The Annals of Probability
1998, Vol. 26, No. 1, 316–345

NO EIGENVALUES OUTSIDE THE SUPPORT OF THE LIMITING SPECTRAL DISTRIBUTION OF LARGE-DIMENSIONAL SAMPLE COVARIANCE MATRICES

By Z. D. Bai[1] and Jack W. Silverstein[2]

National University of Singapore and North Carolina State University

Dedicated to the 60th birthday of Min-Te Chao

Let $B_n = (1/N)T_n^{1/2} X_n X_n^* T_n^{1/2}$, where X_n is $n \times N$ with i.i.d. complex standardized entries having finite fourth moment and $T_n^{1/2}$ is a Hermitian square root of the nonnegative definite Hermitian matrix T_n. It is known that, as $n \to \infty$, if n/N converges to a positive number and the empirical distribution of the eigenvalues of T_n converges to a proper probability distribution, then the empirical distribution of the eigenvalues of B_n converges a.s. to a nonrandom limit. In this paper we prove that, under certain conditions on the eigenvalues of T_n, for any closed interval outside the support of the limit, with probability 1 there will be no eigenvalues in this interval for all n sufficiently large.

1. Introduction. For $n = 1, 2, \ldots$ let $X = X_n = (X_{ij})$, $T = T_n$ and $T_n^{1/2}$ denote, respectively, an $n \times N$ matrix consisting of i.i.d. standardized complex entries ($EX_{11} = 0$, $E|X_{11}|^2 = 1$), an $n \times n$ nonnegative definite matrix and any square root of T. For any square matrix A having real eigenvalues, let F^A denote the empirical distribution function (e.d.f.) of its eigenvalues. The matrix $B_n = (1/N)T_n^{1/2} XX^* T_n^{1/2}$ can be viewed as the sample covariance matrix of a broad class of random vectors, $T_n^{1/2} X_{\bullet 1}$ ($X_{\bullet j}$ denoting the jth column of X). Previous work on understanding the behavior of the eigenvalues of B_n when n and N are large but have the same order of magnitude has been on F^{B_n} and on the extreme eigenvalues when $T = I$, the identity matrix. Assuming $N = N(n)$ with $n/N \to c > 0$ as $n \to \infty$ and $F^T \to_{\mathscr{D}} H$, a proper p.d.f., it is known that almost surely F^{B_n} converges weakly to a nonrandom p.d.f. F [see Silverstein (1995)]. Proving this result, along with describing F (which can be explicitly expressed in only a few cases), is best achieved with the aid of the Stieltjes transform, defined for any p.d.f. G by

$$m_G(z) \equiv \int \frac{1}{\lambda - z}\, dG(\lambda), \qquad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C}\colon \operatorname{Im} z > 0\}.$$

Received July 1996.

[1]Research conducted at the Department of Applied Mathematics, National Sun Yat-sen University Kaohsiung, Taiwan.

[2]Supported by NSF Grant DMS-94-04047.

AMS 1991 *subject classifications.* Primary 60F15; secondary 62H99.

Key words and phrases. Random matrix, empirical distribution function of eigenvalues, Stieltjes transform.

Because of the inversion formula

$$G([a,b]) = \frac{1}{\pi} \lim_{\eta \to 0} \int_a^b \operatorname{Im} m_G(\xi + i\eta)\, d\xi$$

(a, b continuity points of G), weak convergence of p.d.f.'s can be proven by showing convergence of Stieltjes transforms.

For each $z \in \mathbb{C}^+$, $m = m_F(z)$ is a solution to the equation

$$m = \int \frac{1}{t(1-c-czm) - z}\, dH(t),$$

which is unique in the set $\{m \in \mathbb{C}\colon -(1-c)/z + cm \in \mathbb{C}^+\}$. Let $\underline{B}_n = (1/N)X^*TX$. Since the spectra of B_n and $\underline{B}_n$ differ by $|n-N|$ zero eigenvalues, it follows that

$$F^{\underline{B}_n} = \left(1 - \frac{n}{N}\right) I_{[0,\infty)} + \frac{n}{N} F^{B_n},$$

from which we get

$$m_{F^{\underline{B}_n}}(z) = -\frac{(1-n/N)}{z} + \frac{n}{N} m_{F^{B_n}}(z), \qquad z \in \mathbb{C}^+,$$

and, with $\underline{F}$ denoting the limit of $F^{\underline{B}_n}$, we have

$$\underline{F} = (1-c) I_{[0,\infty)} + cF$$

and

$$m_{\underline{F}}(z) = -\frac{(1-c)}{z} + c m_F(z), \qquad z \in \mathbb{C}^+.$$

It follows that

$$m_F = -z^{-1} \int \frac{1}{1 + t m_{\underline{F}}}\, dH(t)$$

for each $z \in \mathbb{C}^+$, $\underline{m} = m_{\underline{F}}(z)$ is the unique solution in $\mathbb{C}^+$ to the equation

$$\underline{m} = -\left(z - c \int \frac{t\, dH(t)}{1 + t\underline{m}}\right)^{-1} \tag{1.1}$$

and $m_{\underline{F}}(z)$ has an inverse, explicitly given by

$$z(\underline{m}) = -\frac{1}{\underline{m}} + c \int \frac{t\, dH(t)}{1 + t\underline{m}}. \tag{1.2}$$

Much of the analytic behavior of F can be inferred from these equations [see Silverstein and Choi (1995)]. Indeed, continuous dependence of F on c and H is readily apparent from (1.2) and the inversion formula, and it can be shown that $F \to_{\mathscr{D}} H$ as $c \to 0$. Moreover, it is shown in Silverstein and Choi (1995) that, away from zero, F has a continuous density. As an example Figure 1a is the graph of the density when $c = 0.1$ and H places mass 0.2, 0.4 and 0.4 at, respectively, 1, 3 and 10.

The focus of this paper is on intervals of $\mathbb{R}^+$ lying outside the support of F. The inverse (1.2) can be used to identify these intervals, mainly because, on any such interval, $m_{\underline{F}}$ exists and is increasing. Consequently, its inverse

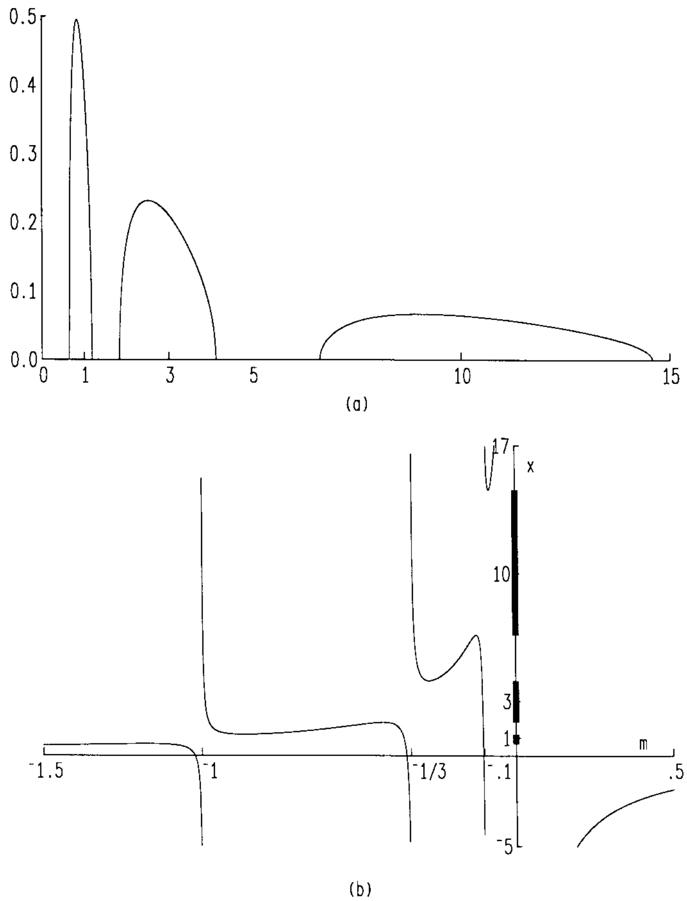

FIG. 1. (a) *Graph of the limiting density when* $c = 0.1$ *and* H *places mass* 0.2, 0.4 *and* 0.4 *at, respectively*, 1, 3 *and* 10. (b) *Graph of* $x = -m^{-1} + c\int t(1+tm)^{-1}\,dH(t)$ *corresponding to* (a). *The bold lines on the vertical axis indicate the support of the density, the set in* $\mathbb{R}^+$ *remaining after removing intervals where the graph is increasing. Using the fact that the density at* $x \in \mathbb{R}^+$ *is equal to* $(c\pi)^{-1}$ *times the imaginary part of* $m_{\underline{F}}(x)$ [*see Silverstein and Choi* (1995)], *the graph in* (a) *was created by applying Newton's method to* (1.2) *for values of* $z = x$ *in the support.*

will also exist and will be increasing on the range of this interval. Silverstein and Choi (1995) confirm each $\underline{m}$ in this range is such that $-1/\underline{m}$ lies outside the support of H. Therefore, plotting (1.2) on $\mathbb{R}$ and observing the range of values where it is increasing will yield the complement of the support of $\underline{F}$ and, together with c (to determine whether there is any mass at zero), the complement of the support of F. Figure 1b provides an illustration. It is the graph of (1.2) corresponding to the density in Figure 1a.

For large n one would intuitively expect no eigenvalues to appear on a closed interval outside the support of F. This, of course, cannot be inferred from the limiting result on F^{B_n}. The two important cases when $T = I$ have been settled. Here the support of F lies on $[(1-\sqrt{c})^2, (1+\sqrt{c})^2]$, with the addition of 0 when $c > 1$. When the entries of X come from the upper-left portion of a doubly infinite array of independent random variables having finite fourth moment, Yin, Bai and Krishnaiah (1988) and Bai and Yin (1993) show, respectively, the largest eigenvalue of B_n converges a.s. to $(1+\sqrt{c})^2$, and the $\min(n, N)$th largest (which is the smallest eigenvalue when $c < 1$) converges a.s. to $(1-\sqrt{c})^2$ [we remark here that in Bai, Silverstein and Yin (1988), it is proven that $E|X_{11}|^4 < \infty$ is necessary for the former to hold].

Extensive computer simulations, performed in order to show the importance of the spectral limiting results to the detection problem in array signal processing [Silverstein and Combettes (1992)], resulted in no eigenvalues appearing where there is no mass in the limit. Under reasonably mild conditions, this paper will provide a proof of this phenomenon, again in the form of a limit theorem as $n \to \infty$.

It will be necessary to impose stronger conditions on the eigenvalues of T_n than simply weak convergence of F^{T_n} to H. For this, if we let $F^{c,H}$ denote $\underline{F}$ and $c_n = n/N$, then F^{c_n, H_n} is the "limiting" nonrandom d.f. associated with the "limiting" ratio c_n and d.f. H_n. As will be seen, the conditions on H_n are reflected in F^{c_n, H_n}.

THEOREM 1.1. *Assume:*

(a) *X_{ij}, $i, j = 1, 2, \ldots$, are i.i.d. random variables in $\mathbb{C}$ with $EX_{11} = 0$, $E|X_{11}|^2 = 1$ and $E|X_{11}|^4 < \infty$.*

(b) *$N = N(n)$ with $c_n = n/N \to c > 0$ as $n \to \infty$.*

(c) *For each n, $T = T_n$ is $n \times n$ Hermitian nonnegative definite satisfying $H_n \equiv F^{T_n} \to_{\mathscr{D}} H$, a p.d.f.*

(d) *$\|T_n\|$, the spectral norm of T_n, is bounded in n.*

(e) *$B_n = (1/N) T_n^{1/2} X_n X_n^* T_n^{1/2}$, $T_n^{1/2}$ any Hermitian square root of T_n, $\underline{B}_n = (1/N) X_n^* T_n X_n$, where $X = X_n = (X_{ij})$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, N$.*

(f) *The interval $[a, b]$ with $a > 0$ lies outside the support of $F^{c,H}$ and F^{c_n, H_n} for all large n.*

Then $P($no eigenvalue of B_n appears in $[a, b]$ for all large n $) = 1$.

Using the results on the extreme eigenvalues of $(1/N)XX^*$, we see that the interval can also be unbounded. In particular, we have

COROLLARY. *If $\|T_n\|$ converges to the largest number in the support of H, then $\|B_n\|$ converges a.s. to the largest number in the support of F. If the smallest eigenvalue of T_n converges to the smallest number in the support of H, then $c < 1$ $(c > 1)$ implies the smallest eigenvalue of B_n $(\underline{B}_n)$ converges to the smallest number in the support of F $(\underline{F})$.*

Theorem 1.1 is proven by showing the convergence of Stieltjes transforms at an appropriate rate, uniform with respect to the real part of z over certain intervals, while the imaginary part of z converges to 0. Besides relying on standard results on matrices, the proof requires well-known bounds on moments of martingale difference sequences, as well as an extension of Rosenthal's inequality to random quadratic forms. The proof of the latter will be given in the Appendix. Statements of most of the mathematical tools needed will be given in the next section. Section 3 establishes a rate of convergence of $F^{\underline{B}_n}$, needed in proving the convergence of the Stieltjes transforms. The latter will be broken down into two parts (Sections 4 and 5), while Section 6 completes the proof.

It is mentioned here that Theorem 1.1 is actually only part of the important phenomena observed in simulations. It can be shown that on any interval J_H with endpoints outside the support of H, there corresponds, for c sufficiently small, an interval $J_{F,c}$ with endpoints being boundary points of the support of F satisfying $F(J_{F,c}) = H(J_H)$. This should be viewed in the finite but large dimensional case as the eigenvalues of B_n being a "smoothed" deformation of the eigenvalues of T_n, continuous in the ratio of dimension to sample size. Simulations reveal that the number of eigenvalues of B_n appearing in J_{F_n, c_n} is exactly the same as the number of eigenvalues of T_n in J_{H_n}. The formulation of the conjecture naturally arising from this is simply

$$n(F_n(J_{F,c}) - F^{c_n, H_n}(J_{F,c})) \to 0 \quad \text{a.s.}$$

Its truth is currently being investigated.

2. Mathematical tools. We list in this section results needed to prove Theorem 1.1. Throughout the rest of the paper constants appearing in inequalities are represented by K and occasionally subscripted with the variables they depend on. They are nonrandom and may take on different values from one appearance to the next.

The referenced results below concerning moments of sums of complex random variables were originally proven for real variables. Extension to the complex case is straightforward.

LEMMA 2.1 [Burkholder (1973)]. *Let $\{X_k\}$ be a complex martingale difference sequence with respect to the increasing σ-field $\{\mathscr{F}_k\}$. Then, for $p \geq 2$,*

$$E\Big|\sum X_k\Big|^p \leq K_p\Big(E\Big(\sum E(|X_k|^2|\mathscr{F}_{k-1})\Big)^{p/2} + E\sum |X_k|^p\Big).$$

LEMMA 2.2 [Burkholder (1973)]. *With $\{X_k\}$ as above, we have, for $p > 1$,*

$$E\Big|\sum X_k\Big|^p \leq K_p E\Big(\sum |X_k|^2\Big)^{p/2}.$$

LEMMA 2.3 [Rosenthal (1970)]. *If $\{X_k\}$ are independent nonnegative, then, for $p \geq 1$,*

$$E\Big(\sum X_k\Big)^p \leq K_p\Big(\Big(\sum EX_k\Big)^p + \sum EX_k^p\Big).$$

LEMMA 2.4 [Dilworth (1993)]. *With $\{\mathscr{F}_k\}$ as above, $\{X_k\}_{k\geq 1}$ a sequence of integrable random variables and $1 \leq q \leq p < \infty$, we have*

$$E\bigg(\sum_{k=1}^{\infty}\big|E(X_k|\mathscr{F}_k)\big|^q\bigg)^{p/q} \leq \bigg(\frac{p}{q}\bigg)^{p/q} E\bigg(\sum_{k=1}^{\infty}|X_k|^q\bigg)^{p/q}.$$

The following lemma is found in most probability textbooks.

LEMMA 2.5 (Kolmogorov's inequality for submartingales). *If $X_1, \dots, X_m$ is a submartingale, then, for any $\alpha > 0$,*

$$P\Big(\max_{k \le m} X_k \ge \alpha\Big) \le \frac{1}{\alpha} E(|X_m|).$$

The next one has a straightforward proof.

LEMMA 2.6. *If, for all $t > 0$, $P(|X| > t)t^p \le K$ for some positive p, then, for any positive $q < p$,*

$$E|X|^q \le K^{q/p}\left(\frac{p}{p-q}\right).$$

LEMMA 2.7. *For $X = (X_1, \dots, X_n)^T$ i.i.d. standardized (complex) entries, C $n \times n$ matrix (complex), we have, for any $p \ge 2$,*

$$E|X^*CX - \operatorname{tr} C|^p \le K_p\big((E|X_1|^4 \operatorname{tr} CC^*)^{p/2} + E|X_1|^{2p} \operatorname{tr}(CC^*)^{p/2}\big).$$

The proof is given in the Appendix.

LEMMA 2.8 [Corollary 7.3.8 of Horn and Johnson (1985)]. *For $r \times s$ matrices A and B with respective singular values $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_q$, $\tau_1 \ge \tau_2 \ge \cdots \ge \tau_q$, where $q = \min(r, s)$, we have*

$$|\sigma_k - \tau_k| \le \|B - A\| \quad \text{for all } k = 1, 2, \dots, q.$$

LEMMA 2.9 [(3.3.41) of Horn and Johnson (1991)]. *For $n \times n$ Hermitian $A = (a_{ij})$ with eigenvalues $\lambda_1, \dots, \lambda_n$, and convex f, we have*

$$\sum_{i=1}^{n} f(a_{ii}) \le \sum_{i=1}^{n} f(\lambda_i).$$

LEMMA 2.10 [Lemma 2.6 of Silverstein and Bai (1995)]. *Let $z \in \mathbb{C}^+$ with $v = \operatorname{Im} z$, A and B $n \times n$ with B Hermitian and $r \in \mathbb{C}^n$. Then*

$$\big|\operatorname{tr}\big((B - zI)^{-1} - (B + rr^* - zI)^{-1}\big)A\big| = \left|\frac{r^*(B - zI)^{-1}A(B - zI)^{-1}r}{1 + r^*(B - zI)^{-1}r}\right| \le \frac{\|A\|}{v}.$$

LEMMA 2.11 [Lemma 2.3 of Silverstein (1995)]. *For $z = x + iv \in \mathbb{C}^+$ let $m_1(z)$, $m_2(z)$ be Stieltjes transforms of any two p.d.f.'s, A and B $n \times n$ with A Hermitian nonnegative definite and $r \in \mathbb{C}^n$. Then:*

(a) $$\big\|(m_1(z)A + I)^{-1}\big\| \le \max(4\|A\|/v, 2),$$

(b) $$\begin{aligned}&\big|\operatorname{tr} B((m_1(z)A + I)^{-1} - (m_2(z)A + I)^{-1})\big| \\ &\qquad \le |m_2(z) - m_1(z)| n \|B\|\, \|A\| (\max(4\|A\|/v, 2))^2,\end{aligned}$$

(c)
$$\begin{aligned}&\left|r^*B(m_1(z)A+I)^{-1}r-r^*B(m_2(z)A+I)^{-1}r\right|\\&\quad\le|m_2(z)-m_1(z)|\,\|r\|^2\|B\|\,\|A\|(\max(4\|A\|/v,2))^2\end{aligned}$$

($\|r\|$ denoting the Euclidean norm on r).

LEMMA 2.12 [Lemma 2.4 of Silverstein and Bai (1995)]. *For $n\times n$ Hermitian A and B,*

$$\|F^A-F^B\|\le\frac{1}{n}\operatorname{rank}(A-B),$$

$\|\cdot\|$ here denoting the sup norm on functions.

Basic properties on matrices will be used throughout the paper, the two most common being: $\operatorname{tr}AB\le\|A\|\operatorname{tr}B$ for Hermitian nonnegative definite A and B, and for A $n\times n$ and $r\in\mathbb{C}^n$, for which both A and $A+rr^*$ are invertible,

$$r^*(A+rr^*)^{-1}=\frac{1}{(1+r^*A^{-1}r)}r^*A^{-1}.$$

At one point in Section 3 the two-dimensional Stieltjes transform is needed. Its definition and relevant properties are given here. For a p.d.f. $F(x,y)$ defined on $\mathbb{R}^2$, it is defined as

$$m(z_1,z_2)=\int\frac{1}{(x-z_1)(y-z_2)}\,dF(x,y)$$

for all $z_1=x_1+iv_1$, $z_2=x_2+iv_2$, $v_1\ne0$, $v_2\ne0$. Due to the inversion formula

$$\begin{aligned}F([a,b]\times[c,d])=-\frac{1}{\pi^2}\lim_{v\downarrow0}\int_{[a,b]\times[c,d]}&m(z_1,z_2)-m(\bar z_1,z_2)\\&-m(z_1,\bar z_2)+m(\bar z_1,\bar z_2)\,dx_1\,dx_2,\end{aligned}$$

$v_1=v_2=v$, whenever $F(\partial([a,b]\times[c,d]))=0$, weak convergence of p.d.f.'s on $\mathbb{R}^2$ is assured once convergence of their Stieltjes transforms is verified on a countable collection of points (z_1,z_2) dense in some open set in $\mathbb{C}^2$.

3. A rate on $F^{\underline{B}_n}$. We begin by simplifying our assumptions.

Because of assumption (d) in Theorem 1.1, we can assume $\|T_n\|=1$.

For $C>0$ let $Y_{ij}=X_{ij}I_{[|X_{ij}|\le C]}-EX_{ij}I_{[|X_{ij}|\le C]}$, $Y=(Y_{ij})$ and $\tilde B_n=(1/N)T_n^{1/2}Y_nY_n^*T_n^{1/2}$. Denote the eigenvalues of B_n and $\tilde B_n$ by λ_k and $\tilde\lambda_k$ (in decreasing order). Since these are the squares of the kth largest singular values of $(1/\sqrt{N})T_n^{1/2}X_n$ and $(1/\sqrt{N})T_n^{1/2}Y_n$ (respectively), we find, using Lemma 2.8,

$$\max_{k\le n}\left|\lambda_k^{1/2}-\tilde\lambda_k^{1/2}\right|\le(1/\sqrt{N})\left\|X_n-Y_n\right\|.$$

Since $X_{ij} - Y_{ij} = X_{ij} I_{[|X_{ij}|>C]} - E X_{ij} I_{[|X_{ij}|>C]}$, from Yin, Bai and Krishnaiah (1988) we have, with probability 1,

$$\limsup_{n\to\infty} \max_{k\le n} \left|\lambda_k^{1/2} - \tilde{\lambda}_k^{1/2}\right| \le (1+\sqrt{c}) E^{1/2} \left|X_{11}\right|^2 I_{[|X_{11}|>C]}.$$

Because of assumption (a) we can make the above bound arbitrarily small by choosing C sufficiently large. Thus, in proving Theorem 1.1 it is enough to consider the case where the underlying variables are uniformly bounded.

In this case it is proven in Yin, Bai and Krishnaiah (1988) that there exists a sequence $\{k_n\}$ satisfying $k_n/\log n \to \infty$ such that, for any $\eta > (1+\sqrt{c})^2$,

$$E\left\|(1/N) X_n X_n^*\right\|^{k_n} \le \eta^{k_n}$$

for all n sufficiently large. It follows then that $\lambda_{\max}$, the largest eigenvalue of B_n, satisfies

$$P\left(\lambda_{\max} \ge K\right) = o(N^{-\ell}) \tag{3.1}$$

for any $K > (1+\sqrt{c})^2$ and any positive ℓ.

Also, since $\operatorname{tr}(CC^*)^{p/2} \le (\operatorname{tr} CC^*)^{p/2}$, we get from Lemma 2.7 when X_1 is bounded

$$E|X_{\bullet 1}^* C X_{\bullet 1} - \operatorname{tr} C|^p \le K_p (\operatorname{tr} CC^*)^{p/2}, \tag{3.2}$$

where K_p also depends on the distribution of $X_{\bullet 1}$. From (3.2) we easily get

$$E|X_{\bullet 1}^* C X_{\bullet 1}|^p \le K_p ((\operatorname{tr} CC^*)^{p/2} + |\operatorname{tr} C|^p). \tag{3.3}$$

Throughout the paper, the variable $z = x + iv$ will be the argument of any Stieltjes transform. Let $m_n = m_{F^{B_n}}$ and $\underline{m}_n = m_{F^{\underline{B}_n}}$. For $j = 1, 2, \ldots, N$, let $q_j = (1/\sqrt{n}) X_{\bullet j}$ ($X_{\bullet j}$ denoting the jth column of X), $r_j = (1/\sqrt{N}) T_n^{1/2} X_{\bullet j}$ and $B_{(j)} = B_{(j)}^n = B_n - r_j r_j^*$.

In Silverstein (1995) the formula

$$\underline{m}_n(z) = -\frac{1}{N} \sum_{j=1}^{N} \frac{1}{z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)}$$

is derived. It is easy to verify

$$\operatorname{Im} r_j^*((1/z) B_{(j)} - I)^{-1} r_j \ge 0.$$

Therefore, for each j,

$$\frac{1}{|z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)|} \le \frac{1}{v}. \tag{3.4}$$

It is also shown in Silverstein (1995) that

$$\begin{aligned} &\frac{1}{n} \operatorname{tr}(-z\underline{m}_n(z) T_n - zI)^{-1} - m_n(z) \\ &\equiv w_n(z) \\ &= \frac{1}{N} \sum_{j=1}^{N} \frac{-1}{z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)} d_j, \end{aligned} \tag{3.5}$$

where

$$d_j = q_j^* T_n^{1/2}(B_{(j)} - zI)^{-1}(\underline{m}_n(z)T_n + I)^{-1}T_n^{1/2}q_j \\ - \frac{1}{n}\operatorname{tr}(\underline{m}_n(z)T_n + I)^{-1}T_n(B_n - zI)^{-1}.$$

The next task is to prove for $v = v_n \geq N^{-1/17}$ and for any subsets $S_n \subset [0, \infty)$ containing at most n elements the almost sure convergence of

$$\max_{x\in S_n} \frac{|w_n(z)|}{v_n^5}$$

to 0. Let

$$\underline{m}_{(j)}(z) = -\frac{(1-c_n)}{z} + c_n m_{F^{B_{(j)}}}(z).$$

From Lemma 2.10 we have

$$\max_{j\leq N}\left|\underline{m}_n(z) - \underline{m}_{(j)}(z)\right| \leq \frac{1}{Nv}. \tag{3.6}$$

Moreover, it is easy to verify that $\underline{m}_{(j)}(z)$ is the Stieltjes transform of a p.d.f., so that $|\underline{m}_{(j)}(z)| \leq v^{-1}$.

Write for each $j \leq N$, $d_j = d_j^1 + d_j^2 + d_j^3 + d_j^4$, where

$$\begin{aligned} d_j^1 &= q_j^* T_n^{1/2}(B_{(j)} - zI)^{-1}(\underline{m}_n(z)T_n + I)^{-1}T_n^{1/2}q_j \\ &\quad - q_j^* T_n^{1/2}(B_{(j)} - zI)^{-1}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n^{1/2}q_j, \\ d_j^2 &= q_j^* T_n^{1/2}(B_{(j)} - zI)^{-1}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n^{1/2}q_j \\ &\quad - \frac{1}{n}\operatorname{tr}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n(B_{(j)} - zI)^{-1}, \\ d_j^3 &= \frac{1}{n}\operatorname{tr}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n(B_{(j)} - zI)^{-1} \\ &\quad - \frac{1}{n}\operatorname{tr}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n(B_n - zI)^{-1} \end{aligned}$$

and

$$\begin{aligned} d_j^4 &= \frac{1}{n}\operatorname{tr}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n(B_n - zI)^{-1} \\ &\quad - \frac{1}{n}\operatorname{tr}(\underline{m}_n(z)T_n + I)^{-1}T_n(B_n - zI)^{-1}. \end{aligned}$$

In view of (3.4), it is sufficient to show the a.s. convergence of

$$\max_{j\leq N,\, x\in S_n} \frac{|d_j^i|}{v^6} \tag{3.7}$$

to 0 for $i = 1, 2, 3, 4$.

Using $\|(A-zI)^{-1}\| \le 1/v$ for any Hermitian matrix A, we get from Lemma 2.11(c) and (3.6)

$$|d_j^1| \le 16 \frac{\|X_{\bullet j}\|^2}{n} \frac{1}{Nv^4}.$$

Using (3.2), it follows that, for any $\varepsilon > 0$, $p \ge 2$ and all n sufficiently large,

$$P\left(\max_{j\le N,\, x\in S_n} \frac{|d_j^1|}{v^6} > \varepsilon\right) \le nP\left(\max_{j\le N}\left|\frac{\|X_{\bullet j}\|^2}{n} - 1\right| \frac{16}{Nv^{10}} > \frac{\varepsilon}{2}\right)$$
$$\le K_p \frac{nN}{(Nv^{10})^p} \varepsilon^{-p} n^{-p/2},$$

so (3.7) $\to$ 0 a.s. when $i = 1$ and for any $v_n \in (N^{-1/10}, 1]$.

Using Lemma 2.10 and Lemma 2.11(a), we find

$$v^{-6}|d_j^3| \le \frac{1}{nv^8},$$

so that (3.7) $\to$ 0 a.s. for $i = 3$ and for any $v_n = N^{-\delta}$ with $\delta \in [0, 1/8)$.

We get from Lemma 2.11(b) and (3.6)

$$v^{-6}|d_j^4| \le 16 \frac{1}{Nv^{10}},$$

so that (3.7) $\to$ 0 a.s. for $i = 4$, and for any $\delta \in [0, 1/10)$.

Using (3.2), we find, for any $p \ge 2$,

$$\begin{aligned}
E|v^{-6}d_j^2|^p &\le K_p \frac{1}{v^{6p}n^p}\Big(\operatorname{tr} T_n^{1/2}(B_{(j)} - zI)^{-1}(\underline{m}_{(j)}(z)T_n + I)^{-1} \\
&\qquad\qquad \times T_n(\overline{\underline{m}}_{(j)}(z)T_n + I)^{-1}(B_{(j)} - \bar zI)^{-1}T_n^{1/2}\Big)^{p/2} \\
&= K_p \frac{1}{v^{6p}n^p}\Big(\operatorname{tr}(\underline{m}_{(j)}(z)T_n + I)^{-1}T_n(\overline{\underline{m}}_{(j)}(z)T_n + I)^{-1} \\
&\qquad\qquad \times (B_{(j)} - \bar zI)^{-1}T_n(B_{(j)} - zI)^{-1}\Big)^{p/2} \\
&\qquad\qquad\qquad\qquad \text{[using Lemma 2.11(a)]} \\
&\le K_p \frac{1}{v^{6p}n^p}\frac{1}{v^{2\cdot p/2}}\big(\operatorname{tr}(B_{(j)} - \bar zI)^{-1}T_n(B_{(j)} - zI)^{-1}\big)^{p/2} \\
&= K_p \frac{1}{(nv^7)^p}\big(\operatorname{tr} T_n(B_{(j)} - zI)^{-1}(B_{(j)} - \bar zI)^{-1}\big)^{p/2} \\
&\le K_p \frac{1}{(nv^7)^p}\left(\frac{n}{v^2}\right)^{p/2} \\
&= K_p \frac{1}{(n^{1/2}v^8)^p}.
\end{aligned}$$

We then have, for any $\varepsilon > 0$ and $p \geq 2$,

$$P\Big(\max_{j\leq N,\, x\in S_n} |v^{-6} d_j^2| > \varepsilon\Big) \leq K_p \frac{1}{\varepsilon^p} \frac{nN}{(n^{1/2} v^8)^p}.$$

Thus, $\max_{x\in S_n} |w_n(z)| v^{-5} \to 0$ a.s. for any nonnegative $\delta \leq 1/17$ since we have shown for any positive ℓ, we have, for all p sufficiently large and for all $\varepsilon > 0$,

$$P\Big(\max_{x\in S_n} |w_n(z)| v_n^{-5} > \varepsilon\Big) \leq K_p \varepsilon^{-p} n^{-\ell}.$$

Moreover, for the sequence $\{\mu_n\}$ with $\mu_n = N^{1/68}$, we have, for $v_n = N^{-\delta}$ with any $\delta \leq 1/17$,

$$P\Big(\mu_n \max_{x\in S_n} |w_n(z)| v_n^{-5} > \varepsilon\Big) \leq K_p \varepsilon^{-p} n^{-\ell}. \tag{3.8}$$

We now rewrite w_n totally in terms of $\underline{m}_n$. With $H_n \equiv F^T$ and using the identity

$$\underline{m}_n(z) = -\frac{(1-c)}{z} + c m_n(z),$$

we have

$$\begin{aligned} w_n &= \frac{1}{c}\bigg(-\frac{c}{z}\int \frac{dH_n(t)}{1+t\underline{m}_n} - \underline{m}_n - \frac{(1-c)}{z}\bigg) \\ &= \frac{\underline{m}_n}{cz}\bigg(-\frac{c}{\underline{m}_n}\int \frac{dH_n(t)}{1+t\underline{m}_n} - z - \frac{(1-c)}{\underline{m}_n}\bigg) \\ &= \frac{\underline{m}_n}{cz}\bigg(-z - \frac{1}{\underline{m}_n} + c\int \frac{t\,dH_n(t)}{1+t\underline{m}_n}\bigg). \end{aligned}$$

Let

$$\omega = -z - \frac{1}{\underline{m}_n} + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n}.$$

Then $\omega = w_n z c_n / \underline{m}_n$.

Returning now to F^{c_n, H_n} and $F^{c, H}$, let $\underline{m}_n^0 = m_{F^{c_n, H_n}}$ and $\underline{m}^0 = m_{F^{c,H}}$. Then $\underline{m}^0$ solves (1.1), its inverse is given by (1.2),

$$\underline{m}_n^0 = \frac{1}{-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n^0}}, \tag{3.9}$$

and the inverse of $\underline{m}_n^0$, denoted z_n^0, is given by

$$z_n^0(\underline{m}) = -\frac{1}{\underline{m}} + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}}. \tag{3.10}$$

From (3.10) and the inversion formula for Stieltjes transforms, it is obvious that $F^{c_n, H_n} \to_{\mathscr{D}} F^{c, H}$ as $n \to \infty$. Therefore, from assumption (f), an $\underline{\epsilon} > 0$ exists for which $[a - 2\underline{\epsilon}, b + 2\underline{\epsilon}]$ also satisfies (f). This interval will stay uniformly bounded away from the boundary of the support of F^{c_n, H_n} for all large n, so that for these n both $\sup_{x\in[a-2\underline{\epsilon},\, b+2\underline{\epsilon}]}(d/dx)\underline{m}_n^0(x)$ is bounded and $-1/\underline{m}_n^0(x)$ for $x \in [a-2\underline{\epsilon}, b+2\underline{\epsilon}]$ stays uniformly away from the support of H_n. Therefore, for all n sufficiently large,

$$\sup_{x\in[a-2\underline{\epsilon},\, b+2\underline{\epsilon}]}\left(\frac{d}{dx}\underline{m}_n^0(x)\right)\int\frac{t^2\,dH_n(t)}{(1+t\underline{m}_n^0(x))^2} \le K. \tag{3.11}$$

Let $a' = a-\underline{\epsilon}$, $b' = b+\underline{\epsilon}$. On either $(-\infty, a']$ or $[b', \infty)$, each collection of functions in λ, $\{(\lambda-x)^{-1}\colon x\in[a,b]\}$, $\{(\lambda-x)^{-2}\colon x\in[a,b]\}$, form a uniformly bounded, equicontinuous family. It is straightforward then to show

$$\lim_{n\to\infty}\sup_{x\in[a,\,b]}|\underline{m}_n^0(x)-\underline{m}^0(x)| = 0 \tag{3.12}$$

and

$$\lim_{n\to\infty}\sup_{x\in[a,b]}\left|\frac{d}{dx}\underline{m}_n^0(x)-\frac{d}{dx}\underline{m}^0(x)\right| = 0 \tag{3.13}$$

[see, e.g., Billingsley (1968), Problem 8, page 17]. Since, for all $x\in[a,b]$, $\lambda\in[a',b']^c$ and positive v,

$$\left|\frac{1}{\lambda-(x+iv)}-\frac{1}{\lambda-x}\right| < \frac{v}{\underline{\epsilon}^2},$$

we have, for any sequence of positive v_n converging to 0,

$$\lim_{n\to\infty}\sup_{x\in[a,\,b]}\left|\underline{m}_n^0(x+iv_n)-\underline{m}_n^0(x)\right| = 0. \tag{3.14}$$

Similarly,

$$\lim_{n\to\infty}\sup_{x\in[a,\,b]}\left|\frac{\operatorname{Im}\underline{m}_n^0(x+iv_n)}{v_n}-\frac{d}{dx}\underline{m}_n^0(x)\right| = 0. \tag{3.15}$$

Expressions (3.11), (3.12), (3.14) and (3.15) will be needed in the latter part of Section 5.

Let $\underline{m}_2^0 = \operatorname{Im}\underline{m}_n^0$. We then have from (3.9)

$$\underline{m}_2^0 = \frac{v_n+\underline{m}_2^0 c_n\int\dfrac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}}{\left|-z+c_n\int\dfrac{t\,dH_n(t)}{1+t\underline{m}_n^0}\right|^2}. \tag{3.16}$$

For any real x, by Lemma 2.11(a),

$$\begin{aligned}\underline{m}_2^0 c_n\int\frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2} &= c_n\operatorname{Im}\left(\int\frac{t\,dH_n(t)}{1+t\underline{m}_n^0}\right)\\ &\le c_n\|T(I+T\underline{m}_n^0)^{-1}\| \le \frac{4c_n}{v_n}.\end{aligned}$$

It follows that

$$\left(\frac{\underline{m}_2^0 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}}{v_n + \underline{m}_2^0 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}}\right)^{1/2} < 1 - K v_n^2 \tag{3.17}$$

for some positive constant K.

Let $\underline{m}_n = \underline{m}_1 + i\underline{m}_2$, where $\underline{m}_1 = \operatorname{Re}\underline{m}_n$, $\underline{m}_2 = \operatorname{Im}\underline{m}_n$. We have $\underline{m}_n$ satisfying

$$\underline{m}_n = \frac{1}{-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n} - \omega} \tag{3.18}$$

and

$$\underline{m}_2 = \frac{v_n + \underline{m}_2 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n|^2} + \operatorname{Im}\omega}{\left|-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n} - \omega\right|^2}. \tag{3.19}$$

From (3.9) and (3.18), we get

$$\underline{m}_n - \underline{m}_n^0 = \frac{(\underline{m}_n - \underline{m}_n^0) c_n \int \frac{t^2\,dH_n(t)}{(1+t\underline{m}_n)(1+t\underline{m}_n^0)}}{\left(-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n} - \omega\right)\left(-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n^0}\right)} + \underline{m}_n \underline{m}_n^0 \omega. \tag{3.20}$$

From Cauchy-Schwarz, (3.16), (3.17) and (3.19), we get, when $|\operatorname{Im}\omega/v_n| < 1$,

$$\begin{aligned}
&\left| c_n \frac{\int \frac{t^2\,dH_n(t)}{(1+t\underline{m}_n)(1+t\underline{m}_n^0)}}{\left(-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n} - \omega\right)\left(-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n^0}\right)} \right| \\
&\quad \le \left(c_n \frac{\int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n|^2}}{\left|-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n} - \omega\right|^2} \right)^{1/2} \left(c_n \frac{\int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}}{\left|-z + c_n \int \frac{t\,dH_n(t)}{1+t\underline{m}_n^0}\right|^2} \right)^{1/2} \\
&\quad = \left(\frac{\underline{m}_2 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n|^2}}{v_n + \underline{m}_2 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n|^2} + \operatorname{Im}\omega} \right)^{1/2} \left(\frac{\underline{m}_2^0 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}}{v_n + \underline{m}_2^0 c_n \int \frac{t^2\,dH_n(t)}{|1+t\underline{m}_n^0|^2}} \right)^{1/2}
\end{aligned} \tag{3.21}$$

$$\le \left(\frac{\underline{m}_2^0 c_n \int \frac{t^2\, dH_n(t)}{|1+t\underline{m}_n^0|^2}}{v_n + \underline{m}_2^0 c_n \int \frac{t^2\, dH_n(t)}{|1+t\underline{m}_n^0|^2}} \right)^{1/2}$$

$$\le 1 - \underline{K} v_n^2.$$

We claim that on the set $\{\lambda_{\max} \le K_1\}$, where $K_1 > (1+\sqrt{c})^2$, for all n sufficiently large, $|\underline{m}_n| \ge \frac{1}{2}\mu_n^{-1} v_n$ whenever $|x| \le \mu_n v_n^{-1}$. Indeed, when $x \le -v_n$ or $x \ge \lambda_{\max} + v_n$,

$$|\underline{m}_n| \ge |\mathrm{Re}\, \underline{m}_n| \ge \frac{K_1 + \mu_n v_n^{-1}}{(K_1 + \mu_n v_n^{-1})^2 + v_n^2} \ge \frac{1}{2\mu_n v_n^{-1}}$$

for n large. When $-v_n < x < \lambda_{\max} + v_n$,

$$|\underline{m}_n| \ge |\mathrm{Im}\, \underline{m}_n| \ge \frac{v_n}{(K_1 + v_n)^2 + v_n^2} \ge \mu_n^{-1} v_n^{-1}$$

for n large. Thus, the claim is proven.

Therefore, when $|x| \le \mu_n v_n^{-1}$, on the set $\{|w_n| \le v_n^4\} \cap \{\lambda_{\max} \le K_1\}$ we have for n large $|z| \le 2\mu_n v_n^{-1}$ and

$$|\mathrm{Im}(\omega)| \le |c_n z w_n / \underline{m}_n| \le K \mu_n^2 v_n^{-2} |w_n| < v_n.$$

Therefore, by (3.20) and (3.21), we have

$$\begin{aligned} |\underline{m}_n - \underline{m}_n^0| &\le \underline{K}^{-1} v_n^{-2} |\underline{m}_n \underline{m}_n^0 \omega| \\ &= \underline{K}^{-1} v_n^{-2} |c_n z \underline{m}_n^0 w_n| \le K' v_n^{-4} \mu_n |w_n|. \end{aligned}$$

It is easy to verify that for n large, when either $|x| > \mu_n v_n^{-1}$, $|w_n| > v_n^4$ or $\lambda_{\max} > K_1$,

$$|\underline{m}_n - \underline{m}_n^0| \le 3\mu_n^{-1} v_n + 2 v_n^{-1} (I_{[|w_n| > v_n^4]} + I_{[\lambda_{\max} > K_1]}).$$

Therefore, for n large, we have

$$\begin{aligned} &\max_{x \in S_n} v_n^{-1} |\underline{m}_n(z) - \underline{m}_n^0| \\ &\qquad \le K' \mu_n \max_{x \in S_n} |w_n| v_n^{-5} + 3\mu_n^{-1} + 2 v_n^{-2} \max_{x \in S_n} \big(I_{[|w_n| > v_n^4]} + I_{[\lambda_{\max} > K_1]} \big). \end{aligned}$$

Therefore, from (3.1) and (3.8) we find, for any positive ε and ℓ,

$$P\Big(v_n^{-1} \max_{x \in S_n} |\underline{m}_n(z) - \underline{m}_n^0| > \varepsilon \Big) \le K_p \varepsilon^{-p} n^{-\ell} \tag{3.22}$$

for all p sufficiently large, whenever $\delta \le 1/17$.

We now assume the n elements of S_n to be equally spaced between $-\sqrt{n}$ and $\sqrt{n}$. Since, for $|x_1 - x_2| \le 2n^{-1/2}$,

$$|\underline{m}_n(x_1 + i v_n) - \underline{m}_n(x_2 + i v_n)| \le 2n^{-1/2} v_n^{-2},$$

$$|\underline{m}_n^0(x_1 + i v_n) - \underline{m}_n^0(x_2 + i v_n)| \le 2n^{-1/2} v_n^{-2},$$

and when $|x| \geq \sqrt{n}$, for n large,

$$|\underline{m}_n(x+iv_n)| \leq 2n^{-1/2} + v_n^{-1} I_{[\lambda_{\max} > K_1]},$$
$$|\underline{m}_n^0(x+iv_n)| \leq 2n^{-1/2},$$

we conclude from (3.22) and (3.1), that, for any positive ε and ℓ,

$$P\Big(v_n^{-1} \sup_{x\in\mathbb{R}} |\underline{m}_n(x+iv_n) - \underline{m}_n^0(x+iv_n)| > \varepsilon\Big) \leq K_p \varepsilon^{-p} n^{-\ell} \tag{3.23}$$

for all sufficiently large p, whenever $\delta \leq 1/17$.

Let $E_0(\cdot)$ denote expectation and $E_k(\cdot)$ denote conditional expectation with respect to the σ-field generated by $r_1, \ldots, r_k$. Let $\ell, \ell' > 0$ be arbitrary. Choose $\underline{\ell} > \ell$, let p be suitably large so that (3.23) holds with ℓ replaced by $\underline{\ell}$ and set $r = \ell p/(\underline{\ell}\ell')$. Since $E_k(v_n^{-\ell'} \sup_{x\in\mathbb{R}} |\underline{m}_n(x+iv_n) - \underline{m}_n^0(x+iv_n)|^{\ell'})$, $k = 0, \ldots, N$, forms a martingale, it follows from Jensen's inequality, Lemmas 2.5 and 2.6 and (3.23) that, for any positive ε,

$$\begin{aligned}
P\Big(\max_{k\leq N} E_k\Big(v_n^{-\ell'} \sup_{x\in\mathbb{R}}\big|\underline{m}_n(x+iv_n) - \underline{m}_n^0(x+iv_n)\big|^{\ell'}\Big) > \varepsilon\Big)&\\
\leq \varepsilon^{-r} E\Big(v_n^{-r\ell'} \sup_{x\in\mathbb{R}}\big|\underline{m}_n(x+iv_n) - \underline{m}_n^0(x+iv_n)\big|^{r\ell'}\Big)&\\
\leq \varepsilon^{-r} K_p^{\ell/\underline{\ell}} \frac{\underline{\ell}}{\underline{\ell}-\ell} n^{-\ell},&
\end{aligned}$$

whenever $\delta \leq 1/17$. In particular, we have, for $\delta \leq 1/17$,

$$\lim_{n\to\infty} \max_{k\leq N} \frac{E_k(\sup_{x\in\mathbb{R}} |\underline{m}_n(x+iv_n) - \underline{m}_n^0(x+iv_n)|^2)}{v_n^2} = 0 \quad \text{with probability 1.} \tag{3.24}$$

Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ be the eigenvalues of $\underline{B}_n$ and write

$$\underline{m}_j = \underline{m}_j^{\text{out}} + \underline{m}_j^{\text{in}}, \qquad j = 1, 2,$$

where

$$\underline{m}_2^{\text{out}}(x+iv) = \frac{1}{N} \sum_{\lambda_j \in [a', b']} \frac{v}{(x-\lambda_j)^2 + v^2},$$
$$\underline{m}_1^{\text{out}}(x+iv) = \frac{1}{N} \sum_{\lambda_j \in [a', b']} \frac{x-\lambda_j}{(x-\lambda_j)^2 + v^2}.$$

Define the sequence $\{G_m\}_{m=1}^{\infty}$ of functions on $\mathbb{R}^2$ by

$$G_{\sum_{j=1}^{n-1}(N(j)+1)+k}(x_1, x_2) = E_k F^{\underline{B}_n}(x_1) F^{\underline{B}_n}(x_2)$$

for $k = 0, 1, \ldots, N(n)$. Clearly each G_m is a probability distribution function on $\mathbb{R}^2$, and when $m = \sum_{j=1}^{n-1}(N(j)+1)+k$, the two-dimensional Stieltjes transform, $m_m^{(G)}(x_1+iv_1, x_2+iv_2)$ of G_m is $E_k \underline{m}_n(x_1+iv_1)\underline{m}_n(x_2+iv_2)$. Obviously, when $\delta = 0$, (3.22) implies that, with probability 1, $\sup_{x_1, x_2\in\mathbb{R}} |m_m^{(G)}(x_1+$

$iv_1, x_2 + iv_2) - \underline{m}_n^0(x_1 + iv_1)\underline{m}_n^0(x_2 + iv_2)| \to 0$ as $m \to \infty$ for countably many (v_1, v_2) forming a dense subset of an open set in the first quadrant (bounded uniformly away from the two axes). We conclude that, with probability 1, $G_m(x_1, x_2)$ converges weakly to $F^{c,H}(x_1)F^{c,H}(x_2)$.

Since the integrands of

$$\int_{[a',b']^c \times [a',b']^c} \frac{dE_k F^{\underline{B}_n}(x_1) F^{\underline{B}_n}(x_2)}{((x-x_1)^2 + v^2)((x-x_2)^2 + v^2)}$$

and

$$\int_{[a',b']^c} \frac{dE_k F^{\underline{B}_n}(x_1)}{(x-x_1)^2 + v^2}$$

on their respective domains are uniformly bounded and equicontinuous for $x \in [a, b]$, it follows as in (3.13) that

$$\max_k \sup_{x\in[a,b]} E_k \left| \frac{\underline{m}_2^{\text{in}}(x + iv_n)}{v} - \frac{d}{dx}\underline{m}^0(x) \right|^2 \to 0 \quad \text{a.s.} \tag{3.25}$$

for any $v = v_n \to 0$.

Therefore, from (3.24) and (3.25) we have

$$\max_{k\le N,\ x\in[a,b]} v_n^{-2} E_k\big(\underline{m}_2^{\text{out}}(x + iv)\big)^2 \to 0 \quad \text{a.s.} \tag{3.26}$$

From (3.26) we can infer a bound on the number of eigenvalues in $[a, b]$. Notice $NF^{\underline{B}_n}(A)$ is the number of eigenvalues of $\underline{B}_n$ in the set A. Let e_n denote the left-hand side of (3.26). For any $x \in [a, b]$,

$$\begin{aligned} e_n &\ge \frac{1}{N^2} \max_{k\le N} E_k \Bigg(\sum_{\lambda_j \in [a,b]\cap[x-v_n,\, x+v_n]} \frac{1}{(x-\lambda_j)^2 + v_n^2} \Bigg)^2 \\ &\ge \max_{k\le N} \frac{N^2 E_k (F^{\underline{B}_n}\{[a,b]\cap[x-v_n, x+v_n]\})^2}{4v_n^4 N^2}, \end{aligned}$$

and since the number of intervals of length $2v_n$ needed to cover $[a, b]$ is $\lceil (b - a)/2v_n \rceil$, we find $E_k(F^{\underline{B}_n}\{[a,b]\})^2 \le (b-a)^2 v_n^2 e_n$. Therefore,

$$\max_{k\le N} E_k\big(F^{\underline{B}_n}\{[a,b]\}\big)^2 = o_{\text{a.s.}}(v_n^2) = o_{\text{a.s.}}(N^{-2/17}),$$

which implies

$$\max_{k\le N} E_k\big(F^{\underline{B}_n}\{[a,b]\}\big) = o_{\text{a.s.}}(v_n) = o_{\text{a.s.}}(N^{-1/17}).$$

The above arguments apply to $[a', b']$ as well, so we also have

$$\max_{k\le N} E_k\big(F^{\underline{B}_n}\{[a',b']\}\big)^2 = o_{\text{a.s.}}(v_n^2) = o_{\text{a.s.}}(N^{-2/17}) \tag{3.27}$$

and

$$\max_{k\le N} E_k\big(F^{\underline{B}_n}\{[a',b']\}\big) = o_{\text{a.s.}}(v_n) = o_{\text{a.s.}}(N^{-1/17}). \tag{3.28}$$

4. Convergence of $m_n - Em_n$. We now restrict $\delta = 1/68$, that is, $v = v_n = N^{-1/68}$.

Our goal is to show that

$$\sup_{x\in[a,b]} Nv_n|m_n - Em_n| \to 0 \quad \text{a.s. as } n \to \infty. \tag{4.1}$$

Write $D = B_n - zI$, $D_j = D - r_j r_j^*$ and $D_{j\underline{j}} = D - (r_j r_j^* + r_{\underline{j}} r_{\underline{j}}^*)$, $j \neq \underline{j}$. Then $m_n = (1/n)\,\mathrm{tr}(D^{-1})$. Let us also denote

$$\alpha_j = r_j^* D_j^{-2} r_j - N^{-1}\,\mathrm{tr}(D_j^{-2}T_n), \qquad a_j = N^{-1}\,\mathrm{tr}(D_j^{-2}T_n),$$

$$\beta_j = \frac{1}{1 + r_j^* D_j^{-1} r_j}, \qquad b_n = \frac{1}{1 + N^{-1}E\,\mathrm{tr}(T_n D_1^{-1})},$$

$$\gamma_j = r_j^* D_j^{-1} r_j - N^{-1}E(\mathrm{tr}(D_j^{-1}T_n)), \qquad \hat{\gamma}_j = r_j^* D_j^{-1} r_j - N^{-1}\,\mathrm{tr}(D_j^{-1}T_n).$$

We first derive bounds on moments of γ_j and $\hat{\gamma}_j$. Using (3.2), we find, for all $p \geq 2$,

$$E|\hat{\gamma}_j|^p \leq K_p N^{-p} E(\mathrm{tr}\, T_n^{1/2} D_j^{-1} T_n \bar{D}_j^{-1} T_n^{1/2})^{p/2} \leq K_p N^{-p/2} v_n^{-p}. \tag{4.2}$$

Using Lemmas 2.2 and 2.10, we have, for $p \geq 2$,

$$\begin{aligned} E|\gamma_j - \hat{\gamma}_j|^p &= E|\gamma_1 - \hat{\gamma}_1|^p \\ &= E\left|\frac{1}{N}\sum_{j=2}^{N} E_j\,\mathrm{tr}\, T_n D_1^{-1} - E_{j-1}\,\mathrm{tr}\, T_n D_1^{-1}\right|^p \\ &= E\left|\frac{1}{N}\sum_{j=2}^{N} E_j\,\mathrm{tr}\, T_n(D_1^{-1} - D_{1j}^{-1}) - E_{j-1}\,\mathrm{tr}\, T_n(D_1^{-1} - D_{1j}^{-1})\right|^p \\ &= E\left|\frac{1}{N}\sum_{j=2}^{N}(E_j - E_{j-1})\frac{r_j^* D_{1j}^{-1} T_n D_{1j}^{-1} r_j}{1 + r_j^* D_{1j}^{-1} r_j}\right|^p \\ &\leq K_p \frac{1}{N^p} E\left(\sum_{j=2}^{N}\left|(E_j - E_{j-1})\frac{r_j^* D_{1j}^{-1} T_n D_{1j}^{-1} r_j}{1 + r_j^* D_{1j}^{-1} r_j}\right|^2\right)^{p/2} \\ &\leq K_p N^{-p/2} v_n^{-p}. \end{aligned}$$

Therefore,

$$E|\gamma_j|^p \leq K_p N^{-p/2} v_n^{-p}. \tag{4.3}$$

We next prove that b_n is bounded for all n. We have b_n and β_1 both bounded in absolute value by $|z|/v_n$ [see (3.4)]. From the equation relating $\underline{m}_n$ to the β_j's [above (3.4)], we have $E\beta_1 = -zE\underline{m}_n$. Using (3.24), we get

$$\sup_{x\in[a,b]} \left|E(\underline{m}_n(z)) - \underline{m}_n^0(z)\right| = o(v_n).$$

Since $\underline{m}_n^0$ is bounded for all n, $x \in [a, b]$ and v, we have $\sup_{x\in[a,b]} |E\beta_1| \leq K$.

Since $b_n = \beta_1 + \beta_1 b_n \gamma_1$, we get

$$\sup_{x\in[a,b]} |b_n| = \sup_{x\in[a,b]} |E\beta_1 + E\beta_1 b_n \gamma_1| \le K + K_2^{1/2} v_n^{-3} N^{-1/2} \le K.$$

Since $|m_n(x_1 + iv_n) - m_n(x_2 + iv_n)| \le |x_1 - x_2| v_n^{-2}$, we see that (4.1) will follow from

$$\max_{x\in S_n} N v_n |m_n - Em_n| \to 0 \quad \text{a.s.},$$

where S_n now contains n^2 elements, equally spaced in $[a, b]$.

We write

$$\begin{aligned}
m_n - Em_n &= \frac{1}{n}\sum_{j=1}^{N} E_j \operatorname{tr} D^{-1} - E_{j-1} \operatorname{tr} D^{-1} \\
&= \frac{1}{n}\sum_{j=1}^{N} [E_j - E_{j-1}]\left(\frac{r_j^* D_j^{-2} r_j}{1 + r_j^* D_j^{-1} r_j}\right) \\
&= \frac{1}{n}\sum_{j=1}^{N} (E_j - E_{j-1}) \frac{r_j^* D_j^{-2} r_j}{1 + N^{-1} E \operatorname{tr} T_n D_j^{-1}} \\
&\quad + \frac{1}{n}\sum_{j=1}^{N} (E_j - E_{j-1}) \frac{r_j^* D_j^{-2} r_j (N^{-1} E \operatorname{tr} T_n D_j^{-1} - r_j^* D_j^{-1} r_j)}{(1 + N^{-1} E \operatorname{tr} T_n D_j^{-1})^2} \\
&\quad + \frac{1}{n}\sum_{j=1}^{N} (E_j - E_{j-1}) \frac{r_j^* D_j^{-2} r_j (N^{-1} E \operatorname{tr} T_n D_j^{-1} - r_j^* D_j^{-1} r_j)^2}{(1 + N^{-1} E \operatorname{tr} T_n D_j^{-1})^2 (1 + r_j^* D_j^{-1} r_j)} \\
&= \frac{b_n}{n}\sum_{j=1}^{N} E_j \alpha_j - \frac{b_n^2}{n}\sum_{j=1}^{N} E_j a_j \hat{\gamma}_j \\
&\quad - \frac{b_n^2}{n}\sum_{j=1}^{N} (E_j - E_{j-1})(\alpha_j \gamma_j - r_j^* D_j^{-2} r_j \beta_j \gamma_j^2) \\
&\equiv W_1 - W_2 - W_3.
\end{aligned}$$

Let F_{nj} be the spectral distribution of the matrix $\sum_{k\ne j} r_k r_k^*$. From Lemma 2.12 and (3.27), we get

$$\max_j E_j (F_{nj}([a', b']))^2 = o(N^{-2/17}) = o(v_n^8) \quad \text{a.s.} \tag{4.4}$$

Define

$$\mathscr{B}_j = I_{[E_{j-1} F_{nj}([a',b']) \le v_n^4] \cap [E_{j-1}(F_{nj}([a',b']))^2 \le v_n^8]}.$$

Then $\mathscr{B}_j = I_{[E_j F_{nj}([a',b']) \le v_n^4] \cap [E_j (F_{nj}([a',b']))^2 \le v_n^8]}$ a.s. and we have

$$P\left(\bigcup_{j=1}^{N} [\mathscr{B}_j = 0] \text{ i.o.}\right) = 0.$$

Therefore, we have, for any $\varepsilon > 0$,

$$
\begin{aligned}
&P\Big(\max_{x\in S_n}|Nv_nW_1| > \varepsilon \text{ i.o.}\Big)\\
&\quad \le P\bigg(\bigg(\bigg[\max_{x\in S_n}\bigg|v_n\sum_{j=1}^N E_j(\alpha_j)\bigg| > \underline{\varepsilon}\bigg]\bigcap_{j=1}^N[\mathscr{B}_j = 1]\bigg)\cup\bigg(\bigcup_{j=1}^N[\mathscr{B}_j = 0]\bigg)\text{ i.o.}\bigg)\\
&\quad \le P\bigg(\max_{x\in S_n}\bigg|v_n\sum_{j=1}^N E_j(\alpha_j)\mathscr{B}_j\bigg| > \underline{\varepsilon}\text{ i.o.}\bigg),
\end{aligned}
$$

where $\underline{\varepsilon} = \inf_n n\varepsilon/(Nb_n) > 0$ since b_n is bounded. Note that, for each $x \in \mathbb{R}$, $\{E_j(\alpha_j)\mathscr{B}_j\}$ forms a martingale difference sequence.

By Lemma 2.1 and (3.2), we have, for each $x \in [a, b]$ and $p \ge 2$,

$$
\begin{aligned}
&E\bigg|v_n\sum_{j=1}^N E_j(\alpha_j)\mathscr{B}_j\bigg|^p\\
&\quad \le K_p\bigg(E\bigg(\sum_{j=1}^N E_{j-1}|v_nE_j(\alpha_j)\mathscr{B}_j|^2\bigg)^{p/2} + \sum_{j=1}^N E|v_nE_j(\alpha_j)\mathscr{B}_j|^p\bigg)\\
&\quad \le K_p\bigg(E\bigg(\sum_{j=1}^N E_{j-1}v_n^2N^{-2}\mathscr{B}_j\operatorname{tr}\Big(T_n^{1/2}D_j^{-2}T_n\bar{D}_j^{-2}T_n^{1/2}\Big)\bigg)^{p/2} + Nv_n^pE|\alpha_1|^p\bigg)\\
&\quad \le K_p\bigg(v_n^pN^{-p}E\bigg(\sum_{j=1}^N \mathscr{B}_jE_{j-1}\operatorname{tr}\Big(D_j^{-2}\bar{D}_j^{-2}\Big)\bigg)^{p/2}\\
&\qquad\qquad + Nv_n^pN^{-p}E\Big(\operatorname{tr}\Big(T_n^{1/2}D_1^{-2}T_n\bar{D}_1^{-2}T_n^{1/2}\Big)\Big)^{p/2}\bigg)\\
&\quad \le K_p\bigg(v_n^pN^{-p}E\bigg(\sum_{j=1}^N \mathscr{B}_jE_{j-1}\operatorname{tr}\Big(D_j^{-2}\bar{D}_j^{-2}\Big)\bigg)^{p/2} + v_n^{-p}N^{1-p/2}\bigg).
\end{aligned}
$$

Let λ_{kj} denote the kth smallest eigenvalue of $\sum_{k\ne j} r_kr_k^*$. We have

$$
\begin{aligned}
\sum_{j=1}^N \mathscr{B}_jE_{j-1}\operatorname{tr}D_j^{-2}\bar{D}_j^{-2} &= \sum_{j=1}^N \mathscr{B}_jE_{j-1}\bigg[\sum_{\lambda_{kj}\notin[a'b']}\frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2}\\
&\qquad\qquad + \sum_{\lambda_{kj}\in[a'b']}\frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2}\bigg]\\
&\le \sum_{j=1}^N (n\underline{\epsilon}^{-4} + \mathscr{B}_jv_n^{-4}E_{j-1}nF_{nj}([a', b'])) \le KN^2.
\end{aligned}
$$

Therefore,

$$
P\bigg(\max_{x\in S_n}\bigg|v_n\sum_{j=1}^N E_j(\alpha_j)\mathscr{B}_j\bigg| > \varepsilon\bigg) \le n^2K_{p,\varepsilon}N^{-p/68},
$$

which is summable when $p > 204$. Therefore, $\max_{x\in S_n}|W_1| = o(1/Nv_n)$ a.s.

Proving

$$\max_{x\in S_n} |W_2| = o(1/Nv_n) \quad \text{a.s.} \tag{4.5}$$

is handled the same way. We get, using Lemma 2.10, (3.2) and the fact that $|a_j| \le (n/N)v_n^{-2}$,

$$\begin{aligned}
&E\left|v_n \sum_{j=1}^N E_j(a_j\hat{\gamma}_j)\mathscr{B}_j\right|^p \\
&\quad \le K_p\left(E\left(\sum_{j=1}^N E_{j-1}|v_n E_j(a_j\hat{\gamma}_j)\mathscr{B}_j|^2\right)^{p/2} + \sum_{j=1}^N E|v_n E_j(a_j\hat{\gamma}_j)\mathscr{B}_j|^p\right) \\
&\quad \le K_p\left(v_n^p N^{-p} E\left(\sum_{j=1}^N \mathscr{B}_j E_{j-1}\left(|a_j|^2 \operatorname{tr} D_j^{-1}\bar{D}_j^{-1}\right)\right)^{p/2}\right. \\
&\qquad\qquad \left. + N^{1-p} v_n^{-p}\left(\operatorname{tr} D_j^{-1}\bar{D}_j^{-1}\right)^{p/2}\right) \\
&\quad \le K_p\left(v_n^p N^{-p} E\left(\sum_{j=1}^N \mathscr{B}_j E_{j-1}\left(|a_j|^2 \operatorname{tr} D_j^{-1}\bar{D}_j^{-1}\right)\right)^{p/2} + v_n^{-2p} N^{1-p/2}\right).
\end{aligned}$$

This time

$$\begin{aligned}
&\sum_{j=1}^N \mathscr{B}_j E_{j-1}\left(|a_j|^2 \operatorname{tr} D_j^{-1}\bar{D}_j^{-1}\right) \\
&\quad \le \sum_{j=1}^N \mathscr{B}_j E_{j-1} N^{-2} n \sum_k \frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2} \sum_k \frac{1}{(\lambda_{kj}-x)^2+v_n^2} \\
&\quad \le \sum_{j=1}^N \mathscr{B}_j N^{-2} n E_{j-1}(n\underline{\epsilon}^{-4} + v_n^{-4} n F_{nj}([a,b])) \\
&\qquad\quad \times (n\underline{\epsilon}^{-2} + v_n^{-2} n F_{nj}([a,b])) \le KN^2,
\end{aligned}$$

so that (4.5) also holds.

Using Lemmas 2.2 and 2.10 and (3.2) and (4.3), we get

$$\begin{aligned}
&E\left|v_n \sum_{j=1}^N (E_j - E_{j-1})(\alpha_j\gamma_j - r_j D_j^{-2} r_j \beta_j \gamma_j^2)\right|^p \\
&\quad \le K_p v_n^p N^{p/2}(E|\alpha_1\gamma_1|^p + v_n^{-p} E|\gamma_1|^{2p}) \\
&\quad \le K_p v_n^p N^{p/2}\left(N^{-p}\left(E\left(\operatorname{tr} D_j^{-2}\bar{D}_j^{-2}\right)^p\right)^{1/2} N^{-p/2} v_n^{-p} + v_n^{-3p} N^{-p}\right) \\
&\quad \le K_p v_n^p N^{p/2}(N^{-p} N^{p/2} v_n^{-2p} N^{-p/2} v_n^{-p} + v_n^{-3p} N^{-p}) = 2K_p N^{-p/2} v_n^{-2p}.
\end{aligned}$$

Thus, we get $\max_{x\in S_n} |W_3| = o(1/Nv_n)$ a.s. and, consequently, (4.1).

5. Convergence of expected value. Our next goal is to show that, for $v = N^{-1/68}$,

$$\sup_{x\in[a,b]} \left| E\underline{m}_n - \underline{m}_n^0 \right| = O(1/N). \tag{5.1}$$

We begin by deriving an identity similar to (3.5). Write $B_n - zI - (-zE\underline{m}_n(z)T_n - zI) = \sum_{j=1}^N r_j r_j^* - (-zE\underline{m}_n(z))T_n$. Taking first inverses and then expected value, we get

$$\begin{aligned}
&\left(-zE\underline{m}_n T_n - zI\right)^{-1} - E(B_n - zI)^{-1} \\
&\quad = \left(-zE\underline{m}_n T_n - zI\right)^{-1} E\left[\sum_{j=1}^N r_j r_j^* - (-zE\underline{m}_n(z))T_n (B_n - zI)^{-1} \right] \\
&\quad = -z^{-1} \sum_{j=1}^N E\beta_j \bigg[(E\underline{m}_n(z)T_n + I)^{-1} r_j r_j^* (B_{(j)} - zI)^{-1} \\
&\qquad\qquad - \frac{1}{N}(E\underline{m}_n(z)T_n + I)^{-1} T_n E(B_n - zI)^{-1} \bigg] \\
&\quad = -z^{-1} N E\beta_1 \bigg[(E\underline{m}_n(z)T_n + I)^{-1} r_1 r_1^* D_1^{-1} \\
&\qquad\qquad - \frac{1}{N}(E\underline{m}_n(z)T_n + I)^{-1} T_n E D^{-1} \bigg].
\end{aligned}$$

Taking the trace on both sides and dividing by $-N/z$, we get

$$\begin{aligned}
&c_n \int \frac{dH_n(t)}{1 + tE\underline{m}_n} + zc_n E(m_n(z)) \\
&\quad = E\beta_1 \left[r_1^* D_1^{-1} (E\underline{m}_n T_n + I)^{-1} r_1 - \frac{1}{N} E\,\mathrm{tr}(E\underline{m}_n T_n + I)^{-1} T_n D^{-1} \right].
\end{aligned} \tag{5.2}$$

We first show

$$\begin{aligned}
\sup_{x\in[a,b]} N^{-1} \Big| & E\,\mathrm{tr}(E\underline{m}_n T_n + I)^{-1} T_n D^{-1} \\
& - E\,\mathrm{tr}(E\underline{m}_n T_n + I)^{-1} T_n D_1^{-1} \Big| = O(N^{-1}).
\end{aligned} \tag{5.3}$$

From (4.4) we get

$$\sup_{x\in[a,b]} E\left(\mathrm{tr}\, D_1^{-1} \bar{D}_1^{-1}\right)^2 \le E(n\underline{\epsilon}^{-2} + v_n^{-2} n F_{n1}([a', b']))^2 \le KN^2 \tag{5.4}$$

and

$$\sup_{x\in[a,b]} E\,\mathrm{tr}\, D_1^{-2} \bar{D}_1^{-2} \le E(n\underline{\epsilon}^{-4} + v_n^{-4} n F_{n1}([a', b'])) \le KN. \tag{5.5}$$

Also, because of (3.24) and the fact that $-1/\underline{m}_n^0(z)$ stays uniformly away from the eigenvalues of T_n for all $x \in [a, b]$, we must have

$$\sup_{x\in[a,b]} \|(E\underline{m}_n(z)T_n + I)^{-1}\| \le K. \tag{5.6}$$

Therefore, from (3.3), (4.3), (5.4)–(5.6) and the fact that $\sup_{x\in[a,b]}|b_n|$ is bounded, we get

$$\begin{aligned}
&\text{Left-hand side of (5.3)}\\
&= N^{-1}\sup_{x\in[a,b]}\big|E\beta_1 r_1^* D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}r_1)\big|\\
&\le N^{-1}\sup_{x\in[a,b]}\big(|b_n|\cdot\big|Er_1^*D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}r_1\big|\\
&\qquad\qquad + E\big|\beta_1 b_n\gamma_1 r_1^*D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}r_1\big|\big)\\
&\le KN^{-1}\sup_{x\in[a,b]}\big(N^{-1}\big|E\,\mathrm{tr}\,T_n^{1/2}D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}T_n^{1/2}\big|\\
&\qquad\qquad + v_n^{-1}(E|\gamma_1|^2)^{1/2}\big(E\big|r_1^*D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}r_1\big|^2\big)^{1/2}\big)\\
&\le KN^{-1}\sup_{x\in[a,b]}\Big(N^{-1}E\,\mathrm{tr}\,D_1^{-1}\bar{D}_1^{-1}\\
&\qquad\qquad + v_n^{-1}N^{-1/2}v_n^{-1}N^{-1}\Big(E\,\mathrm{tr}\,D_1^{-2}\bar{D}_1^{-2}+E\Big(\mathrm{tr}\,D_1^{-1}\bar{D}_1^{-1}\Big)^2\Big)^{1/2}\Big)\\
&\le KN^{-1}.
\end{aligned}$$

Thus, (5.3) holds.

From (3.2), (5.4) and (5.6), we get

$$\begin{aligned}
&\sup_{x\in[a,b]} E\big|r_1^*D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}r_1 - N^{-1}\,\mathrm{tr}\,D_1^{-1}\big(E\underline{m}_nT_n+I\big)^{-1}T_n\big|^2\\
&\qquad \le KN^{-2}\sup_{x\in[a,b]} ED_1^{-1}\bar{D}_1^{-1}\\
&\qquad \le KN^{-1}.
\end{aligned}\tag{5.7}$$

Next we show

$$\begin{aligned}
&\sup_{x\in[a,b]} N^{-1}E\Big|\mathrm{tr}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}\\
&\qquad - E\,\mathrm{tr}\big(E\underline{m}_nT_n+I\big)^{-1}T_nD_1^{-1}\Big|^2 \le KN^{-1}.
\end{aligned}\tag{5.8}$$

Let

$$\beta_{1j} = \frac{1}{1+r_j^*D_{1j}^{-1}r_j}, \qquad b_{1n} = \frac{1}{1+N^{-1}E\,\mathrm{tr}(T_nD_{12}^{-1})}$$

and

$$\gamma_{1j} = r_j^*D_{1j}^{-1}r_j - N^{-1}E\big(\mathrm{tr}\big(D_{1j}^{-1}T_n\big)\big).$$

As in the previous section, both β_{1j} and b_{1n} are bounded in absolute value by $|z|/v_n$ and γ_{1j} satisfies the same bound as in (4.3). Moreover, if we let $X_{(1)}$ denote X without its first column, then one can easily derive

$$\begin{aligned}\frac{1}{N-1}\operatorname{tr}\left(\frac{1}{N-1}X_{(1)}^*T_nX_{(1)}-\left(\frac{N}{N-1}z\right)I\right)^{-1} &= \frac{1}{N}\operatorname{tr}\left(\frac{1}{N}X_{(1)}^*T_nX_{(1)}-zI\right)^{-1}\\ &= -\frac{1}{z(N-1)}\sum_{j=2}^{N}\beta_{1j},\end{aligned}$$

and conclude that $\sup_{x\in[a,b]}|E\beta_{1j}|$ and, consequently, $\sup_{x\in[a,b]}|b_{1n}|$ are bounded.

It is also clear that the bounds in (4.4), (5.4) and (5.5) hold when two columns of X are removed. Moreover, with F_{n12} denoting the e.d.f. of $\sum_{j\neq 1,2} r_j r_j^*$, we get

$$\begin{aligned}\sup_{x\in[a,b]} E\left(\operatorname{tr} D_{12}^{-1}\bar{D}_{12}^{-1}\right)^4 &\le E(n\underline{\epsilon}^{-2}+v_n^{-2}nF_{n12}([a',b']))^4\\ &\le KN^4(\underline{\epsilon}^{-8}+v_n^{-8}E(F_{n12}([a',b']))^2)\le KN^4\end{aligned}$$

and

$$\sup_{x\in[a,b]} E\left(\operatorname{tr} D_1^{-2}\bar{D}_1^{-2}\right)^2 \le E(n\underline{\epsilon}^{-4}+v_n^{-4}nF_{n1}([a',b']))^2\le KN^2.$$

With these facts and (3.3) and (5.6), we have

Left-hand side of (5.8)

$$\begin{aligned}
&= \sup_{x\in[a,b]} N^{-2}\sum_{j=2}^{N} E\left|(E_j-E_{j-1})\operatorname{tr}\left(E\underline{m}_nT_n+I\right)^{-1}T_nD_1^{-1}\right|^2\\
&\le 2N^{-2}\sup_{x\in[a,b]}\sum_{j=2}^{N}E\left|\beta_{1j}r_j^*D_{1j}^{-1}\left(E\underline{m}_nT_n+I\right)^{-1}T_nD_{1j}^{-1}r_j\right|^2\\
&\le 2N^{-1}\sup_{x\in[a,b]}E\left|(b_{1n}+\beta_{12}b_{1n}\gamma_{12})r_2^*D_{12}^{-1}\left(E\underline{m}_nT_n+I\right)^{-1}T_nD_{12}^{-1}r_2\right|^2\\
&\le KN^{-1}\Big(\sup_{x\in[a,b]}E\left|r_2^*D_{12}^{-1}\left(E\underline{m}_nT_n+I\right)^{-1}T_nD_{12}^{-1}r_2\right|^2\\
&\qquad\qquad + v_n^{-2}\left(E|\gamma_{12}|^4E\left|r_2^*D_{12}^{-1}\left(E\underline{m}_nT_n+I\right)^{-1}T_nD_{12}^{-1}r_2\right|^4\right)^{1/2}\Big)\\
&\le KN^{-3}\sup_{x\in[a,b]}\Big(E\left(\operatorname{tr} D_{12}^{-2}\bar{D}_{12}^{-2}\right)+E\left(\operatorname{tr} D_{12}^{-1}\bar{D}_{12}^{-1}\right)^2\\
&\qquad\qquad + v_n^{-2}N^{-1}v_n^{-2}\left(E\operatorname{tr}\left(D_{12}^{-2}\bar{D}_{12}^{-2}\right)^2+E\left(\operatorname{tr} D_{12}^{-1}\bar{D}_{12}^{-1}\right)^4\right)^{1/2}\Big)\\
&\le KN^{-3}\left(N^2+Nv_n^{-4}\right)\le KN^{-1}.
\end{aligned}$$

Thus, we get (5.8).

Notice we get the same result if $(E\underline{m}_n T_n + I)^{-1}$ is removed from all the expressions; that is, we have just shown

$$\sup_{x\in[a,b]} E|\gamma_1 - \hat{\gamma}_1|^2 \le KN^{-1}.$$

Moreover, from (4.2) and (5.4), when $p = 2$,

$$\sup_{x\in[a,b]} E|\hat{\gamma}_1|^2 \le \sup_{x\in[a,b]} KN^{-2} E \operatorname{tr} D_1^{-1}\bar{D}_1^{-1} \le KN^{-1}.$$

Therefore,

$$\sup_{x\in[a,b]} E|\gamma_1|^2 \le KN^{-1}. \tag{5.9}$$

From (4.3), (5.2), (5.3) and (5.7)–(5.9), we get

$$\begin{aligned}
\sup_{x\in[a,b]} &\left| c_n \int \frac{dH_n(t)}{1 + tE\underline{m}_n} + zc_n E(m_n(z)) \right| \\
&\le KN^{-1} + \sup_{x\in[a,b]} \left| E\beta_1 \left[r_1^* D_1^{-1} (E\underline{m}_n T_n + I)^{-1} r_1 \right.\right. \\
&\qquad\qquad \left.\left. - \frac{1}{N} E \operatorname{tr}\big(E\underline{m}_n T_n + I\big)^{-1} T_n D_1^{-1} \right] \right| \\
&= KN^{-1} + \sup_{x\in[a,b]} |b_n|^2 \left| E(\gamma_1 - \beta_1\gamma_1^2) \left[r_1^* D_1^{-1} \big(E\underline{m}_n T_n + I\big)^{-1} r_1 \right.\right. \\
&\qquad\qquad \left.\left. - \frac{1}{N} E \operatorname{tr}\big(E\underline{m}_n T_n + I\big)^{-1} T_n D_1^{-1} \right] \right| \\
&\le K\Big(N^{-1} + \sup_{x\in[a,b]} \big(E|\gamma_1|^2 + v_n^{-2} E|\gamma_1|^4 \big)^{1/2} N^{-1/2} \Big) \\
&\le K\big(N^{-1} + (N^{-1} + v_n^{-2} N^{-2} v_n^{-4})^{1/2} N^{-1/2} \big) \\
&\le KN^{-1}.
\end{aligned}$$

As in Section 3 we let

$$w_n = -\frac{1}{z} \int \frac{dH_n(t)}{1 + tE\underline{m}_n(z)} - E(m_n(z))$$

and

$$\omega_n = -z - \frac{1}{E\underline{m}_n} + c_n \int \frac{t\, dH_n(t)}{1 + tE\underline{m}_n}.$$

Then

$$\sup_{x\in[a,b]} |w_n| \le KN^{-1},$$

$\omega_n = w_n z c_n / E\underline{m}_n$ and (3.20), together with the steps leading to (3.21), holds with $\underline{m}_n$ replaced with its expected value. From (3.10) it is clear that $\underline{m}_n^0$ must be uniformly bounded away from 0 for all $x \in [a, b]$ and all n. From (3.24) we see that $E\underline{m}_n$ must also satisfy this same property. Therefore,

$$\sup_{x\in[a,b]} |\omega_n| \le KN^{-1}.$$

Using (3.11), (3.12), (3.14) and (3.15), it follows that $\sup_{x\in[a,b]}|\underline{m}_n^0|$ is bounded in n and

$$\sup_{x\in[a,b]} \frac{\underline{m}_2^0 c_n \int \frac{t^2\, dH_n(t)}{|1+t\underline{m}_n^0|^2}}{v_n + \underline{m}_2^0 c_n \int \frac{t^2\, dH_n(t)}{|1+t\underline{m}_n^0|^2}}$$

is bounded away from 1 for all n. Therefore, we get, for all n sufficiently large,

$$\sup_{x\in[a,b]} |E\underline{m}_n - \underline{m}_n^0| \le Kc_n z \underline{m}_n^0 w_n \le KN^{-1},$$

which is (5.1).

6. Completing the proof. From the last two sections, we get

$$\sup_{x\in[a,b]} \left|\underline{m}_n(z) - \underline{m}_n^0(z)\right| = o(1/Nv_n) \quad \text{a.s.,} \tag{6.1}$$

when $v_n = N^{-1/68}$. It is clear from the arguments used in Sections 3–5 that (6.1) is true when the imaginary part of z is replaced by a constant multiple of v_n. In fact, we have

$$\max_{k\in\{1,2,\dots,34\}} \sup_{x\in[a,b]} \left|\underline{m}_n(x+i\sqrt{k}v_n) - \underline{m}_n^0(x+i\sqrt{k}v_n)\right| = o(1/Nv_n) = o(v_n^{67}) \quad \text{a.s.}$$

We take the imaginary part and get

$$\max_{k\in\{1,2,\dots,34\}} \sup_{x\in[a,b]} \left| \int \frac{d(F^{\underline{B}_n}(\lambda) - F^{c_n, H_n}(\lambda))}{(x-\lambda)^2 + kv_n^2} \right| = o(v_n^{66}) \quad \text{a.s.}$$

Upon taking differences, we find

$$\max_{k_1\neq k_2} \sup_{x\in[a,b]} \left| \int \frac{v_n^2\, d(F^{\underline{B}_n}(\lambda) - F^{c_n,H_n}(\lambda))}{((x-\lambda)^2 + k_1 v_n^2)((x-\lambda)^2 + k_2 v_n^2)} \right|$$
$$= o(v_n^{66}) \quad \text{a.s.,}$$

$$\max_{\substack{k_1, k_2, k_3 \\ \text{distinct}}} \sup_{x\in[a,b]} \left| \int \frac{(v_n^2)^2\, d(F^{\underline{B}_n}(\lambda) - F^{c_n,H_n}(\lambda))}{((x-\lambda)^2 + k_1 v_n^2)((x-\lambda)^2 + k_2 v_n^2)((x-\lambda)^2 + k_3 v_n^2)} \right|$$
$$= o(v_n^{66}) \quad \text{a.s.,}$$

$$\vdots$$

$$\sup_{x\in[a,b]}\left|\int \frac{(v_n^2)^{33}\, d(F^{\underline{B}_n}(\lambda) - F^{c_n,H_n}(\lambda))}{((x-\lambda)^2+v_n^2)((x-\lambda)^2+2v_n^2)\cdots((x-\lambda)^2+34v_n^2)}\right|$$
$$= o(v_n^{66}) \quad \text{a.s.}$$

Thus,

$$\sup_{x\in[a,b]}\left|\int \frac{d(F^{\underline{B}_n}(\lambda) - F^{c_n,H_n}(\lambda))}{((x-\lambda)^2+v_n^2)((x-\lambda)^2+2v_n^2)\cdots((x-\lambda)^2+34v_n^2)}\right| = o(1) \quad \text{a.s.}$$

We split up the integral and get

$$\begin{aligned}\sup_{x\in[a,b]}\Bigg|\int & \frac{I_{[a',b']^c}\, d(F^{\underline{B}_n}(\lambda) - F^{c_n,H_n}(\lambda))}{((x-\lambda)^2+v_n^2)((x-\lambda)^2+2v_n^2)\cdots((x-\lambda)^2+34v_n^2)} \\ &+ \sum_{\lambda_j\in[a',b']} \frac{v_n^{68}}{((x-\lambda_j)^2+v_n^2)((x-\lambda_j)^2+2v_n^2)\cdots((x-\lambda_j)^2+34v_n^2)}\Bigg| \\ &= o(1) \quad \text{a.s.}\end{aligned} \tag{6.2}$$

Now if, for each term in a subsequence satisfying (6.2), there is at least one eigenvalue contained in $[a,b]$, then the sum in (6.2), with x evaluated at these eigenvalues, will be uniformly bounded away from 0. Thus, at these same x values, the integral in (6.2) must also stay uniformly bounded away from 0. But the integral converges to 0 a.s. since the integrand is bounded and, with probability 1, both $F^{\underline{B}_n}$ and F^{c_n,H_n} converge weakly to the same limit having no mass on $\{a',b'\}$. Thus, with probability one, no eigenvalues of B_n will appear in $[a,b]$ for all n sufficiently large. This completes the proof of Theorem 1.1. □

APPENDIX

We give here a proof of Lemma 2.7. We first prove the following.

LEMMA A.1. *For* $X=(X_1,\dots,X_n)^T$ i.i.d. *standardized (complex) entries,* B $n\times n$ *Hermitian nonnegative definite matrix, we have, for any* $p\ge 1$,

$$E|X^*BX|^p \le K_p\big((\operatorname{tr} B)^p + E|X_1|^{2p}\operatorname{tr} B^p\big). \tag{A.1}$$

PROOF. Notice the result is trivially true for $p=1$. For $p>1$ we have

$$\begin{aligned}E|X^*BX|^p &\le K_p\Bigg(E\Bigg|\sum_{i=1}^n |X_i|^2 B_{ii}\Bigg|^p + E\Bigg|\sum_{i=2}^n \overline{X}_i \sum_{j<i} X_j B_{ij}\Bigg|^p \\ &\qquad + E\Bigg|\sum_{j=2}^n X_j \sum_{i<j}\overline{X}_i B_{ij}\Bigg|^p\Bigg) \\ &= K_p\Bigg(E\Bigg|\sum_{i=1}^n |X_i|^2 B_{ii}\Bigg|^p + 2E\Bigg|\sum_{i=2}^n \overline{X}_i \sum_{j<i} X_j B_{ij}\Bigg|^p\Bigg).\end{aligned} \tag{A.2}$$

Using Lemmas 2.3 and 2.9,

$$E\left|\sum_{i=1}^{n}|X_i|^2 B_{ii}\right|^p \le K_p\left((\operatorname{tr} B)^p + \sum_{i=1}^{n} E|X_1|^{2p}(B_{ii})^p\right)$$
$$\le K_p\big((\operatorname{tr} B)^p + E|X_1|^{2p}\operatorname{tr} B^p\big).$$

For $1 < p \le 2$ we have, using Lemma 2.2,

$$E\left|\sum_{i=2}^{n}\overline{X}_i\sum_{j<i}X_jB_{ij}\right|^p \le K_pE\left(\sum_{i=2}^{n}\left|\overline{X}_i\sum_{j<i}X_jB_{ij}\right|^2\right)^{p/2}$$
$$\le K_p\left(\sum_{i=2}^{n}E\left|\overline{X}_i\sum_{j<i}X_jB_{ij}\right|^2\right)^{p/2}$$
$$= K_p\left(\sum_{j<i}|B_{ij}|^2\right)^{p/2} \le K_p(\operatorname{tr} B^2)^{p/2} \le K_p(\operatorname{tr} B)^p.$$

Therefore, (A.1) is true for $p \in [1,2]$. We proceed by induction on k, where $p \in [2^k, 2^{k+1}]$. Assume (A.1) is true for $p \in [2^{k-1}, 2^k]$ and suppose $p \in [2^k, 2^{k+1}]$. Since the first term in (A.2) is bounded by the right-hand side of (A.1), we need only consider the second term. Let $\mathscr{F}_i = \sigma(X_1,\ldots,X_i)$. We have, by Lemma 2.1,

$$E\left|\sum_{i=2}^{n}\overline{X}_i\sum_{j<i}X_jB_{ij}\right|^p$$
$$\le K_p\left(E\left(\sum_{i=2}^{n}\left|\sum_{j<i}X_jB_{ij}\right|^2\right)^{p/2} + E|X_1|^p\sum_{i=2}^{n}E\left|\sum_{j<i}X_jB_{ij}\right|^p\right).$$

Using Lemma 2.4 (with $q = 2$), we have

$$E\left(\sum_{i=2}^{n}\left|\sum_{j<i}X_jB_{ij}\right|^2\right)^{p/2} = E\left(\sum_{i=2}^{n}\left|E\left(\sum_{j=1}^{n}X_jB_{ij}\middle|\mathscr{F}_{i-1}\right)\right|^2\right)^{p/2}$$
$$\le K_pE\left(\sum_{i=2}^{n}\left|\sum_{j=1}^{n}X_jB_{ij}\right|^2\right)^{p/2}$$
$$\le K_pE(X^*B^2X)^{p/2}$$
$$\le K_p\big((\operatorname{tr} B^2)^{p/2} + E|X_1|^{2(p/2)}\operatorname{tr} B^{2(p/2)}\big)$$

(by the inductive hypothesis)

$$\le K_p\big((\operatorname{tr} B)^p + E|X_1|^{2p}\operatorname{tr} B^p\big)$$

(using $1 \le E|X_1|^s \le (E|X_1|^{2s})^{1/2} \le E|X_1|^{2s}$ for $s \ge 2$).

Using Lemma 2.1, we have

$$\begin{aligned}
&E|X_1|^p \sum_{i=2}^{n} E\left|\sum_{j<i} X_j B_{ij}\right|^p \\
&\quad \le K_p E|X_1|^p \sum_{i=2}^{n}\left(\left(\sum_{j<i}|B_{ij}|^2\right)^{p/2} + E|X_1|^p \sum_{j<i}|B_{ij}|^p\right) \\
&\quad \le K_p E|X_1|^p (1+E|X_1|^p) \sum_{i=2}^{n}\left(\sum_{j<i}|B_{ij}|^2\right)^{p/2} \\
&\quad \le K_p E|X_1|^p (1+E|X_1|^p) \sum_{i=1}^{n}((B^2)_{ii})^{p/2} \\
&\quad \le K_p E|X_1|^p (1+E|X_1|^p)\, \mathrm{tr}\, B^p \qquad \text{(by Lemma 2.9)} \\
&\quad \le K_p E|X_1|^{2p}\, \mathrm{tr}\, B^p.
\end{aligned}$$

Therefore, (A.1) is true for $p \in [2^k, 2^{k+1}]$ and the proof of the lemma is complete.

We can now prove Lemma 2.7. We have

$$\begin{aligned}
&E|X^* C X - \mathrm{tr}\, C|^p \\
&\quad \le K_p\left(E\left|\sum_{i=1}^{n}(|X_i|^2-1)C_{ii}\right|^p\right. \\
&\qquad\qquad \left. + E\left|\sum_{i=2}^{n} \bar{X}_i \sum_{j<i} X_j C_{ij}\right|^p + E\left|\sum_{j=2}^{n} X_j \sum_{i<j} \bar{X}_i C_{ij}\right|^p\right).
\end{aligned}$$

Using Lemma 2.1,

$$\begin{aligned}
&E\left|\sum_{i=1}^{n}(|X_i|^2-1)C_{ii}\right|^p \\
&\quad \le K_p\left(\left(\sum_{i=1}^{n} E(|X_i|^2-1)^2|C_{ii}|^2\right)^{p/2} + \sum_{i=1}^{n} E||X_1|^2-1|^p|C_{ii}|^p\right) \\
&\quad \le K_p\left((E|X_1|^4\, \mathrm{tr}\, CC^*)^{p/2} + E|X_1|^{2p}\sum_{i=1}^{n}|C_{ii}|^p\right)
\end{aligned}$$

(using $(E||X_1|^2-1|^p)^{1/p} \le (E|X_1|^{2p})^{1/p} + 1 \le 2(E|X_1|^{2p})^{1/p}$). From Lemma 2.9 we have

$$\sum_{i=1}^{n}|C_{ii}|^p \le \sum_{i=1}^{n}(CC^*)_{ii}^{p/2} \le \sum_{i=1}^{n}\lambda_i(CC^*)^{p/2}.$$

Therefore,

$$E\left|\sum_{i=1}^{n}(|X_i|^2-1)C_{ii}\right|^p \le K_p\big((E|X_1|^4\, \mathrm{tr}\, CC^*)^{p/2} + E|X_1|^{2p}\, \mathrm{tr}(CC^*)^{p/2}\big).$$

Using Lemma 2.1,

$$E\left|\sum_{i=2}^{n}\overline{X}_i\sum_{j<i}X_jC_{ij}\right|^p$$

$$\le K_p\left(E\left(\sum_{i=2}^{n}\left|\sum_{j<i}X_jC_{ij}\right|^2\right)^{p/2}+E|X_1|^p\sum_{i=2}^{n}E\left|\sum_{j<i}X_jC_{ij}\right|^p\right).$$

Using Lemma 2.4 (with $q=2$),

$$E\left(\sum_{i=2}^{n}\left|\sum_{j<i}X_jC_{ij}\right|^2\right)^{p/2}$$

$$=E\left(\sum_{i=2}^{n}\left|E\left(\sum_{j=1}^{n}X_jC_{ij}\middle|\mathscr{F}_{i-1}\right)\right|^2\right)^{p/2}$$

$$\le K_pE\left(\sum_{i=2}^{n}\left|\sum_{j=1}^{n}X_jC_{ij}\right|^2\right)^{p/2}\le K_pE(X^*C^*CX)^{p/2}$$

$$\le K_p\big((\operatorname{tr}C^*C)^{p/2}+E|X_1|^{2p}\operatorname{tr}(C^*C)^{p/2}\big)\qquad\text{(by Lemma A.1)}$$

$$=K_p\big((\operatorname{tr}CC^*)^{p/2}+E|X_1|^{2p}\operatorname{tr}(CC^*)^{p/2}\big).$$

Using Lemma 2.1, we have

$$E|X_1|^p\sum_{i=2}^{n}E\left|\sum_{j<i}X_jC_{ij}\right|^p$$

$$\le K_pE|X_1|^p\sum_{i=2}^{n}\left(\left(\sum_{j<i}|C_{ij}|^2\right)^{p/2}+E|X_1|^p\sum_{j<i}|C_{ij}|^p\right)$$

$$\le K_pE|X_1|^p(1+E|X_1|^p)\sum_{i=2}^{n}\left(\sum_{j<i}|C_{ij}|^2\right)^{p/2}$$

$$\le K_pE|X_1|^p(1+E|X_1|^p)\sum_{i=1}^{n}((CC^*)_{ii})^{p/2}$$

$$\le K_pE|X_1|^p(1+E|X_1|^p)\operatorname{tr}(CC^*)^{p/2}\qquad\text{(by Lemma 2.9)}$$

$$\le K_pE|X_1|^{2p}\operatorname{tr}(CC^*)^{p/2}.$$

Therefore, $E|\sum_{i=2}^{n}\overline{X}_i\sum_{j<i}X_jC_{ij}|^p$ is bounded by the right-hand side of the inequality in Lemma 2.7. Similarly, $E|\sum_{j=2}^{n}X_j\sum_{i<j}\overline{X}_iC_{ij}|^p$ is also bounded by the right-hand side of the inequality, and the proof of the lemma is complete. □

Acknowledgments. The authors would like to thank Pawel Hitczenko for pointing out some of the moment bounds in Section 2, and to an anonymous referee for some helpful suggestions. Part of this work was done while

J. W. Silverstein visited the Department of Applied Mathematics at National Sun Yat-sen University. He thanks the members of the department for their hospitality, and the National Science Council of Taiwan for their financial support.

REFERENCES

BAI, Z. D., SILVERSTEIN, J. W. and YIN, Y. Q. (1988). A note on the largest eigenvalue of a large dimensional sample covariance matrix. *J. Multivariate Anal.* 26 166–168.

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

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

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

DILWORTH, S. J. (1993). Some probabilistic inequalities with applications to functional analysis. *Banach Spaces. Contemp. Math.* 144 53–67.

HORN, R. A. and JOHNSON, C. R. (1985). *Matrix Analysis*. Cambridge Univ. Press.

HORN, R. A. and JOHNSON, C. R. (1991). *Topics in Matrix Analysis*. Cambridge Univ. Press.

ROSENTHAL, H. P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* 273–303.

SILVERSTEIN, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* 5 331–339.

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

SILVERSTEIN, J. W. and COMBETTES, P. L. (1992). Signal detection via spectral theory of large dimensional random matrices. *IEEE Trans. Signal Process.* 40 2100–2105.

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

DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
10 KENT RIDGE CRESCENT
SINGAPORE 119260
E-MAIL: matbaizd@leonis.nus.sg

DEPARTMENT OF MATHEMATICS
BOX 8205
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NORTH CAROLINA 27695-8205
E-MAIL: jack@jack.math.ncsu.edu