(4.50) \quad |s_n(z) - s(z)| = |s(z + \delta) - s(z) + \delta| \leq 2 + v \leq 3.$$

Thus, by (4.49) and the second inequality of (4.47),

$$(4.51) \quad |\delta| \leq \frac{(8M + 5)\Delta_1}{nv^3} |z + s_n(z)|^{-2}.$$

Therefore, by (4.50), we have

$$\begin{aligned}
 (4.52) \quad \int_{-\infty}^{\infty} |\delta| du &\leq \frac{(8M + 5)\Delta_1}{nv^3} \int_{-\infty}^{\infty} |z + s_n(z)|^{-2} du \\
 &\leq 3 \frac{(8M + 5)\Delta_1}{nv^3} \left[\int_{-\infty}^{\infty} |s_n(z) - s(z)|^2 du \right. \\
 &\quad \left. + \int_{-\infty}^{\infty} |s(z)|^2 du + \int_{-\infty}^{\infty} |\delta|^2 du \right] \\
 &\leq 3 \frac{(8M + 5)\Delta_1}{nv^3} \left[\int_{-\infty}^{\infty} |s_n(z) - s(z)|^2 du + 2\pi + v \int_{-\infty}^{\infty} |\delta| du \right] \\
 &\leq 3 \frac{(8M + 6)\Delta_1}{nv^3} \left[2\pi + 3 \int_{-\infty}^{\infty} |s_n(z) - s(z)| du \right],
 \end{aligned}$$

where an upper bound for the integral of $|s(z)|^2$ is established in Corollary 3.2 in Section 3.

Recall the proof of the first inequality of (4.33), where the only condition used is that $|\delta| < v$. Therefore, by (4.49) we have

$$\begin{aligned}
 \int_{-\infty}^{\infty} |s_n(z) - s(z)| du &\leq 2 \int_{-\infty}^{\infty} |\delta| du + \eta v \\
 (4.53) \qquad &\leq (\eta + 12\pi\Delta_1(8M + 6)D^{-4})v \\
 &\quad + 18 \frac{(8M + 6)\Delta_1}{nv^3} \int_{-\infty}^{\infty} |s_n(z) - s(z)| du \\
 &\leq 2(\eta + 12\pi\Delta_1(8M + 6)D^{-4})v.
 \end{aligned}$$

Applying Theorem 2.1, the proof is complete. \square

Acknowledgments. The author would like to express his thanks to Professor Alan J. Izenman for his help in the preparation of the paper, and also thanks to Professor J. W. Silverstein and an Associate Editor for their helpful comments.

REFERENCES

- ARNOLD, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.* **20** 262–268.
- BAI, Z. D. (1993). Convergence rate of expected spectral distributions of large random matrices. Part II. Sample covariance matrices. *Ann. Probab.* **21** 649–672.
- BAI, Z. D., SILVERSTEIN, J. W. and YIN, Y. Q. (1988). A note on the largest eigenvalue of a large dimensional sample matrix. *J. Multivariate Anal.* **26** 166–168.
- 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. (1990). Limit of the smallest eigenvalue of large dimensional sample covariance matrix. Technical Report 90-05, Center for Multivariate Analysis, Pennsylvania State Univ.
- BAI, Z. D., YIN, Y. Q. and KRISHNAIAH, P. R. (1986). On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. *J. Multivariate Anal.* **19** 189–200.
- BAI, Z. D., YIN, Y. Q. and KRISHNAIAH, P. R. (1987). On limiting spectral distribution function of the eigenvalues of a multivariate F matrix. *Teor. Veroyatnost. i Primenen. (Theory Probab. Appl.)* **32** 537–548.
- 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.
- GIRKO, V. L. (1975). *Random Matrices*. Vishcha Shkola, Kiev (in Russian).
- GIRKO, V. L. (1989). Asymptotics of the distribution of the spectrum of random matrices. *Russian Math. Surveys* **44** 3–36.
- GRENNANDER, U. (1963). *Probabilities on Algebraic Structures*. Almqvist and Wiksell, Stockholm.
- GRENNANDER, U. and SILVERSTEIN, J. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* **32** 499–519.
- JONSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38.
- PASTUR, L. A. (1972). On the spectrum of random matrices. *Teoret. Mat. Fiz.* **10** 102–112 (*Teoret. Mat. Phys.* **10** 67–74).

- PASTUR, L. A. (1973). Spectra of random self-adjoint operators. *Uspehi Mat. Nauk* **28** 4–63 (*Russian Math. Surveys* **28** 1–67).
- SILVERSTEIN, J. W. (1985a). The limiting eigenvalue distribution of a multivariate F matrix. *SIAM J. Math. Anal.* **16** 641–646.
- SILVERSTEIN, J. W. (1985b). The smallest eigenvalue of a large dimensional Wishart matrix. *Ann. Probab.* **13** 1364–1368.
- VOICULESCU, DAN (1990). Non-commutative random variables and spectral problems in free product C^* -algebras. *Rocky Mountain J. Math.* To appear.
- VOICULESCU, DAN (1991). Limit laws for random matrices and free products. *Invent. Math.* **104** 201–220.
- 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.
- WIGNER, E. P. (1955). Characteristic vectors bordered matrices with infinite dimensions. *Ann. of Math.* **62** 548–564.
- WIGNER, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Ann. of Math.* **67** 325–327.
- XU, Y. C. (1982). *An Introduction to Algebra*. Shanghai Sci. & Tech. Press, Shanghai, China (in Chinese).
- YIN, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20** 50–68.
- YIN, Y. Q., BAI, Z. D. and KRISHNAIAH, P. R. (1983). Limiting behavior of the eigenvalues of a multivariate F matrix. *J. Multivariate Anal.* **13** 508–516.
- 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.
- YIN, Y. Q. and KRISHNAIAH, P. R. (1983). A limit theorem for the eigenvalues of product of two random matrices. *J. Multivariate Anal.* **13** 489–507.

DEPARTMENT OF STATISTICS
 341 SPEAKMAN HALL
 TEMPLE UNIVERSITY
 PHILADELPHIA, PENNSYLVANIA 19122