The Annals of Probability
1999, Vol. 27, No. 3, 1536–1555

EXACT SEPARATION OF EIGENVALUES 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 Eightieth Birthday of C. R. Rao

Let $B_n = (1/N)T_n^{1/2}X_nX_n^*T_n^{1/2}$ 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 was shown in an earlier paper by the authors that, under certain conditions on the eigenvalues of T_n, with probability 1 no eigenvalues lie in any interval which is outside the support of the limiting empirical distribution (known to exist) for all large n. For these n the interval corresponds to one that separates the eigenvalues of T_n. The aim of the present paper is to prove exact separation of eigenvalues; that is, with probability 1, the number of eigenvalues of B_n and T_n lying on one side of their respective intervals are identical for all large n.

1. Introduction. The main result in this paper completes the analysis begun in Bai and Silverstein (1998) (hereafter referred to as BS (1998)) on the location of eigenvalues of the $n \times n$ matrix $B_n = (1/N)T_n^{1/2}X_nX_n^*T_n^{1/2}$ when n is large. Here $X_n = (X_{ij})$ is $n \times N$ consisting of i.i.d. standardized complex entries ($\mathsf{E}X_{11} = 0$, $\mathsf{E}|X_{11}|^2 = 1$), T_n is an $n \times n$ nonnegative definite matrix and $T_n^{1/2}$ is any Hermitian square root of T_n. It is assumed that $N = N(n)$ with $n/N \to c > 0$ as $n \to \infty$ and, with F^A denoting the empirical distribution function (e.d.f.) of the eigenvalues of any matrix A having real eigenvalues, it is also assumed that $F^{T_n} \to_{\mathscr{D}} H$, a cumulative distribution function (c.d.f.). It follows [Silverstein (1995)] that with probability 1, $F^{B_n} \to_{\mathscr{D}} F$, a nonrandom c.d.f. With the additional assumption that all X_n come from the upper left portion of a doubly infinite array of independent random variables having finite fourth moment, along with some additional conditions on F^{T_n}, it is shown in BS (1998) the almost sure absence of eigenvalues of B_n in any closed interval which lies outside the support of F in $\mathbb{R}^+$ for all n sufficiently large. The result is stated in Theorem 1.1 below. The aim of this paper is to prove that the proper number of eigenvalues lie on either side of these intervals.

The precise meaning of the last statement as well as the significance of the two results become apparent when B_n is viewed as the sample covariance

Received September 1998; revised February 1999.

[1]Supported in part by National University of Singapore Grant RP 3972712.

[2]Supported by NSF Grants DMS-94-04047 and DMS-97-03591.

AMS 1991 *subject classifications*. Primary 15A52, 60F15; secondary 62H99.

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

matrix of N samples of the random vector $T_n^{1/2}X_{\bullet 1}$ ($X_{\bullet j}$ denoting the jth column of X_n). From the law of large numbers, for N large relative to n, B_n will with high probability be close to T_n, the population covariance matrix. Thus for small c (the limiting ratio of vector dimension to sample size), on an interval $J \subset \mathbb{R}^+$ for which no eigenvalues of T_n appear for all n large, it seems reasonable to expect the same to occur for B_n on some interval $[a, b]$ close to J, with the number of eigenvalues of B_n on one side of $[a, b]$ matching up with those of T_n on the same side of J. As will be presently seen, these statements can be proved quite easily for c sufficiently small, provided the eigenvalues of T_n are bounded in n. However, c need not be small for the support of F to split.

To prove the above for c small, we use results on the eigenvalues of B_n when $T = I$, the identity matrix, and the following lemma.

LEMMA 1.1 [Fan (1951)]. *For rectangular matrix A and positive integer $i \le$ rank A, let λ_i^A denote the i^{th} largest singular value of A. Define λ_i^A to be zero for all $i >$ rank A. Let m, n be arbitrary nonnegative integers.*

Then, for A, B rectangular for which AB is defined,

$$\lambda_{m+n+1}^{AB} \le \lambda_{m+1}^{A}\lambda_{n+1}^{B}.$$

Extending the notation introduced in Lemma 1.1 to eigenvalues, and for notational convenience, defining $\lambda_0^A = \infty$, suppose $\lambda_{i_n}^{T_n}$ and $\lambda_{i_n+1}^{T_n}$ lie, respectively, to the right and left of J. From Lemma 1.1 we have (using the fact that the spectra of B_n and $(1/N)X_nX_n^*T_n$ are identical)

$$\lambda_{i_n+1}^{B_n} \le \lambda_1^{(1/N)X_nX_n^*}\lambda_{i_n+1}^{T_n} \quad \text{and} \quad \lambda_{i_n}^{B_n} \ge \lambda_n^{(1/N)X_nX_n^*}\lambda_{i_n}^{T_n}. \tag{1.1}$$

It is well known [dating back to Marčenko and Pastur (1967)] that the limiting spectral distribution of $(1/N)X_nX_n^*$ has support $[(1-\sqrt{c})^2, (1+\sqrt{c})^2]$, with the addition of 0 when $c > 1$. Moreover, when the entries of X_n have finite fourth moment and arise (as stated above) from one doubly infinite array we have the following lemma.

LEMMA 1.2 [Yin, Bai and Krishnaiah (1988), Bai and Yin (1993)]. *With probability 1, the largest eigenvalue of $(1/N)X_nX_n^*$ converges to $(1+\sqrt{c})^2$, while the $\min(n, N)$th largest eigenvalue (the smallest when $c < 1$) converges to $(1-\sqrt{c})^2$.*

Thus from (1.1) and Lemma 1.2, interval $[a, b]$ exists which splits the eigenvalues in exactly the same manner as J, and its endpoints can be made arbitrarily close to those of J by choosing c sufficiently small.

The goal of this paper is to extend the above result of exact separation whenever the support of F splits, regardless of the size of c. As an example of its relevancy, consider the detection problem in array signal processing. An unknown number q of sources emit signals onto an array of n sensors in a noise-filled environment ($q < n$). From the population covariance matrix R

arising from the vector of random values recorded from the sensors, the value q can be determined because the multiplicity of the smallest eigenvalue of R, attributed to the noise, is $n - q$. The matrix R is approximated by a sample covariance matrix $\widehat{R}$ which, with a sufficiently large sample, will have, with high probability, $n - q$ *noise* eigenvalues clustering near each other and to the left of the other eigenvalues. The problem is, for n sizable the number of samples needed for $\widehat{R}$ to adequately approximate R would be prohibitively large. However, if for n large the number of samples were to be merely on the same order of magnitude as n, then, under certain conditions, it is shown in Silverstein and Combettes (1992) that $F^{\widehat{R}}$ would, with high probability, be close to the nonrandom limiting c.d.f. F. Moreover, it can be shown that for c sufficiently small, the support of F will split into two parts, with mass $(n-q)/n$ on the left, q/n on the right. In Silverstein and Combettes (1992) extensive computer simulations were performed to demonstrate that, at the least, the proportion of sources to sensors can be reliably estimated. It came as a surprise to find that, not only were there no eigenvalues outside the support of F, except those near the boundary of the support [verified in BS (1998)], but the *exact* number of eigenvalues appeared on intervals slightly larger than those within the support of F (the aim of this paper). Thus, the simulations demonstrate that, in order to detect the *number* of sources in the large dimensional case, it is not necessary for $\widehat{R}$ to be close to R; the number of samples only needs to be large enough so that the support of F splits.

To establish exact separation whenever there is an interval $[a, b]$ outside the support of the limiting F, an interval J must be identified which is naturally associated with it. It is at this point necessary to review properties of F. The best way of understanding F is through the limiting e.d.f. of the eigenvalues of $\underline{B}_n \equiv (1/N)X_n^* T_n X_n$ and properties of its Stieltjes transform $m_{\underline{B}_n}$, which for any c.d.f. G is defined by

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

Since the spectra of B_n and $\underline{B}_n$ differ by $|n - N|$ zeros, it is easy to verify

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

(I_A denoting the indicator function of the set A), from which we get

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

Let $F^{c,H}$ denote the a.s. limit of $F^{\underline{B}_n}$. Thus

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

and

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

The main result in BS (1998) can now be stated.

THEOREM 1.1 [Theorem 1.1 of BS (1998)]. *Assume*:

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

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

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

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

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

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

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

Note: The phrase "in an open interval" was inadvertently left out of the original paper. Also, we should point out that the original paper included the condition that $[a, b]$ also lie outside the support of $F^{c, H}$. Assumption (f) presented here implies this condition.

Our main result will make the same assumptions as those in Theorem 1.1.

Attention is drawn to $F^{c, H}$ partly because $m_{F^{c,H}}(z)$ is invertible, with inverse

$$z_{c, H}(m) \equiv -\frac{1}{m} + c\int \frac{t}{1 + tm} dH(t) \tag{1.2}$$

[see BS (1998)].

From (1.2) much of the analytic behavior of F can be derived [see Silverstein and Choi (1995)]. This includes the continuous dependence of F on c and H, the fact that F has a continuous density on $\mathbb{R}^+$ and, most importantly for this paper, a way of understanding the support of F. On any closed interval outside the support of $F^{c, H}$, $m_{F^{c,H}}$ exists and is increasing. Therefore, on the range of this interval its inverse exists and is also increasing. In Silverstein and Choi (1995) the converse is shown to be true, along with some other results. We summarize the relevant facts in the following

LEMMA 1.3 [Silverstein and Choi (1995)]. *Let for any c.d.f.* G S_G *denote its support and* S'_G *the complement of its support. If* $x \in S'_{F^{c,H}}$ *then* $m = m_{F^{c,H}}(x)$ *satisfies* (1) $m \in \mathbb{R}\backslash\{0\}$, (2) $-m^{-1} \in S'_H$ *and* (3) $z'_{c, H}(m) > 0$. *Conversely, if* m *satisfies* (1)–(3), *then* $x = z_{c, H}(m) \in S'_{F^{c,H}}$.

Thus by plotting $z_{c, H}(m)$ for $m \in \mathbb{R}$, the range of values where it is increasing yields $S'_{F^{c,H}}$ [see Figure 1 of BS (1998) for an example]. Of course, the supports of F and $F^{c, H}$ are identical on $\mathbb{R}^+$. As for whether F places any mass at 0, it is shown in Silverstein and Choi (1995) that

$$F^{c, H}(0) = \max(0, 1 - c[1 - H(0)]),$$

which implies

$$F(0) = \begin{cases} H(0), & c[1 - H(0)] \le 1, \\ 1 - c^{-1}, & c[1 - H(0)] > 1. \end{cases} \tag{1.3}$$

Assume $m_{F^{c,H}}(b) < 0$. Because of assumption (f) in Theorem 1.1 and Lemma 1.3 the interval

$$[-1/m_{F^{c,H}}(a), -1/m_{F^{c,H}}(b)]$$

is contained in S'_{H_n} for all large n. We take J to be this interval.

Let for large n integer $i_n \ge 0$ be such that

$$\lambda_{i_n}^{T_n} > -1/m_{F^{c,H}}(b) \quad \text{and} \quad \lambda_{i_n+1}^{T_n} < -1/m_{F^{c,H}}(a). \tag{1.4}$$

It will be seen that only when $m_{F^{c,H}}(b) < 0$ will exact separation occur.

To understand why interval J should be linked to $[a, b]$, we need to analyze the dependence of intervals in $S'_{F^{c,H}}$ on c. We state this dependence in the following lemma, the proof given in Sections 2 and 5.

LEMMA 1.4. (a) *If (t_1, t_2) is contained in S'_H with $t_1, t_2 \in \partial S_H$ and $t_1 > 0$, then there is a $c_0 > 0$ for which $c < c_0$ implies that there are two values $m_c^1 < m_c^2$ in $[-t_1^{-1}, -t_2^{-1}]$ for which $(z_{c,H}(m_c^1), z_{c,H}(m_c^2)) \subset S'_{F^{c,H}}$, with endpoints lying in $\partial S_{F^{c,H}}$, and $z_{c,H}(m_c^1) > 0$. Moreover,*

$$z_{c,H}(m_c^i) \to t_i \quad \text{as } c \to 0 \tag{1.5}$$

for $i = 1, 2$. The endpoints vary continuously with c shrinking down to a point as $c \uparrow c_0$ while $z_{c,H}(m_c^2) - z_{c,H}(m_c^1)$ is monotone in c.

(b) *If $(t_3, \infty) \subset S'_H$ with $0 < t_3 \in \partial S_H$, then there exists $m_c^3 \in [-1/t_3, 0)$ such that $z_{c,H}(m_c^3)$ is the largest number in $S_{F^{c,H}}$. As c decreases from ∞ to 0, (1.5) holds for $i = 3$ with convergence monotone from ∞ to t_3.*

(c) *If $c[1 - H(0)] < 1$ and $(0, t_4) \subset S'_H$ with $t_4 \in \partial S_H$, then there exists $m_c^4 \in (-\infty, -1/t_4]$ such that $z_{c,H}(m_c^4)$ is the smallest positive number in $S_{F^{c,H}}$, and (1.5) holds with $i = 4$, the convergence being monotone from 0 as c decreases from $[1 - H(0)]^{-1}$.*

(d) *If $c[1 - H(0)] > 1$, then, regardless of the existence of $(0, t_4) \subset S'_H$, there exists $m_c > 0$ such that $z_{c,H}(m_c) > 0$ and is the smallest number in $S_{F^{c,H}}$. It decreases from ∞ to 0 as c decreases from ∞ to $[1 - H(0)]^{-1}$.*

(e) *If $H = I_{[0,\infty)}$, that is, H places all mass at 0, then $F = F^{c, I_{[0,\infty)}} = I_{[0,\infty)}$.*

All intervals in $S'_{F^{c,H}} \cap [0, \infty)$ arise from one of the above. Moreover, disjoint intervals in S'_H yield disjoint intervals in $S'_{F^{c,H}}$.

Thus it is clear how important a role the Stieltjes transform (and its inverse) plays in associating intervals in $S'_{F^c,H}$ with the eigenvalues of T_n.

The main result can now be stated.

THEOREM 1.2. *Assume* (a)–(f) *of Theorem* 1.1.

(i) *If* $c[1 - H(0)] > 1$, *then* x_0, *the smallest value in the support of* $F^{c,H}$, *is positive*, *and with probability* 1 $\lambda_N^{B_n} \to x_0$ *as* $n \to \infty$. *The number* x_0 *is the maximum value of the function* $z_{c,H}(m)$ *for* $m \in \mathbb{R}^+$.

(ii) *If* $c[1 - H(0)] \le 1$, *or* $c[1 - H(0)] > 1$ *but* $[a, b]$ *is not contained in* $[0, x_0]$, *then* $m_{F^{c,H}}(b) < 0$ *and with* i_n *defined as in* (1.4) *we have*

$$\mathsf{P}\Big(\lambda_{i_n}^{B_n} > b \text{ and } \lambda_{i_n+1}^{B_n} < a \text{ for all large } n\Big) = 1.$$

Conclusion (i) should not be surprising since in this case $N < n$ for large n and so $\lambda_{N+1}^{B_n} = 0$. Therefore exact separation should not be expected to occur for $[a, b] \subset [0, x_0]$. Notice that this result is consistent with (1.3). Essentially, the $n - N$ smallest eigenvalues of T_n are tranferred (via B_n) to zero. What is worth noting is that when $c[1 - H(0)] > 1$ and F and (consequently) H each has at least two nonconnected members in their support in $\mathbb{R}^+$, the numbers of eigenvalues of B_n and T_n will match up in each respective member, *except* the left-most member. Thus the transference to zero is affecting only this member.

The proof of Theorem 1.2 will be given in the following sections. The proof of both parts rely heavily on Theorem 1.1 and Lemma 1.2. The proof of (ii) involves systematically increasing the number of columns of X_n, keeping track on the movements of the eigenvalues of the new matrices, until the limiting c is sufficiently small that the result obtained at the beginning of this section can be used.

2. Proof of Theorem 1.2(i). We see that x_0 must coincide with the boundary point in (d) of Lemma 1.4. Most of (d) will be proved in the following

LEMMA 2.1. *If* $c[1 - H(0)] > 1$, *then the smallest value in the support of* F^{c_n,H_n} *is positive for all* n *large*, *and it converges to the smallest value*, *also positive*, *in the support of* $F^{c,H}$ *as* $n \to \infty$.

PROOF. Assume $c[1 - H(0)] > 1$. Write

$$z_{c,H}(m) = \frac{1}{m}\left(-1 + c\int \frac{tm}{1+tm}\, dH(t)\right),$$

$$z'_{c,H}(m) = \frac{1}{m^2}\left(1 - c\int \left(\frac{tm}{1+tm}\right)^2 dH(t)\right).$$

As m increases in $\mathbb{R}^+$, the two integrals increase from 0 to $1 - H(0)$, which implies $z_{c,H}(m)$ increases from $-\infty$ to a maximum value and decreases to zero. Let $\hat{m}$ denote the number where the maximum occurs. Then by Lemma 1.3, $x_0 \equiv z_{c,H}(\hat{m})$ is the smallest value in the support of $F^{c,H}$. We see that $\hat{m}$ is m_c in (d) of Lemma 1.4.

We have

$$c\int\left(\frac{t\hat{m}}{1+t\hat{m}}\right)^2 dH(t)=1.$$

From this it is easy to verify

$$z_{c,H}(\hat{m})=c\int\frac{t}{(1+t\hat{m})^2}\,dH(t).$$

Therefore $z_{c,H}(\hat{m})>0$.

Since $\limsup_n H_n(0)\le H(0)$, we have $c_n(1-H_n(0))>1$ for all large n. We consider now only these n and we let $\hat{m}_n$ denote the value where the maximum of $z_{c_n,H_n}(m)$ occurs in $\mathbb{R}^+$. We see that $z_{c_n,H_n}(\hat{m}_n)$ is the smallest positive value in the support of F^{c_n,H_n}. It is clear that for all positive m $z_{c_n,H_n}(m)\to z_{c,H}(m)$ and $z'_{c_n,H_n}(m)\to z'_{c,H}(m)$ as $n\to\infty$, uniformly on any closed subset of $\mathbb{R}^+$. Thus, for any positive m_1, m_2 such that $m_1<\hat{m}<m_2$, we have for all n large

$$z'_{c_n,H_n}(m_1)>0>z'_{c_n,H_n}(m_2)$$

which implies $m_1<\hat{m}_n<m_2$. Therefore, $\hat{m}_n\to\hat{m}$ and, in turn, $z_{c_n,H_n}(\hat{m}_n)\to x_0$ as $n\to\infty$. □

We now prove when $c[1-H(0)]>1$,

$$\lambda_N^{B_n}\to x_0 \quad \text{a.s. as } n\to\infty. \tag{2.1}$$

Assume first that T_n is nonsingular with $\lambda_n^{T_n}$ uniformly bounded away from 0. Using Lemma 1.1, we find

$$\lambda_N^{(1/N)X_nX_n^*}\le\lambda_N^{B_n}\lambda_1^{T_n^{-1}}=\lambda_N^{B_n}\left(\lambda_n^{T_n}\right)^{-1}.$$

Since by Lemma 1.2 $\lambda_N^{(1/N)X_nX_n^*}\to(1-\sqrt{c})^2$ a.s. as $n\to\infty$ we conclude that $\liminf_n\lambda_N^{B_n}>0$ a.s. Since, by Lemma 2.1, the interval $[a,b]$ in Theorem 1.1 can be made arbitrarily close to $(0,x_0)$, we get

$$\liminf_n\lambda_N^{B_n}\ge x_0 \quad \text{a.s.}$$

However, since $F^{B_n}\to_{\mathscr{D}}F$ a.s., we must have

$$\limsup_n\lambda_N^{B_n}\le x_0 \quad \text{a.s.}$$

Thus we get (2.1).

For general T_n, let for $\varepsilon>0$ suitably small T_n^ε denote the matrix resulting from replacing all eigenvalues of T_n less than ε with ε. Let $H_n^\varepsilon=F^{T_n^\varepsilon}=I_{[\varepsilon,\infty)}H_n$. Then $H_n^\varepsilon\to_{\mathscr{D}}H^\varepsilon\equiv I_{[\varepsilon,\infty)}H$. Let B_n^ε denote the sample covariance matrix corresponding to T_n^ε.

Let $\hat{m}^\varepsilon$ denote the value where the maximum of $z_{c,H^\varepsilon}(m)$ occurs on $\mathbb{R}^+$. Then

$$\lambda_N^{B_n^\varepsilon}\to z_{c,H^\varepsilon}(\hat{m}^\varepsilon) \quad \text{a.s. as } n\to\infty. \tag{2.2}$$

Using Corollary 7.3.8 of Horn and Johnson (1985), we have

$$\begin{aligned}\left|\lambda_N^{B_n^\varepsilon}-\lambda_N^{B_n}\right| &= \left|\lambda_N^{(1/N)X_n^*T_n^\varepsilon X_n}-\lambda_N^{(1/N)X_n^*T_nX_n}\right| \\ &\le \left\|\frac{1}{N}X_n^*(T_n^\varepsilon-T_n)X_n\right\| \\ &\le \left\|\frac{1}{N}X_nX_n^*\right\|\varepsilon.\end{aligned} \tag{2.3}$$

Since $c[1-H^\varepsilon(0)]=c>1$, we get from Lemma 2.1,

$$x^\varepsilon(\hat{m}^\varepsilon)\to z_{c,H}(\hat{m}) \quad \text{as } \varepsilon\to 0. \tag{2.4}$$

Therefore, from (2.2)–(2.4) and the a.s. convergence of $\lambda_1^{(1/N)X_nX_n^*}$ (Lemma 1.2), we get $\liminf_n \lambda_N^{B_n}>0$ a.s. which, as above, implies (2.1). □

The proof of Theorem 1.2(ii) will be given in the following sections.

3. Convergence of a random quadratic form. We begin this section by simplifying the conditions on the entries of X_n. For $C>0$ let $Y_{ij}=X_{ij}I_{[|X_{ij}|\le C]}-\mathsf{E}X_{ij}I_{[|X_{ij}|\le C]}$, $Y_n=(Y_{ij})$ and $\tilde{B}_n=(1/N)T^{1/2}Y_nY_n^*T^{1/2}$. It is shown in BS (1998) that with probability 1,

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

It is clear that we can make the above bound arbitrarily small by choosing C sufficiently large. Thus, in proving Theorem 1.2(ii) it is enough to consider the case where the underlying variables are bounded. Moreover, it is evident from the proofs in Bai and Yin (1993) and BS (1998) that Lemma 1.2 is still true, as well as the conclusion to Theorem 1.1 with only X_{11} bounded, standardized, and no assumptions on the relationship between X_n for varying n (i.e., the entries of X_n need not come from the same doubly infinite array).

Another simplifying assumption is on the size of $\|T_n\|$. Since it is assumed to be bounded, we may assume $\|T_n\|\le 1$.

For this section we need the following two results previously proved.

LEMMA 3.1 [(3.3) of BS (1998)]. *Let $X_{\bullet 1}$ denote the first column of X_n. Then for any $p\ge 2$ and $n\times n$ matrix C (complex), there exists $K_p>0$ depending only on p and the distribution of X_{11} such that*

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

LEMMA 3.2 [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 $q\in\mathbb{C}^n$. Then*

$$\left|\operatorname{tr}\left((B-zI)^{-1}-(B+\tau qq^*-zI)^{-1}\right)A\right|\le\frac{\|A\|}{v}.$$

The goal of this section is to prove a limiting result on a random quadratic form involving the resolvent of B_n.

LEMMA 3.3. *Let x be any point in $[a, b]$ and $m = m_{F^{c,H}}(x)$. Let $\widetilde{X} \in \mathbb{C}^n$ be distributed the same as $X_{\bullet 1}$ and independent of X_n. Set $r = r_n = (1/\sqrt{N})T_n^{1/2}\widetilde{X}$. Then*

$$r^*(xI - B_n)^{-1}r \to 1 + \frac{1}{xm} \quad \text{a.s. as } n \to \infty. \tag{3.1}$$

PROOF. Let B_n^{N+1} denote $(1/N)T_n^{1/2}X_n^{N+1}X_n^{N+1*}T_n^{1/2}$, where X_n^{N+1} is $n \times (N+1)$ and contains i.i.d. entries (same distribution as X_{11}), and $\underline{B}_n^{N+1} = (1/N)X_n^{N+1*}T_nX_n^{N+1}$. Let $z = x + iv$, $v > 0$. For Hermitian A, let m_A denote the Stieltjes transform of the spectral distribution of A. We have

$$m_A(z) = \frac{1}{n}\operatorname{tr}(zI - A)^{-1}.$$

Therefore, using Lemma 3.2, we have

$$\left|m_{B_n}(z) - m_{B_n^{N+1}}(z)\right| \le \frac{1}{nv}.$$

From

$$m_{\underline{B}_n}(z) = -\frac{1 - n/N}{z} + \frac{n}{N}m_{B_n}(z)$$

and

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

we conclude

$$\left|m_{\underline{B}_n}(z) - m_{\underline{B}_n^{N+1}}(z)\right| \le \frac{(2c_n + 1)}{v(N+1)}. \tag{3.2}$$

For $j = 1, 2, \ldots, N+1$, let $r_j = (1/\sqrt{N})T_n^{1/2}X_{\bullet j}$ ($X_{\bullet j}$ denoting the jth column of X_n^{N+1}) and $B_{(j)} = B_n^{N+1} - r_jr_j^*$. Notice $B_{(N+1)} = B_n$.

Generalizing formula (2.2) in Silverstein (1995), we find for any $n \times M$ matrix C with jth column denoted by c_j and $C_{(j)}$ denoting C without the jth column,

$$m_{C^*C}(z) = -\frac{1}{M}\sum_{j=1}^{M}\frac{1}{z(1 + c_j^*(C_{(j)}C_{(j)}^* - zI)^{-1}c_j)}.$$

It is easy to verify

$$\operatorname{Im} c_j^*((1/z)C_{(j)}C_{(j)}^* - I)^{-1}c_j \ge 0,$$

which implies

$$\frac{1}{|z(1+c_j^*(C_{(j)}C_{(j)}^*-zI)^{-1}c_j)|}\le\frac{1}{v}. \tag{3.3}$$

Thus we have

$$m_{\underline{B}_n^{N+1}}(z)=-\frac{1}{N+1}\sum_{j=1}^{N+1}\frac{1}{z(1+r_j^*(B_{(j)}-zI)^{-1}r_j)}. \tag{3.4}$$

Let $\mu_n(z)=-1/z(1+r^*(B_n-zI)^{-1}r)$, where $r=r_{N+1}$.

Our present goal is to show that for any $i\le N+1$, $\varepsilon>0$, $z=z_n=x+v_n$ with $v_n=N^{-\delta}$, $\delta\in[0,1/3)$ and $p>2$, we have for all n sufficiently large,

$$\mathsf{P}\big(|m_{\underline{B}_n(z)}-\mu_n(z)|>\varepsilon\big)\le K_p\left(\frac{|z|}{\varepsilon v_n^3}\right)^p\frac{n^{p/2}}{N^{p-1}}. \tag{3.5}$$

We have from (3.4)

$$\begin{aligned}
&m_{\underline{B}_n^{N+1}}(z)-\mu_n(z)\\
&\quad=-\frac{1}{(N+1)z}\sum_{j=1}^{N}\left(\frac{1}{(1+r_j^*(B_{(j)}-zI)^{-1}r_j)}-\frac{1}{(1+r^*(B_n-zI)^{-1}r)}\right)\\
&\quad=-\frac{1}{(N+1)z}\sum_{j=1}^{N}\frac{r^*(B_n-zI)^{-1}r-r_j^*(B_{(j)}-zI)^{-1}r_j}{(1+r^*(B_n-zI)^{-1}r)(1+r_j^*(B_{(j)}-zI)^{-1}r_j)}.
\end{aligned}$$

Using (3.3) we find

$$\big|m_{\underline{B}_n^{N+1}}(z)-\mu_n(z)\big|\le\frac{|z|}{v_n^2}\max_{j\le N}\big|r^*(B_n-zI)^{-1}r-r_j^*(B_{(j)}-zI)^{-1}r_j\big|. \tag{3.6}$$

Write

$$\begin{aligned}
&r^*(B_n-zI)^{-1}r-r_j^*(B_{(j)}-zI)^{-1}r_j\\
&\quad=r^*(B_n-zI)^{-1}r-\frac{1}{N}\operatorname{tr}T_n^{1/2}(B_n-zI)^{-1}T_n^{1/2}\\
&\qquad-\left(r_j^*(B_{(j)}-zI)^{-1}r_j-\frac{1}{N}\operatorname{tr}T_n^{1/2}(B_{(j)}-zI)^{-1}T_n^{1/2}\right)\\
&\qquad+\frac{1}{N}\operatorname{tr}((B_n-zI)^{-1}-(B_{(j)}-zI)^{-1})T_n.
\end{aligned}$$

Using Lemma 3.2 we find

$$\frac{1}{N}\big|\operatorname{tr}((B_n-zI)^{-1}-(B_{(j)}-zI)^{-1})T_n\big|\le\frac{2}{Nv_n}. \tag{3.7}$$

Using Lemma 3.1 we have for any $j \le N+1$ and $p \ge 2$,

$$\begin{aligned}
&\mathsf{E}\left| r_j^*(B_{(j)} - zI)^{-1} r_j - \frac{1}{N}\mathrm{tr} T_n^{1/2}(B_{(j)} - zI)^{-1} T_n^{1/2} \right|^p \\
&\quad \le K_p \frac{1}{N^p} \mathsf{E}(\mathrm{tr} T_n^{1/2}(B_{(j)} - zI)^{-1} T_n (B_{(j)} - \bar{z}I)^{-1} T_n^{1/2})^{p/2} \\
&\quad \le K_p \frac{1}{N^p} \mathsf{E}(\mathrm{tr}(B_{(j)} - zI)^{-1} T_n (B_{(j)} - \bar{z}I)^{-1})^{p/2} \\
&\quad \le K_p \frac{1}{N^p} \mathsf{E}(\mathrm{tr}(B_{(j)} - zI)^{-1} (B_{(j)} - \bar{z}I)^{-1})^{p/2} \\
&\quad \le K_p \frac{1}{N^p} \left(\frac{n}{v_n^2} \right)^{p/2}.
\end{aligned} \tag{3.8}$$

Therefore, from (3.2), (3.6)–(3.8), we get (3.5).

Setting $v_n = N^{-1/17}$, from (3.24) of BS (1998) we have

$$m_{\underline{B}_n}(x + iv_n) - m_{F^{c_n, H_n}}(x + iv_n) \to 0 \quad \text{a.s. as } n \to \infty.$$

Since $m_{F^{c_n, H_n}}(x + iv_n) \to m$ as $n \to \infty$, we have

$$m_{\underline{B}_n}(x + iv_n) \to m \quad \text{a.s. as } n \to \infty.$$

When $p > 68/11$, the bound in (3.5) is summable and we conclude

$$|\mu_n(z_n) - m| \to 0 \quad \text{a.s. as } n \to \infty.$$

Therefore

$$\left| r(z_n I - B_n)^{-1} r - \left(1 + \frac{1}{xm}\right) \right| \to 0 \quad \text{a.s. as } n \to \infty. \tag{3.9}$$

Let d_n denote the distance between x and the nearest eigenvalue of B_n. Then, because of Theorem 1.1 there exists a nonrandom $d > 0$ such that, almost surely, $\liminf_n d_n \ge d$.

Write $\widetilde{X} = X_{\bullet N+1}$. Then when $d_n > 0$,

$$\left| r^*(zI - B_n)^{-1} r - r^*(xI - B_n)^{-1} r \right| \le \frac{v_n}{d_n^2} \frac{\widetilde{X}^* \widetilde{X}}{N}. \tag{3.10}$$

Using Lemma 3.1 we have for any $\varepsilon > 0$ and $p = 3$,

$$\mathsf{P}\left(\left| \frac{1}{n} \widetilde{X}^* \widetilde{X} - 1 \right| > \varepsilon \right) \le K_3 \frac{1}{\varepsilon^3} n^{-3/2},$$

which gives us

$$\left| \frac{1}{n} \widetilde{X}^* \widetilde{X} - 1 \right| \to 0 \quad \text{a.s. as } n \to \infty. \tag{3.11}$$

Therefore from (3.9)–(3.11), we get (3.1). $\square$

4. Spread of eigenvalues. In this section we assume the sequence $\{S_n\}$ of Hermitian matrices to be arbitrary except their eigenvalues lie in the fixed interval $[d, e]$. To simplify notation we arrange the eigenvalues of S_n in nondecreasing order, denoting them as $s_1 \le \cdots \le s_n$. Our goal is to prove the following lemma.

LEMMA 4.1. *For any $\varepsilon > 0$ we have for all M sufficiently large,*

$$\limsup_{n\to\infty} \lambda_1^{(1/N)Y_n^* S_n Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_n Y_n} < \varepsilon \quad a.s., \tag{4.1}$$

where Y_n is $n \times [N/M]$ containing i.i.d. elements distributed the same as X_{11} ($[\,\cdot\,]$ denotes the greatest integer function).

PROOF. We verify first a basic inequality.

LEMMA 4.2. *Suppose A and B are $n \times n$ Hermitian. Then*

$$\lambda_1^{A+B} - \lambda_n^{A+B} \le \lambda_1^A - \lambda_n^A + \lambda_1^B - \lambda_n^B.$$

PROOF. Let unit vectors $x, y \in \mathbb{C}^n$ be such that $x^*(A+B)x = \lambda_1^{A+B}$ and $y^*(A+B)y = \lambda_n^{A+B}$. Then

$$\lambda_1^{A+B} - \lambda_n^{A+B} = x^*Ax + x^*Bx - (y^*Ay + y^*By) \le \lambda_1^A + \lambda_1^B - \lambda_n^A - \lambda_n^B. \qquad \square$$

We continue now with the proof of Lemma 4.1. Since each S_n can be written as the difference between two nonnegative Hermitian matrices, because of Lemma 4.2 we may as well assume $d \ge 0$. Choose any positive α so that

$$\frac{e(e-d)}{\alpha} < \frac{\varepsilon}{24c}. \tag{4.2}$$

Choose any positive integer L_1 satisfying

$$\frac{\alpha}{L_1}(1+\sqrt{c})^2 < \frac{\varepsilon}{3}. \tag{4.3}$$

Choose any $M > 1$ so that

$$\frac{Mc}{L_1} > 1 \quad \text{and} \quad 4\sqrt{\frac{cL_1}{M}}\,e < \frac{\varepsilon}{3}. \tag{4.4}$$

Let

$$L_2 = \left[\frac{Mc}{L_1}\right] + 1. \tag{4.5}$$

Assume $n \ge L_1 L_2$. For $k = 1, 2, \dots, L_1$ let $l_k = \{s_{[(k-1)n/L_1]+1}, \dots, s_{[kn/L_1]}\}$ and $\mathscr{L}_1 = \{l_k\colon s_{[kn/L_1]} - s_{[(k-1)n/L_1]+1} \le \alpha/L_1\}$. For any $l_k \notin \mathscr{L}_1$, define for $j = 1, 2, \dots, L_2$ $l_{kj} = \{s_{[(k-1)n/L_1 + (j-1)n/(L_1L_2)]+1}, \dots, s_{[(k-1)n/L_1 + jn/(L_1L_2)]}\}$, and let $\mathscr{L}_2$ be the collection of all the latter sets. Notice the number of elements in $\mathscr{L}_2$ is bounded by $L_1 L_2 (e-d)/\alpha$.

For $l \in \mathscr{L}_1 \cup \mathscr{L}_2$ write

$$S_{n,l} = \sum_{s_i \in l} s_i e_i e_i^* \ (e_i \text{ unit eigenvector of } S_n \text{ corresponding to } s_i),$$

$$A_{n,l} = \textstyle\sum_{s_i \in l} e_i e_i^*,$$

$$\bar{s}_l = \max_i\{s_i \in l\} \quad \text{and} \quad \underline{s}_l = \min_i\{s_i \in l\}.$$

We have

$$\underline{s}_l Y^* A_{n,l} Y \le Y^* S_{n,l} Y \le \bar{s}_l Y^* A_{n,l} Y, \tag{4.6}$$

where "$\le$" denotes partial ordering on Hermitian matrices (i.e., $A \le B \Leftrightarrow B - A$ is nonnegative definite).

Using Lemma 4.2 and (4.6) we have

$$\begin{aligned}
\lambda_1^{(1/N)Y_n^* S_n Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_n Y_n} &\le \sum_l \Big[\lambda_1^{(1/N)Y_n^* S_{n,l} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_{n,l} Y_n}\Big] \\
&\le \sum_l \Big[\bar{s}_l \lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} - \underline{s}_l \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n}\Big] \\
&= \sum_l \bar{s}_l \Big(\lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n}\Big) \\
&\quad + \sum_l (\bar{s}_l - \underline{s}_l) \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n}.
\end{aligned}$$

From (4.5) we have

$$\lim_{n\to\infty} \frac{[n/L_1 L_2]}{[N/M]} = \frac{Mc}{L_1 L_2} < 1. \tag{4.7}$$

Therefore for $l \in \mathscr{L}_2$ we have for all n sufficiently large,

$$\operatorname{rank} A_{n,l} \le \left[\frac{n}{L_1 L_2}\right] + 1 < \left[\frac{N}{M}\right],$$

where we have used the fact that for $a, r > 0$ $[a+r] - [a] = [r]$ or $[r]+1$. This implies $\lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n} = 0$ for all n large. Thus for these n,

$$\begin{aligned}
&\lambda_1^{(1/N)Y_n^* S_n Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_n Y_n} \\
&\quad \le e L_1 \max_{l \in \mathscr{L}_1} \Big(\lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n}\Big) \\
&\qquad + \frac{e(e-d)L_1 L_2}{\alpha} \max_{l \in \mathscr{L}_2} \lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} + \frac{\alpha}{L_1} \lambda_{[N/M]}^{(1/N)Y_n^* Y_n},
\end{aligned}$$

where for the last term we use the fact that for Hermitian C_i, $\sum \lambda_{\min}^{C_i} \le \lambda_{\min}^{\sum C_i}$.

We have with probability 1,

$$\lambda_{[N/M]}^{(1/[N/M])Y_n^* Y_n} \longrightarrow (1 - \sqrt{Mc})^2.$$

Therefore, from (4.3) we have almost surely

$$\lim_{n\to\infty} \frac{\alpha}{L_1} \lambda_{[N/M]}^{(1/N)Y_n^* Y_n} < \frac{\varepsilon}{3}.$$

We have

$$F^{n, A_l} = \left(1 - \frac{|l|}{n}\right) I_{[0,\infty)} + \frac{|l|}{n} I_{[1,\infty)},$$

where $|l|$ is the size of l, and from the expression for the inverse of the Stieltjes transform of the limiting distribution it is a simple matter to show

$$F^{n/[N/M], F^{A_{n,l}}} = F^{|l|/[N/M], I_{[1,\infty)}}.$$

For $l \in \mathscr{L}_1$ we have

$$F^{n, A_l} \to_{\mathscr{D}} \left(1 - \frac{1}{L_1}\right) I_{[0,\infty)} + \frac{1}{L_1} I_{[1,\infty)} \equiv G.$$

From the corollary to Theorem 1.1 of BS (1998), the first inequality in (4.4), and conclusion (i) of Theorem 1.2 we have the extreme eigenvalues of $(1/[N/M])Y_n^* A_{n,l} Y_n$ converging a.s. to the extreme values in the support of $F^{Mc, G}$. Therefore, from Lemma 1.2 we have with probability 1,

$$\lambda_1^{(1/[N/M])Y_n^* A_{n,l} Y_n} - \lambda_{[N/M]}^{(1/[N/M])Y_n^* A_{n,l} Y_n} \longrightarrow 4\sqrt{\frac{Mc}{L_1}},$$

and from the second inequality in (4.4) we have almost surely

$$\lim_{n\to\infty} eL_1 \max_{l\in\mathscr{L}_1} \left(\lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,l} Y_n}\right) < \frac{\varepsilon}{3}.$$

Finally, from (4.7) we see that for $l \in \mathscr{L}_2 \lim_{n\to\infty} |l|/[N/M] < 1$, so that from (4.2), the first inequality in (4.4), and the corollary to Theorem 1.1 of BS (1998) we have with probability 1,

$$\lim_{n\to\infty} \frac{e(e-d)L_1 L_2}{\alpha} \max_{l\in\mathscr{L}_2} \lambda_1^{(1/N)Y_n^* A_{n,l} Y_n} < \frac{e(e-d)}{\alpha} L_1 L_2 \frac{4}{M} < \frac{\varepsilon}{3}.$$

This completes the proof of Lemma 4.1. □

5. Dependence on c.

We now complete the proof of Lemma 1.4. The following relies on Lemma 1.3 and (1.2), the explicit form of $z_{c,H}$.

For (a) we have $(t_1, t_2) \subset S'_H$ with $t_1, t_2 \in \partial S_H$ and $t_1 > 0$. On $(-t_1^{-1}, -t_2^{-1})$, $z_{c,H}(m)$ is well defined, and its derivative is positive if and only if

$$g(m) \equiv \int \left(\frac{tm}{1+tm}\right)^2 dH(t) < \frac{1}{c}.$$

It is easy to verify that $g''(m) > 0$ for all $m \in (-t_1^{-1}, -t_2^{-1})$. Let $\hat{m}$ be the value in $[-t_1^{-1}, -t_2^{-1}]$ where the minimum of $g(m)$ occurs, the two endpoints being included in case $g(m)$ has a finite limit at either value. By considering where the level line $x = 1/c$ crosses the graph of $x = g(m)$, we see that for $c < c_0 \equiv 1/g(\hat{m})$ there are two values, $m_c^1 < m_c^2$, in $[-t_1^{-1}, -t_2^{-1}]$ for which $z'_{c,H}(m) > 0$ for $m \in (-t_1^{-1}, -t_2^{-1}) \Longleftrightarrow m \in (m_c^1, m_c^2)$. Then by Lemma 1.3, $(z_{c,H}(m_c^1), z_{c,H}(m_c^2)) \subset S'_{F^{c,H}}$, with endpoints lying in the boundary of $S_{F^{c,H}}$. From the identity

$$z_{c,H}(m) = \frac{1}{m}(cg(m) - 1) + c\int \frac{t}{(1+tm)^2}\, dH(t), \tag{5.1}$$

we see that $z_{c,H}(m_c^1) > 0$.

As c decreases to zero, we have $m_c^1 \downarrow -t_1^{-1}$, $m_c^2 \uparrow -t_2^{-1}$, which also includes the possibility that either endpoint will reach its limit for positive c [when $g(m)$ has a limit at an endpoint]. We show now (1.5) for $i = 1, 2$. If eventually $m_c^i = -t_i^{-1}$ then clearly (1.5) holds. Otherwise we must have $cg(m_c^i) = 1$, and so by Cauchy–Schwarz,

$$c\left|\int\left(\frac{tm_c^i}{1+tm_c^i}\right)\right| dH(t) \le c^{1/2},$$

and so again (1.5) holds.

It is straightforward to show

$$\frac{dz_{c,H}(m_c^i)}{dc} = \int \frac{t}{1+tm_c^i}\, dH(t). \tag{5.2}$$

Since $(1+tm)(1+tm') > 0$ for $t \in S_H$ and $m, m' \in (-t_1^{-1}, -t_2^{-1})$ we get from (5.2)

$$\frac{d(z_{c,H}(m_c^2) - z_{c,H}(m_c^1))}{dc} = (m_c^1 - m_c^2)\int \frac{t^2}{(1+tm_c^2)(1+tm_c^1)}\, dH(t) < 0.$$

Therefore

$$z_{c,H}(m_c^2) - z_{c,H}(m_c^1) \uparrow t_2 - t_1 \quad \text{as } c \downarrow 0.$$

Upon sliding the line $x = 1/c$ down to the place where $g(m)$ has its minimum, we see that m_c^1 and m_c^2 approach $\hat{m}$ and so the interval $(z_{c,H}(m_c^1), z_{c,H}(m_c^2))$ shrinks to a point as $c \uparrow c_0$. This establishes (a).

We have a similar argument for (b) where now $m_c^3 \in [-1/t_3, 0)$ such that $z'_{c,H}(m) > 0$ for $m \in (-1/t_3, 0) \Longleftrightarrow m \in (m_c^3, 0)$. Since $z_{c,H}(m) \to \infty$ as $m \uparrow 0$ we have $(z_{c,H}(m_c^3), \infty) \subset S'_{F^{c,H}}$ with $z_{c,H}(m_c^3) \in \partial S_{F^{c,H}}$. Equation (5.2) holds also in this case, and from it and the fact that $(1+tm) > 0$ for $t \in S_H$, $m \in (-1/t_3, 0)$, we see that boundary point $z_{c,H}(m_c^3) \downarrow t_3$ as $c \to 0$. On the other hand, $m_c^3 \uparrow 0$ and, consequently, $z_{c,H}(m_c^3) \uparrow \infty$ as $c \uparrow \infty$. Thus we get (b).

When $c[1 - H(0)] < 1$, we can find $m_c^4 \in (-\infty, -1/t_4]$ such that $z'_{c,H}(m) > 0$ on this interval $\Longleftrightarrow m < m_c^4$. Since $z_{c,H}(m) \to 0$ as $m \downarrow -\infty$ we have $(0, z_{c,H}(m_c^4)) \in S'_{F^{c,H}}$ with $z_{c,H}(m_c^4) \in \partial S_{F^{c,H}}$. From (5.2) we have $z_{c,H}(m_c^4) \uparrow t_4$ as $c \downarrow 0$. Since $g(m)$ is increasing on $(-\infty, -1/t_4)$ we have $m_c^4 \downarrow -\infty$, and consequently, $z_{c,H}(m_c^4) \downarrow 0$ as $c \uparrow [1 - H(0)]^{-1}$. Therefore we get (c).

In light of Section 2, all that is missing for (d) is monotonicity and verifying the limits. Formula (5.2) gives us the former. Since $g(m_c) = 1/c$, we see that m_c ranges from 0 to ∞ as c decreases from ∞ to $[1 - H(0)]^{-1}$. Subsequently from (5.1), $z_{c,H}(m_c)$ ranges from ∞ to 0, which completes (d).

(e) is obvious since $z_{c, I_{[0,\infty)}} = -1/m$ for all $m \neq 0$ and so $m_{F^{c, I_{[0,\infty)}}}(z) = -1/z$, the Stieltjes transform of $I_{[0,\infty)}$.

From Lemma 1.3 we can only get intervals in $S'_{F^{c,H}}$ from intervals arising from (a)–(e). The last statement in Lemma 1.4 follows from Theorem 4.4 of Silverstein and Choi (1995). This completes the proof of Lemma 1.4. □

We finish this section with a lemma important to the final steps in the proof of Theorem 1.2.

LEMMA 5.1. *For any $\hat{c} < c$ and sequence $\{\hat{c}_n\}$ converging to $\hat{c}$, the interval $[z_{\hat{c},H}(m_{F^{c,H}}(a)), z_{\hat{c},H}(m_{F^{c,H}}(b))]$ satisfies assumption* (f) *of Theorem* 1.1 (*with c, c_n replaced by $\hat{c}, \hat{c}_n$*). *Moreover, its length increases from $b-a$ as $\hat{c}$ decreases from c.*

PROOF. According to (f), there exists an $\varepsilon > 0$ such that $[a - \varepsilon, b + \varepsilon] \subset S'_{F^{c_n,H_n}}$ for all large n. From Lemma 1.3 we have for these n,

$$\begin{aligned}[m_{F^{c,H}}(a-\varepsilon), m_{F^{c,H}}(b+\varepsilon)] &\subset A_{c_n, H_n} \\ &\equiv \{m \in \mathbb{R}: m \neq 0, -m^{-1} \in S'_{H_n}, z'_{c_n,H_n}(m) > 0\}.\end{aligned}$$

Since $z'_{c,H}(m)$ increases as c decreases, $[m_{F^{c,H}}(a-\varepsilon), m_{F^{c,H}}(b+\varepsilon)]$ is also contained in $A_{\hat{c}_n, H_n}$. Therefore by Lemma 1.3,

$$(z_{\hat{c},H}(m_{F^{c,H}}(a-\varepsilon)), z_{\hat{c},H}(m_{F^{c,H}}(b+\varepsilon))) \subset S'_{F^{\hat{c}_n,H_n}}.$$

Since $z_{\hat{c},H}$ and $m_{F^{c,H}}$ are monotonic on, respectively, $(m_{F^{c,H}}(a-\varepsilon), m_{F^{c,H}}(b+\varepsilon))$ and $(a-\varepsilon, b+\varepsilon)$ we have

$$[z_{\hat{c},H}(m_{F^{c,H}}(a)), z_{\hat{c},H}(m_{F^{c,H}}(b))] \subset (z_{\hat{c},H}(m_{F^{c,H}}(a-\varepsilon)), z_{\hat{c},H}(m_{F^{c,H}}(b+\varepsilon))),$$

so assumption (f) is satisfied.

Since $z'_{\hat{c}',H}(m) > z'_{\hat{c},H}(m) > z'_{c,H}(m)$ for $\hat{c}' < \hat{c}$, we have

$$\begin{aligned}&z_{\hat{c}',H}(m_{F^{c,H}}(b)) - z_{\hat{c}',H}(m_{F^{c,H}}(a)) \\ &\qquad > z_{\hat{c},H}(m_{F^{c,H}}(b)) - z_{\hat{c},H}(m_{F^{c,H}}(a)) \\ &\qquad > z_{c,H}(m_{F^{c,H}}(b)) - z_{c,H}(m_{F^{c,H}}(a)) = b - a.\end{aligned}$$

□

6. Proof of Theorem 1.2(ii). We begin with some basic lemmas. For the following, A is assumed to be $n \times n$ Hermitian, $\lambda \in \mathbb{R}$ is not an eigenvalue of A and Y is any matrix with n rows.

LEMMA 6.1. *λ is an eigenvalue of $A + YY^* \Longleftrightarrow Y^*(\lambda I - A)^{-1}Y$ has eigenvalue* 1.

PROOF. Suppose $x \in \mathbb{C}^n \backslash \{0\}$ is s.t. $(A + YY^*)x = \lambda x$. It follows that $Y^* x \neq 0$ and

$$Y^*(\lambda I - A)^{-1} Y Y^* x = Y^* x$$

so that $Y^*(\lambda I - A)^{-1} Y$ has eigenvalue 1 (with eigenvector $Y^* x$).

Suppose $Y^*(\lambda I - A)^{-1} Y$ has eigenvalue 1 with eigenvector z. Then $(\lambda I - A)^{-1} Y z \neq 0$ and

$$(A + YY^*)(\lambda I - A)^{-1} Y z = -Yz + \lambda(\lambda I - A)^{-1} Y z + Yz = \lambda(\lambda I - A)^{-1} Y z.$$

Thus $A + YY^*$ has eigenvalue λ [with eigenvector $(\lambda I - A)^{-1} Y z$]. □

LEMMA 6.2. *Suppose* $\lambda_j^A < \lambda$. *If* $\lambda_1^{Y^*(\lambda I - A)^{-1} Y} < 1$, *then* $\lambda_j^{A + YY^*} < \lambda$.

PROOF. Suppose $\lambda_j^{A+YY^*} \geq \lambda$. Then since $\lambda_j^{A + \alpha YY^*}$ is continuously increasing in $\alpha \in \mathbb{R}^+$ [Corollary 4.3.3 of Horn and Johnson (1985)] there is an $\alpha \in (0, 1]$ such that $\lambda_j^{A + \alpha YY^*} = \lambda$. Therefore from Lemma 6.1 $\alpha Y^*(\lambda I - A)^{-1} Y$ has eigenvalue 1, which means $Y^*(\lambda I - A)^{-1} Y$ has an eigenvalue ≥ 1. □

LEMMA 6.3. *For any* $i \in \{1, 2, \ldots, n\}$, $\lambda_1^A \leq \lambda_1^A - \lambda_n^A + A_{ii}$.

For the proof, simply use the fact that $A_{ii} \geq \lambda_n^A$.

We now complete the proof of Theorem 1.2(ii). Because of the conditions of (ii) and Lemma 1.4, we may assume $m_{F^{c,H}}(b) < 0$. For $M > 0$ (its size to be determined later) let for $j = 0, 1, 2, \ldots$ $c^j = c/(1 + j/M)$, and define the intervals

$$[a^j, b^j] = [z_{c^j, H}(m_{F^{c,H}}(a)), z_{c^j, H}(m_{F^{c,H}}(b))].$$

By Lemma 5.1 these intervals increase in length as j increases, and for each j the interval, together with c^j, satisfy assumption (f) of Theorem 1.1 for any sequence c_n^j converging to c^j. Here we take

$$c_n^j = \frac{n}{N + j[N/M]}.$$

Let $m_a = m_{F^{c,H}}(a)$. We have

$$a^j - a = z_{c^j, H}(m_a) - z_{c, H}(m_a) = (c^j - c) \int \frac{t}{1 + t m_a} \, dH(t).$$

Therefore, for each j,

$$a^j \leq \hat{a} \equiv a + c \left| \int \frac{t}{1 + t m_a} \, dH(t) \right|.$$

We also have

$$a^{j+1} - a^j = z_{c^{j+1}, H}(m_a) - z_{c^j, H}(m_a) = (c^{j+1} - c^j) \int \frac{t}{1 + t m_a} \, dH(t).$$

Thus we can find an $M_1 > 0$ so that for any $M \geq M_1$ and any j,

$$|a^{j+1} - a^j| < \frac{b-a}{4}. \tag{6.1}$$

Let $M_2 \geq M_1$ be such that for all $M \geq M_2$,

$$\frac{1}{1+1/M} > \frac{3}{4} + \frac{1}{4}\frac{\hat{a}}{b-a+\hat{a}}.$$

This will ensure that for all N, $j \geq 0$, and $M \geq M_2$,

$$\frac{N + j[N/M]}{N+(j+1)[N/M]} b^j > b^j - \frac{(b^j - a^j)}{4}. \tag{6.2}$$

We see from the proof of Lemma 4.1 that the size of M guaranteeing (4.1) depends only on ε and the endpoints d, e of the interval the spectra of S_n are assumed to lie in. Thus we can find an $M_3 \geq M_2$ such that for all $M \geq M_3$, (4.1) is true for any sequence of S_n with

$$d = -\frac{4}{3(b-a)}, \qquad e = \frac{4}{b-a} \quad \text{and} \quad \varepsilon = \frac{1}{\hat{a}|m_a|}.$$

We now fix $M \geq M_3$.

Let for each j,

$$B_n^j = \frac{1}{N + j[N/M]} T_n^{1/2} X_n^{N+j[N/M]} X_n^{N+j[N/M]*} T_n^{1/2},$$

where $X_n^{N+j[N/M]} = (X_{i\,k})$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, N + j[N/M]$.

Since a^j and b^j can be made arbitrarily close to $-1/m_{F^{c,H}(a)}$ and $-1/m_{F^{c,H}(b)}$ respectively, by making j sufficiently large, we can find a K_1 such that for all $K \geq K_1$,

$$\lambda_{i_n+1}^{T_n} < a^K \quad \text{and} \quad b^K < \lambda_{i_n}^{T_n} \quad \text{for all large } n.$$

Therefore, using (1.1) and Lemma 1.2, we can find a $K \geq K_1$ such that with probability 1,

$$\limsup_{n\to\infty} \lambda_{i_n+1}^{B_n^K} < a^K \quad \text{and} \quad b^K < \liminf_{n\to\infty} \lambda_{i_n}^{B_n^K}. \tag{6.3}$$

We fix this K.

Let

$$E_j = \{\text{no eigenvalue of } B_n^j \text{ appears in } [a^j, b^j] \text{ for all large } n\}.$$

Let

$$l_n^j = \begin{cases} k, & \text{if } \lambda_k^{B_n^j} > b^j,\ \lambda_{k+1}^{B_n^j} < a^j, \\ -1, & \text{if there is an eigenvalue of } B_n^j \text{ in } [a^j, b^j]. \end{cases}$$

For notational convenience, let $\lambda_{-1}^A = \infty$ for Hermitian A.

Define

$$\hat{a}^j = a^j + \tfrac{1}{4}(b^j - a^j),$$
$$\hat{b}^j = b^j - \tfrac{1}{4}(b^j - a^j).$$

Fix $j \in \{0, 1, \dots, K-1\}$. On the same probability space we define for each n large $Y_n = (Y_{ik})$, $i = 1, 2, \dots, n$, $k = 1, \dots, [N/M]$, entries i.i.d., distributed the same as X_{11}, with $\{B_n^j\}_n$ and $\{Y_n\}_n$ independent (no restriction on Y_n for different n). Let $R_n = T_n^{1/2} Y_n$.

Whenever $\hat{a}^j$ is not an eigenvalue of B_n^j, we have by Lemma 6.3,

$$\begin{aligned} &\lambda_1^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} \\ &\quad \le \lambda_1^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} \\ &\qquad - \lambda_{[N/M]}^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} \\ &\qquad + \left(\frac{1}{N + j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n \right)_{11}. \end{aligned} \tag{6.4}$$

From Lemma 3.3 we have

$$\begin{aligned} &\left(\frac{1}{N + j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n \right)_{11} \\ &\quad \to 1 + \frac{1}{\hat{a}^j m_{F^{c^j,H}}(\hat{a}^j)} < 1 + \frac{1}{\hat{a} m_a} \quad \text{a.s. as } n \to \infty. \end{aligned} \tag{6.5}$$

From Lemma 4.1

$$\begin{aligned} &\limsup_{n\to\infty} \lambda_1^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} \\ &\quad - \lambda_{[N/M]}^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} < \frac{1}{\hat{a}|m_a|} \quad \text{a.s.} \end{aligned} \tag{6.6}$$

holds for a fixed realization in E_j with respect to the probability measure on $\{Y_n\}_n$. By Fubini's theorem we subsequently have (6.6) on the probability space generating $\{B_n^j\}_n$ and $\{Y_n\}_n$. Therefore, from (6.4)–(6.6) we find

$$\mathsf{P}\Big(\lambda_1^{(1/(N+j[N/M]))R_n^*(\hat{a}^j I - B_n^j)^{-1} R_n} < 1 \text{ for all large } n \Big) = 1,$$

and since $B_n^j + (1/(N + j[N/M]))R_n R_n^* \sim ((N + (j+1)[N/M])/(N + j[N/M]))B_n^{j+1}$ we get from Lemma 6.2 and the fact that $\mathsf{P}(E_j) = 1$ (from Theorem 1.1), with probability 1,

$$\lambda_{l_n^j+1}^{B_n^{j+1}} < \hat{a}^j \quad \text{for all large } n.$$

Since $\lambda_{l_n^j}^{B_n^j} \le \lambda_{l_n^j}^{B_n^j + (1/(N+j[N/M]))R_n R_n^*}$, we use (6.2) to get

$$\mathsf{P}\Big(\lambda_{l_n^j}^{B_n^{j+1}} > \hat{b}^j \text{ and } \lambda_{l_n^j+1}^{B_n^{j+1}} < \hat{a}^j \text{ for all large } n \Big) = 1.$$

From (6.1) we see that $[\hat{a}^j, \hat{b}^j] \subset [a^{j+1}, b^{j+1}]$. Therefore, combining the above event with E_{j+1}, we conclude

$$\mathsf{P}\Big(\lambda_{l_n^j}^{B_n^{j+1}} > b^{j+1} \text{ and } \lambda_{l_n^j+1}^{B_n^{j+1}} < a^{j+1} \text{ for all large } n\Big) = 1.$$

Therefore, with probability 1, for all n large $[a, b]$ and $[a^K, b^K]$, split the eigenvalues of, respectively, B_n and B_n^K having equal amounts to the left sides of the intervals. Finally, from (6.3) we get (ii). □

Acknowledgment. Part of this work was done while J. W. Silverstein visited the Department of Mathematics at National University of Singapore. He thanks the members of the department for their hospitality.

REFERENCES

BAI, Z. D. and SILVERSTEIN, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345.

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.

FAN, K. (1951). Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proc. Nat. Acad. Sci. U.S.A.* **37** 760–766.

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

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

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

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 Processing* **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