

# Exact Separation of Eigenvalues of Large Dimensional Sample Covariance Matrices

By Z.D. Bai<sup>1</sup> and Jack W. Silverstein<sup>2</sup>

National University of Singapore and North Carolina State University

Dedicated to the Eightieth Birthday of C.R. Rao

---

<sup>1</sup> Research supported in part by Research Grant RP 3972712 from the National University of Singapore

<sup>2</sup> Research supported by the National Science Foundation under grants DMS-9404047 and DMS-9703591

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

Key Words and Phrases. Random matrix, empirical distribution function of eigenvalues, Stieltjes transform.

## Summary

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 is shown in Bai and Silverstein (1998) that, under certain conditions on the eigenvalues of  $T_n$ , with probability one 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 one 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) 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 ( $\mathbf{E}X_{11} = 0$ ,  $\mathbf{E}|X_{11}|^2 = 1$ ),  $T_n$  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  $n/N \rightarrow c > 0$  as  $n \rightarrow \infty$  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  $F^{T_n} \rightarrow_D H$ , a cumulative distribution function (c.d.f.). It follows [Silverstein (1995)] that with probability one,  $F^{B_n} \rightarrow_D F$ , 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 Bai and Silverstein (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 matrix of  $N$  samples of the random vector  $T_n^{1/2}X_{\cdot 1}$  ( $X_{\cdot j}$  denoting the  $j^{\text{th}}$  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 proven 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 1.1 [Fan (1951)]. *For rectangular matrix  $A$  and positive integer  $i \leq \text{rank } A$ , let  $\lambda_i^A$  denote the  $i^{\text{th}}$  largest singular value of  $A$ . Define  $\lambda_i^A$  to be zero for all  $i > \text{rank } A$ . Let  $m, n$  be arbitrary non-negative integers.*

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

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

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)

$$(1.1) \quad \lambda_{i_n+1}^{B_n} \leq \lambda_1^{(1/N)X_n X_n^*} \lambda_{i_n+1}^{T_n} \quad \text{and} \quad \lambda_{i_n}^{B_n} \geq \lambda_n^{(1/N)X_n X_n^*} \lambda_{i_n}^{T_n}.$$

It is well-known [dating back to Marčenko and Pastur (1967)] that the limiting spectral distribution of  $(1/N)X_n X_n^*$  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

LEMMA 1.2 [Yin, Bai, and Krishnaiah (1988), Bai and Yin (1993)]. *With probability one, the largest eigenvalue of  $(1/N)X_n X_n^*$  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 to 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 due to the fact that 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  $\hat{R}$  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  $\hat{R}$  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  $F^{\hat{R}}$  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 Bai and Silverstein (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  $\hat{R}$  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  $\underline{B}_n \equiv (1/N)X_n^*T_nX_n$  and properties of its Stieltjes transform  $m_{\underline{B}_n}$ , which for any c.d.f.  $G$  is defined by

$$m_G(z) \equiv \int \frac{1}{\lambda - z} dG(\lambda) \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \text{Im } z > 0\}.$$

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

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

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

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

The main result in Bai and Silverstein (1998) can now be stated.

**THEOREM 1.1** [Theorem 1.1 of Bai and Silverstein (1998)]. *Assume:*

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

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

(c) For each  $n$   $T_n$  is an  $n \times n$  Hermitian nonnegative definite satisfying  $H_n \equiv F^{T_n} \rightarrow_D H$ , 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<sup>†</sup> outside the support of  $F^{c_n, H_n}$  for all large  $n$ .

---

<sup>†</sup> 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.

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

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

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

[see Bai and Silverstein (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)  $m \in \mathbb{R} \setminus \{0\}$ , (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 Fig. 1 of Bai and Silverstein (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

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

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  $i_n \geq 0$  be such that

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

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  $c < c_0 \implies$  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,

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

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  $\infty$  to 0, (1.5) holds for  $i = 3$  with convergence monotone from  $\infty$  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  $\infty$  to 0 as  $c$  decreases from  $\infty$  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.

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

(2) If  $c[1 - H(0)] \leq 1$ , 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

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

Conclusion (1) 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 transferred (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 non-connected 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 (2) 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 (1).** We see that  $x_0$  must coincide with the boundary point in (d) of Lemma 1.4. Most of (d) will be proven 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  $n \rightarrow \infty$ .*

*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  $\limsup_n H_n(0) \leq H(0)$  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  $m$   $z_{c_n, H_n}(m) \rightarrow z_{c, H}(m)$  and  $z'_{c_n, H_n}(m) \rightarrow z'_{c, H}(m)$  as  $n \rightarrow \infty$ , 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,  $\hat{m}_n \rightarrow \hat{m}$  and, in turn,  $z_{c_n, H_n}(\hat{m}_n) \rightarrow x_0$  as  $n \rightarrow \infty$ .  $\square$

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

$$(2.1) \quad \lambda_N^{B_n} \xrightarrow{a.s.} x_0 \quad \text{as } n \rightarrow \infty.$$

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

$$\lambda_N^{\frac{1}{N} X_n X_n^*} \leq \lambda_N^{B_n} \lambda_1^{T_n^{-1}} = \lambda_N^{B_n} (\lambda_n^{T_n})^{-1}.$$

Since by Lemma 1.2  $\lambda_N^{\frac{1}{N} X_n X_n^*} \xrightarrow{a.s.} (1 - \sqrt{c})^2$  as  $n \rightarrow \infty$  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

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

But since  $F^{B_n} \rightarrow_D F$  a.s. we must have

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

Thus we get (2.1).

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

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

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

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

$$(2.3). \quad |\lambda_N^{B_n^\epsilon} - \lambda_N^{B_n}| = |\lambda_N^{\frac{1}{N} X_n^* T_n^\epsilon X_n} - \lambda_N^{\frac{1}{N} X_n^* T_n X_n}| \leq \|\frac{1}{N} X_n^* (T_n^\epsilon - T_n) X_n\| \leq \|\frac{1}{N} X_n X_n^*\| \epsilon$$

Since  $H^\epsilon \rightarrow_D H$  as  $\epsilon \rightarrow 0$  we get from Lemma 2.1

$$(2.4) \quad z_{c, H^\epsilon}(\hat{m}^\epsilon) \rightarrow z_{c, H}(\hat{m}) \quad \text{as } \epsilon \rightarrow 0.$$

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

The proof of (2) 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  $Y_{ij} = X_{ij}I_{[|X_{ij}| \leq C]} - \mathbf{E}X_{ij}I_{[|X_{ij}| \leq C]}$ ,  $Y_n = (Y_{ij})$  and  $\tilde{B}_n = (1/N)T^{1/2}Y_nY_n^*T^{1/2}$ . It is shown in Bai and Silverstein (1998) that with probability one

$$\limsup_{n \rightarrow \infty} \max_{k \leq n} |(\lambda_k^{B_n})^{1/2} - (\tilde{\lambda}_k^{\tilde{B}_n})^{1/2}| \leq (1 + \sqrt{C})\mathbf{E}^{1/2}|X_{11}|^2 I_{[|X_{11}| > C]}.$$

It is clear that we can make the above bound arbitrarily small by choosing  $C$  sufficiently large. Thus, in proving (2) 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 Bai and Silverstein (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  $n$  (that is, 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  $\|T_n\| \leq 1$ .

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

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

$$\mathbf{E}|X_{\cdot 1}^* C X_{\cdot 1} - \text{tr } C|^p \leq K_p (\text{tr } C C^*)^{p/2}$$

**LEMMA 3.2** [Lemma 2.6 of Silverstein and Bai (1995)]. *Let  $z \in \mathbb{C}^+$  with  $v = \text{Im } z$ ,  $A$  and  $B$   $n \times n$  with  $B$  Hermitian, and  $q \in \mathbb{C}^n$ . Then*

$$|\text{tr}((B - zI)^{-1} - (B + \tau q q^* - zI)^{-1})A| \leq \frac{\|A\|}{v}.$$

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  $\tilde{X} \in \mathbb{C}^n$  be distributed the same as  $X_{\cdot 1}$  and independent of  $X_n$ . Set  $r = r_n = (1/\sqrt{N})T_n^{1/2}\tilde{X}$ . Then*

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

*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

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

Therefore, using Lemma 3.2 we have

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

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

$$(3.2) \quad |m_{\underline{B}_n}(z) - m_{\underline{B}_n^{N+1}}(z)| \leq \frac{(2c_n + 1)}{v(N+1)}.$$

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

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

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

which implies

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

Thus we have

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

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

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

$$(3.5) \quad \mathbb{P}\left(|m_{\underline{B}_n}(z) - \mu_n(z)| > \epsilon\right) \leq K_p \left(\frac{|z|}{\epsilon v_n^3}\right)^p \frac{n^{p/2}}{N^{p-1}}.$$

We have from (3.4)

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

Using (3.3) we find

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

Write

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

Using Lemma 3.2 we find

$$(3.7) \quad \frac{1}{N} |\text{tr} ((B_n - zI)^{-1} - (B_{(j)} - zI)^{-1}) T_n| \leq \frac{2}{N v_n}.$$

Using Lemma 3.1 we have for any  $j \leq N+1$  and  $p \geq 2$

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

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

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

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

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

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

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

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

Therefore

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

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,  $\liminf_n d_n \geq d$ .

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

$$(3.10) \quad |r^*(zI - B_n)^{-1} r - r^*(xI - B_n)^{-1} r| \leq \frac{v_n}{d_n^2} \frac{\tilde{X}^* \tilde{X}}{N}.$$

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

$$\mathbb{P}(|\frac{1}{n} \tilde{X}^* \tilde{X} - 1| > \epsilon) \leq K_3 \frac{1}{\epsilon^3} n^{-3/2}$$

which gives us

$$(3.11) \quad |\frac{1}{n} \tilde{X}^* \tilde{X} - 1| \xrightarrow{a.s.} 0 \quad \text{as } n \rightarrow \infty.$$

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  $s_1 \leq \dots \leq s_n$ . Our goal is to prove

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

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

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

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

*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

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

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

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

Choose any positive integer  $L_1$  satisfying

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

Choose any  $M > 1$  so that

$$(4.4) \quad \frac{Mc}{L_1} > 1 \quad \text{and} \quad 4\sqrt{\frac{cL_1}{M}}e < \frac{\epsilon}{3}.$$

Let

$$(4.5) \quad L_2 = \left\lceil \frac{Mc}{L_1} \right\rceil + 1.$$

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

For  $\ell \in \mathcal{L}_1 \cup \mathcal{L}_2$  write

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

$$\bar{s}_\ell = \max_i \{s_i \in \ell\}, \quad \text{and} \quad \underline{s}_\ell = \min_i \{s_i \in \ell\}.$$

We have

$$(4.6) \quad \underline{s}_\ell Y^* A_{n,\ell} Y \leq Y^* S_{n,\ell} Y \leq \bar{s}_\ell Y^* A_{n,\ell} Y,$$

where “ $\leq$ ” denotes partial ordering on Hermitian matrices (that is,  $A \leq B \iff B - A$  is non-negative definite).

$$\begin{aligned} & \text{Using Lemma 4.2 and (4.6) we have } \lambda_1^{(1/N)Y_n^* S_n Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_n Y_n} \\ & \leq \sum_{\ell} [\lambda_1^{(1/N)Y_n^* S_{n,\ell} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* S_{n,\ell} Y_n}] \leq \sum_{\ell} [\bar{s}_\ell \lambda_1^{(1/N)Y_n^* A_{n,\ell} Y_n} - \underline{s}_\ell \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,\ell} Y_n}] \\ & = \sum_{\ell} \bar{s}_\ell (\lambda_1^{(1/N)Y_n^* A_{n,\ell} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,\ell} Y_n}) + \sum_{\ell} (\bar{s}_\ell - \underline{s}_\ell) \lambda_{[N/M]}^{(1/N)Y_n^* A_{n,\ell} Y_n}. \end{aligned}$$

From (4.5) we have

$$(4.7) \quad \lim_{n \rightarrow \infty} \frac{\left\lfloor \frac{n}{L_1 L_2} \right\rfloor}{\left\lfloor \frac{N}{M} \right\rfloor} = \frac{Mc}{L_1 L_2} < 1.$$

Therefore for  $\ell \in \mathcal{L}_2$  we have for all  $n$  sufficiently large

$$\text{rank } A_{n,\ell} \leq \left\lfloor \frac{n}{L_1 L_2} \right\rfloor + 1 < \left\lfloor \frac{N}{M} \right\rfloor,$$

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

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

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

We have with probability one

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

Therefore, from (4.3) we have almost surely

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

We have

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

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

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

For  $\ell \in \mathcal{L}_1$  we have

$$F^{n, A_\ell} \rightarrow_D \left(1 - \frac{1}{L_1}\right) I_{[0, \infty)} + \frac{1}{L_1} I_{[1, \infty)} \equiv G.$$

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

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

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

$$\lim_{n \rightarrow \infty} eL_1 \max_{\ell \in \mathcal{L}_1} (\lambda_1^{(1/N)Y_n^* A_{n, \ell} Y_n} - \lambda_{[N/M]}^{(1/N)Y_n^* A_{n, \ell} Y_n}) < \frac{\epsilon}{3}.$$

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

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

This completes the proof of Lemma 4.1.  $\square$

**5. Dependence on  $c$ .** We now finish 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  $m \in (-t_1^{-1}, -t_2^{-1}) \iff m \in (m_c^1, m_c^2)$ . 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

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

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

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

and so again (1.5) holds.

It is straightforward to show

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

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  $m \in (-1/t_3, 0) \iff m \in (m_c^3, 0)$ . Since  $z_{c,H}(m) \rightarrow \infty$  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  $c \rightarrow 0$ . 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  $\iff m < m_c^4$ . Since  $z_{c,H}(m) \rightarrow 0$  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  $\infty$  as  $c$  decreases from  $\infty$  to  $[1 - H(0)]^{-1}$ . Subsequently from (5.1)  $z_{c,H}(m_c)$  ranges from  $\infty$  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 follow from Theorem 4.4 of Silverstein and Choi (1995). This completes the proof of Lemma 1.4.  $\square$

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  $\epsilon > 0$  such that  $[a - \epsilon, b + \epsilon] \subset S'_{F^{c_n}, H_n}$  for all large  $n$ . From Lemma 1.3 we have for these  $n$

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

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

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

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

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

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

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

**6. Proof of (2).** 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.  $\lambda$  is an eigenvalue of  $A + YY^* \iff Y^*(\lambda I - A)^{-1}Y$  has eigenvalue 1.

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

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

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  $(\lambda I - A)^{-1}Yz \neq 0$  and

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

Thus  $A + YY^*$  has eigenvalue  $\lambda$  (with eigenvector  $(\lambda I - A)^{-1}Yz$ ).  $\square$

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  $\geq 1$ .  $\square$

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

*Proof.* Simply use the fact that  $A_{ii} \geq \lambda_n^A$ .  $\square$

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

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

Therefore, for each  $j$

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

We also have

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

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

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

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  $M \geq M_2$

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

We see from the proof of Lemma 4.1 that the size of  $M$  guaranteeing (4.1) depends only on  $\epsilon$  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

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

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

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 one

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

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

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

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

Define

$$\begin{aligned} \hat{a}^j &= a^j + \frac{1}{4}(b^j - a^j) \\ \hat{b}^j &= b^j - \frac{1}{4}(b^j - a^j). \end{aligned}$$

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,

$$(6.4) \quad \lambda_1^{\frac{1}{N+j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n} \leq \lambda_1^{\frac{1}{N+j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n} - \lambda_{[N/M]}^{\frac{1}{N+j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n} + (\frac{1}{N+j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n)_{11}.$$

From Lemma 3.3 we have

$$(6.5) \quad (\frac{1}{N+j[N/M]} R_n^* (\hat{a}^j I - B_n^j)^{-1} R_n)_{11} \xrightarrow{a.s.} 1 + \frac{1}{\hat{a}^j m_{F^{cj}, H}(\hat{a}^j)} < 1 + \frac{1}{\hat{a} m_a} \quad \text{as } n \rightarrow \infty.$$

From Lemma 4.1

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

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

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

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

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

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

$$\mathbb{P}(\lambda_{\ell_n^j}^{B_n^{j+1}} > \hat{b}^j \quad \text{and} \quad \lambda_{\ell_{n+1}^j}^{B_n^{j+1}} < \hat{a}^j \quad \text{for all large } n) = 1.$$

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

$$\mathbb{P}(\lambda_{\ell_n^j}^{B_n^{j+1}} > b^{j+1} \quad \text{and} \quad \lambda_{\ell_{n+1}^j}^{B_n^{j+1}} < a^{j+1} \quad \text{for all large } n) = 1.$$

Therefore, with probability one, 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 (2).  $\square$

**Acknowledgments.** 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 empirical 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