No Eigenvalues Outside the Support of Limiting Empirical Spectral Distribution of a Separable Covariance Matrix

Debashis Paul and Jack W. Silverstein

Technical Report #2007-6
July 19, 2007

This material was based upon work supported by the National Science Foundation under Agreement No. DMS-0112069. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science

Statistical and Applied Mathematical Sciences Institute
PO Box 14006
Research Triangle Park, NC 27709-4006
www.samsi.info

No eigenvalues outside the support of limiting empirical spectral distribution of a separable covariance matrix

Debashis Paul[†1] and Jack W. Silverstein[‡2]

†*Department of Statistics, University of California, Davis, CA 95616, USA;*
Email : debashis@wald.ucdavis.edu; Fax : (+01) 530 752 7099

‡*Department of Mathematics, Box 8205, North Carolina State University, Raleigh, NC 27695-8205, USA; and SAMSI, Research Triangle Park, NC 27709-4006, USA;*
Email : jack@math.ncsu.edu; Fax : (+01) 919 513 7336

July 19, 2007

Abstract

We consider a class of matrices of the form $C_n = (1/N)A_n^{1/2}X_nB_nX_n^*A_n^{1/2}$, where X_n is an $n \times N$ matrix consisting of i.i.d. standardized complex entries, $A_n^{1/2}$ is a non-negative definite Hermitian square-root of the non-negative definite matrix A_n, and B_n is diagonal with non-negative diagonal entries. Under the assumption that the distribution of the eigenvalues of A_n and B_n converge to proper probability distributions, as $\frac{n}{N} \to c \in (0,\infty)$, the empirical spectral distribution of C_n converges a.s. to a non-random limit. We show that, under appropriate conditions on the eigenvalues of A_n and B_n, with probability 1, there will be no eigenvalues in any closed interval outside the support of the limiting distribution, for sufficiently large n. The problem is motivated by applications in spatio-temporal statistics and wireless communications.

Keywords : Empirical spectral distribution, Stieltjes transform, separable covariance, CDMA, MIMO.

[1]Paul was supported by the National Science Foundation under Agreement No. DMS-0112069, as part of a research fellowship from SAMSI.

[2]Silverstein's work is supported by U.S. Army Research Office under Grant W911NF-05-1-0244
This material was based upon work supported by the National Science Foundation under Agreement No. DMS-0112069. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

The aim of this paper is to extend the result in Bai and Silverstein (1998) to the eigenvalues of a more general class of random matrices, specifically matrices of the form

$$C_n = (1/N)A_n^{1/2}X_nB_nX_n^*A_n^{1/2},$$

where for $n = 1, 2, \ldots$, X_n is $n \times N$ ($N = N(n)$) consisting of i.i.d. standardized complex entries ($\mathbb{E}X_{11} = 0$, $\mathbb{E}|X_{11}|^2 = 1$), $A_n^{1/2}$ is a nonnegative definite square root of the $n \times n$ Hermitian nonnegative definite matrix A_n, and $B_n = \text{diag}(b_1, b_2, \ldots, b_N)$, each $b_i \geq 0$. The matrices studied in Bai and Silverstein (1998) assume $B_n = I_N$, the $N \times N$ identity matrix. In that case C_n can be viewed as the sample covariance matrix consisting of N samples of the random vector $A_n^{1/2}X_{\cdot 1}$ ($X_{\cdot 1}$ denoting the first column of X_n), which has population covariance matrix A_n. The matrix C_n can then be interpreted as the sample covariance matrix consisting of N weighted samples. There are other ways to interpret the matrix, being important in various applications. One example is the spatio-temporal sampling model to be described in Section 1.2.1. In wireless communications, $H_n = (1/\sqrt{N})A_n^{1/2}X_nB_n^{1/2}$, for general nonnegative definite matrix B_n, is used to model the path gains between different groups of antennas in a multiple-input-multiple-output (MIMO) system (Section 1.2.2). It is typically assumed that X_{11} is complex Gaussian (real and imaginary parts independently distributed as $N(0, 1/2)$), in which case the square of the singular values of H_n has the same distribution as the eigenvalues of C_n (the b_i's being the eigenvalues of B_n).

1.1 Statement of the result

Results have previously been obtained on the limiting behavior of the empirical distribution function, F^{C_n}, of its eigenvalues ($F^{C_n}(x) \equiv$ (number of eigenvalues of $C_n \leq x)/n$), (Burda (2005), Zhang (2006), de Mondvel, Khorunzy and Vasilchuck (1996)), with differing assumptions (the weakest appearing in Zhang (2006)) and varied (but equivalent) forms of expressions for the result. The following limit result is expressed in terms of the Stieltjes transform of F^{C_n}, defined for any distribution function G as

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

Assume that the empirical distribution functions, F^{A_n} and F^{B_n}, converge weakly, as $n \to \infty$, to probability distribution functions, denoted respectively by F^A and F^B, and $c_n \equiv n/N \to c > 0$. Then, with probability 1, F^{C_n} converges weakly to a probability distribution function F whose Stieltjes transform $m(z) = m_F(z)$, for $z \in \mathbb{C}^+$, is given by

$$m(z) = \int \frac{1}{a \int \frac{b}{1+cbe} dF^B(b) - z} dF^A(a), \tag{1}$$

where $e = e(z)$ is the unique solution in $\mathbb{C}^+$ of the equation

$$e = \int \frac{a}{a \int \frac{b}{1+cbe} dF^B(b) - z} dF^A(a). \tag{2}$$

As in Bai and Silverstein (1998), the purpose of this paper is to prove, with additional assumptions, the almost sure non-appearance of eigenvalues of C_n in any interval away from the origin and outside the support of F as $n \to \infty$. Before the result can be formally stated we need one more definition. Let F^{c_n,A_n,B_n} denote the distribution function defined by (1) and (2), that is, in these two expressions replace c, F^A, F^B with c_n, F^{A_n}, F^{B_n}. Then F^{c_n,A_n,B_n} is simply the distribution function with Stieltjes transform given by (1).

The following will be proven:

THEOREM 1: *Assume the following:*

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

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

(c) *For each* n, A_n *is* $n \times n$ *Hermitian nonnegative definite, and* $B_n = diag(b_1, \ldots, b_N)$ *is* $N \times N$, *each* $b_i \geq 0$, *satisfying* $F^{A_n} \xrightarrow{\mathcal{D}} F^A$, $F^{B_n} \xrightarrow{\mathcal{D}} F^B$, *both limits being probability distribution functions.*

(d) $\|A_n\|$ *and* $\|B_n\|$, *the respective spectral norms of* A_n B_n, *are bounded in* n.

(e) $C_n = (1/N)A_n^{1/2}X_nB_nX_n^*A_n^{1/2}$, *where* $A_n^{1/2}$ *any Hermitian square root of* A_n, $X_n = (X_{ij})$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, N$.

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

Then,

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

The applicability of Theorem 1 depends on finding a way to determine the intervals outside the support of F^{c_n,A_n,B_n}, as it exists for sample covariance matrices (Silverstein and Choi (1995)). In the latter case, the limiting Stieltjes transform $m(z)$ has an explicit inverse $z = z(m)$. It is straightforward to verify that a Stieltjes transform is increasing on intervals on the real line outside the support of its distribution function. Its inverse therefore exists on these intervals and is also increasing. Therefore plotting $z(m)$ for m real, and locating on the vertical axis places where the inverse is increasing, yield intervals outside the support. There does not appear to be an explicit inverse for (1). Nevertheless, preliminary work indicates a way to determine an inverse of $m(z)$ associated with an interval outside the support of the limiting spectral distribution. This has been established in the case of another class of random matrices (Dozier and Silverstein (2007)). Work in this area is currently being pursued.

1.2 Motivation

Our results give information on the behavior of the limiting empirical spectral distribution and also on the behavior of individual eigenvalues. Results describing only the limiting behavior of the empirical spectral distribution provide information on the proportion of eigenvalues falling in any interval. But these results do not rule out the possibility of $o(n)$ eigenvalues scattered outside the support of the limiting empirical spectral distribution. The goal of our research is to establish that such a phenomenon does not occur for large enough n. Further research in our framework would allow for precise description of the location of the eigenvalues. In particular, we expect that the results proved here will be key to proving certain *phase transition phenomena* observed in the context of sample covariance matrices with $B_n = I_N$ and A_n having a few large isolated eigenvalues (Baik and Silverstein (2006), Baik, Ben Arous and Péché (2006), El Karoui (2006), Paul (2007)).

1.2.1 Application to spatio-temporal statistics

The data model we are considering here arises in the field of spatio-temporal statistics, where the rows of the $n \times N$ matrix $U_n = A_n^{1/2} X_n B_n^{1/2}$ correspond to indices of spatial locations and the column indices correspond to points in time. This class of models is also known as the *separable covariance model.* This is because, under the assumptions made here on the entries of X_n (i.i.d., mean 0, finite fourth moment), the joint (space-time) covariance of U_n, viewed as an $Nn \times 1$ vector consisting of the columns of the matrix U_n stacked on top of one another, is given by $\Sigma_U = A_n \otimes B_n$, where $\otimes$ denotes the Kronecker product between matrices. Note that, if we further assume Gaussianity for the entries of X, then the joint distribution of U_n is $N_{Nn}(0, A_n \otimes B_n)$. Also, in that setting, we do not require A_n and B_n to be diagonal, but only that they are nonnegative definite. The interpretation of this covariance structure is that the entries of U_n are correlated in time (column), but the pattern of temporal correlation does not vary with location (row). In other words, there is no space-time interaction in the process.

One advantage of this model from a statistical estimation point of view is that, when N is large and n is comparatively small, so that $\frac{n}{N} \to 0$ as $n \to \infty$, it is possible to get quite reliable estimate of A_n from the sample covariance matrix $C_n = \frac{1}{N} U_n U_n^*$. Indeed, in that setting, if moreover $\| A_n \|$ is bounded above, it is not hard to verify that $\| C_n - \frac{1}{N}(\operatorname{tr} B_n) A_n \| \to 0$ a.s., as $n \to \infty$. So, the spectral properties of A_n can be recovered from that of the spectrum of C_n. Of course, the key questions we are addressing here relate to the situation where $\frac{n}{N} \to c \in (0, \infty)$. The behavior of the empirical spectrum in that setting is hitherto unknown.

The results and techniques presented in this paper may prove useful in this problem for a number of different reasons. A statistical problem related to such spatio-temporal processes is to understand the temporal variability of the spatial field. One of the approaches for understanding the temporal variability is to perform an eigen-analysis (in space) of the sample covariance matrix C_n. This is because, the weights of the different eigenvectors of C_n, in representing the columns of U_n (principal components scores), vary in time. These weights therefore capture the temporal variability of the orthogonal components (eigenvectors of C_n) of the spatial process. The eigenfunctions thus obtained are usually referred to as empirical orthogonal functions (particularly in climatology, see, e.g. von Storch and Zwiers (1999)). Understanding the asymptotic behavior of the sample eigenvalues and eigenfunctions therefore is a relevant question, since, under the separable space-time model they give a set of orthogonal components, and their relative strengths, of the spatial variation of the process.

1.2.2 Application to wireless communication

In wireless communications, $H_n = (1/\sqrt{N}) A_n^{1/2} X_n B_n^{1/2}$, for a general nonnegative definite matrix B_n, appears in a variety of models, including both direct-sequence and multiple-carrier code-division multiple-access systems (Tulino and Verdú (2004), sections 3.1-3.2), and in multiple-input-multiple-output (MIMO) systems (Tulino and Verdú (2004), section 3.3). The importance of acquiring more detailed information on the singular values of H_n beyond what the limiting empirical distribution ((1), (2)) reveals, which has been primarily used to estimate capacity, is becoming more apparent.

For example, in Verdú (2002) an estimate of capacity requires knowledge of the largest singular value of H_n, which Theorem 1 provides (the corollary to Theorem 1.1 in Bai and Silverstein (1998) readily follows from Theorem 1). Another example is in MIMO systems, where H_n models the *path gains* between different groups of antennas. It is typically assumed that X_{11} is complex Gaussian (real and imaginary parts independent $N(0,1/2)$), in which case the square of the singular values of H_n has the same distribution as the eigenvalues of C_n (the b_i's being the eigenvalues of B_n). The matrices A_n and B_n are the covariances between the receiver and the transmitter antennas, respectively. They reflect the scenario involving these two groups of antennas, for example, their locations, and the nature of the interference encountered due to their surroundings. The singular values of H_n, or equivalently the eigenvalues of C_n, indicate several important properties of the communication scheme, due to the fact that any information on H_n yields ways to allocate the transmitted signal in an optimal way. For example, if there is a significant number of small eigenvalues, transmission can be achieved after performing a unitary transformation, on the left and/or the right side of H_n, resulting in a reduced number of virtual parallel antennas with little correlation between them. When the number of antennas is sizeable, knowledge of the eigenvalues of C_n, depending only on A_n and B_n, is gained to some extent from the limiting F. It yields the proper proportion of eigenvalues within any interval. However, Theorem 1 is a step toward knowing the location of all the singular values, which provides much more information. For example, it can ensure that no lone eigenvalue above or below the limiting support exists. The importance of the Theorem 1 lies in the determination of spectral behavior of C_n entirely through A_n and B_n.

The essential portion of the proof of Theorem 1 will proceed in the following sections. The results to be obtained here are analogous to those in sections 3-5 of Bai and Silverstein (1998), namely, we will show

$$\sup_{x\in[a,b]} |m_n(z) - m_n^0(z)| = o(1/(nv_n)) \quad \text{a.s.}, \tag{3}$$

where

$$m_n = m_n(z) = m_{F^{C_n}} = (1/n)\text{tr}\,(C_n - zI)^{-1} \tag{4}$$

is the Stieltjes transform of the empirical distribution function of the eigenvalues of C_n,

$$m_n^0 = m_n^0(z) = m_{F^{c_n,A_n,B_n}}(z) \tag{5}$$

and $z = x + iv_n$, where $v_n = \kappa n^{-1/140}$, κ an arbitrary positive constant (fixed for all n).

The steps needed to conclude Theorem 1 from (3) are identical to those in section 6 of Bai and Silverstein (1998), except for the fact that in the latter paper $v_n = N^{-1/68}$. The reader is referred to that section for the details.

Before proceeding, we simplify here some of the assumptions. It is clear from assumption (d) of Theorem 1 that, we can assume that both $\|A_n\|$ and $\|B_n\|$ are bounded by 1 for all n. Also, the argument given at the beginning of section 3 of Bai and Silverstein (1998) carries through in our case. Specifically, for any $C > 0$ let $Y_{ij} = X_{ij}I_{(|X_{ij}|\le C)} - \mathbb{E}X_{ij}I_{(|X_{ij}|\le C)}$ (where I_A denotes the indicator function on the set A), $Y_n = (Y_{ij})$, $i \le n$, $j \le N$, $\tilde{C}_n = (1/N)A_n^{1/2}Y_nB_nY_n^*A_n^{1/2}$, and λ_k, $\tilde{\lambda}_k$ the respective eigenvalues of C_n and $\tilde{C}_n$ in nonincreasing order. Then as in Bai and Silverstein (1998), using the main result in Yin, Bai, and Krishnaiah (1988) on the largest eigenvalue of $(1/N)X_nX_n^*$, we have, with probability 1

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

and because of assumption (a) we can make the bound on the right side arbitrarily small by choosing C sufficiently large. Thus we can assume that the X_{ij} are uniformly bounded.

The rest of the paper is organized as follows. In Section 2, we give the key steps to the derivation of the integral equations for the limiting Stieltjes transforms of associated spectral measures. In Sections 3 and 4 we will show, respectively

$$\sup_{x\in[a,b]} |m_n(z) - \mathbb{E}m_n(z)| = o(1/(nv_n)) \quad \text{a.s.} \tag{6}$$

and

$$\sup_{x\in[a,b]} |\mathbb{E}m_n(z) - Em_n^0(z)| = O(1/n). \tag{7}$$

Some mathematical tools needed in proving these results are given in the Appendix.

2 Integral representation of Stieltjes transforms

Write $X_n = [x_1, \ldots, x_N]$, and let $y_j = (1/\sqrt{N})A_n^{1/2}x_j$. Then we can write

$$C_n = \sum_{j=1}^{N} b_j y_j y_j^*.$$

Fix $z \in \mathbb{C}^+ \equiv \{z = x + iv \in \mathbb{C} : v > 0\}$. Define

$$e_n = e_n(z) = (1/n)\mathrm{tr}\, A_n(C_n - zI)^{-1}, \tag{8}$$

and

$$p_n = -\frac{1}{Nz}\sum_{j=1}^{N} \frac{b_j}{1 + c_n b_j e_n} = \int \frac{-b}{z(1 + c_n b e_n)} dF^{B_n}(b). \tag{9}$$

Notice that both e_n and p_n are Stieltjes transforms of measures on the non-negative reals with total mass $(1/n)\mathrm{tr}\, A_n$ and $(1/N)\mathrm{tr}\, B_n$, respectively. The latter is true since both ze_n and zp_n map $\mathbb{C}^+$ into $\mathbb{C}^+$, and as $z \to \infty$, $zp_n \to -(1/N)\mathrm{tr}\, B_n$ (cf. Krein and Nudelman (1997)). It follows that these quantities are bounded in absolute value by $v^{-1}(1/n)\mathrm{tr}\, A_n$ and $v^{-1}(1/N)\mathrm{tr}\, B_n$, respectively. Another example of a Stieltjes transform is

$$\frac{-b}{z(1 + m(z))},$$

where $b \geq 0$ and $m(z)$ is the Stieltjes transform of a bounded measure on $\mathbb{R}^+$. It follows that this quantity is bounded in absolute value by bv^{-1}.

Let $C_{(j)} = C_n - b_j y_j y_j^*$. We may, without loss of generality, assume that $\max(\|A_n\|, \|B_n\|) \leq 1$. Write

$$C_n - zI + zI + zp_nA_n = \sum_{j=1}^{N} b_j y_j y_j^* + zp_nA_n.$$

Taking inverses and using the definition of C_n and $C_{(j)}$, we have

$$
\begin{aligned}
& (C_n - zI)^{-1} + (zI + zp_nA_n)^{-1} \\
= & \sum_{j=1}^{N} b_j (C_n - zI)^{-1} y_j y_j^* (zI + zp_nA)^{-1} + zp_n (C_n - zI)^{-1} A_n (zI + zp_nA)^{-1} \\
= & \sum_{j=1}^{N} b_j \frac{(C_{(j)} - zI)^{-1} y_j y_j^* (zI + zp_nA)^{-1}}{1 + b_j y_j^* (C_{(j)} - zI)^{-1} y_j} + zp_n (C_n - zI)^{-1} A_n (zI + zp_nA)^{-1}.
\end{aligned}
$$

Taking traces and dividing by n, we have

$$
m_n(z) - \int \frac{1}{a \int \frac{b}{1+c_nbe_n} dF^{B_n}(b) - z} dF^{A_n}(a) = \frac{1}{N} \sum_{j=1}^{N} b_j d_j \equiv w_n^m,
$$

where

$$
d_j = \frac{(1/n) x_j^* A_n^{1/2} (I + p_nA_n)^{-1} (C_{(j)} - zI)^{-1} A_n^{1/2} x_j}{z(1 + b_j y_j^* (C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n) \mathrm{tr}\, (C_n - zI)^{-1} A_n (I + p_nA_n)^{-1}}{z(1 + c_n b_j e_n)}.
$$

Multiplying both sides of the above matrix identity by A_n, and then taking traces and dividing by n, we find

$$
e_n(z) - \int \frac{a}{a \int \frac{b}{1+c_nbe_n} dF^{B_n}(b) - z} dF^{A_n}(a) = \frac{1}{N} \sum_{j=1}^{N} b_j d_j^e \equiv w_n^e,
$$

where

$$
d_j^e = \frac{(1/n) x_j^* A_n^{1/2} (I + p_nA_n)^{-1} A_n (C_{(j)} - zI)^{-1} A_n^{1/2} x_j}{z(1 + b_j y_j^* (C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n) \mathrm{tr}\, A_n (C_n - zI)^{-1} A_n (I + p_nA_n)^{-1}}{z(1 + c_n b_j e_n)}.
$$

2.1 Bound on the approximation error

Notice that for each j, $y_j^* (C_{(j)} - zI)^{-1} y_j$ can be viewed as a Stieltjes transform of a measure on $\mathbb{R}^+$. Therefore

$$
\left| \frac{1}{z(1 + b_j y_j^* (C_{(j)} - zI)^{-1} y_j)} \right| \leq \frac{1}{v}.
$$

For each j, let $e_{(j)} = e_{(j)}(z) = (1/n) \mathrm{tr}\, A_n (C_{(j)} - zI)^{-1}$, and

$$
p_{(j)} = p_{(j)}(z) = \int \frac{-b}{z(1 + c_n b e_{(j)})} dF^{B_n}(b),
$$

both of course being Stieltjes transforms of measures on $\mathbb{R}^+$, along with the integrand for each b.

Using Lemma 1 and Lemma 2(a) we have

$$
|p_n - p_{(j)}| = |e_n - e_{(j)}| c_n \left| \int \frac{b^2}{z(1 + c_n b e_n)(1 + c_n b e_{(j)})} dF^{B_n(b)} \right| \leq \frac{4c_n^2}{nv^3}. \tag{10}
$$

In order to handle both w_n^m, d_j and w_n^e, d_j^e at the same time, we shall denote by E_n either A_n or I_n, and w_n, d_j for now will denote either the original w_n^m, d_j or w_n^e, d_j^e.

Write $d_j = d_j^1 + d_j^2 + d_j^3 + d_j^4$, where

$$d_j^1 = \frac{(1/n)x_j^* A_n^{1/2}(I + p_n A_n)^{-1} E_n (C_{(j)} - zI)^{-1} A_n^{1/2} x_j}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n)x_j^* A_n^{1/2}(I + p_{(j)} A_n)^{-1} E_n (C_{(j)} - zI)^{-1} A_n^{1/2} x_j}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)},$$

$$d_j^2 = \frac{(1/n)x_j^* A_n^{1/2}(I + p_{(j)} A_n)^{-1} E_n (C_{(j)} - zI)^{-1} A_n^{1/2} x_j}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n)\mathrm{tr}\, E_n (C_{(j)} - zI)^{-1} A_n (I + p_{(j)} A_n)^{-1}}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)},$$

$$d_j^3 = \frac{(1/n)\mathrm{tr}\, E_n (C_{(j)} - zI)^{-1} A_n (I + p_{(j)} A_n)^{-1}}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n)\mathrm{tr}\, E_n (C_n - zI)^{-1} A_n (I + p_n A_n)^{-1}}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)},$$

and

$$d_j^4 = \frac{(1/n)\mathrm{tr}\, E_n (C_n - zI)^{-1} A_n (I + p_n A_n)^{-1}}{z(1 + b_j y_j^*(C_{(j)} - zI)^{-1} y_j)} - \frac{(1/n)\mathrm{tr}\, E_n (C_n - zI)^{-1} A_n (I + p_n A_n)^{-1}}{z(1 + c_n b_j e_n)}.$$

From here on, we assume that for all n large, $v = v_n = \kappa n^{-\delta}$ for some $\kappa > 0$ and $\delta \geq 0$. We wish to show that for all $\delta \leq 1/4$, arbitrary subset $S_n \subset [0, \infty)$ containing at most n elements, and arbitrary positive t and ϵ, we have

$$\mathbb{P}(\max_{x \in S_n} |w_n| v_n^{-12} > \epsilon) \leq K_t \epsilon^{-2t} n^{2-t(1-34\delta)} \tag{11}$$

by proving the same bound on each of

$$\mathbb{P}(\max_{j \leq N, x \in S_n} |d_j^i| v_n^{-12} > \epsilon),$$

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

We begin with d_j^1. We get from Lemma 2(c) and (10),

$$|d_j^1| \leq \frac{1}{v_n} \frac{4c_n^2}{n v_n^3} \frac{1}{v_n} \frac{\|x_j\|^2}{n} \frac{16}{v_n^2} = \frac{64 c_n^2}{n v_n^7} \frac{\|x_j\|^2}{n}.$$

So, by Lemma 3 we have, for any $\epsilon > 0$, $p \geq 2$, for all n large

$$\mathbb{P}(\max_{j \leq N, x \in S_n} |d_j^1| v_n^{-12} > \epsilon) \leq n \mathbb{P}\left(\max_{j \leq N} \left|\frac{\|x_j\|^2}{n} - 1\right| \frac{64 c_n^2}{n v_n^{19}} > \frac{\epsilon}{2}\right) \leq K_p \frac{nN}{(n v_n^{19})^p} \epsilon^{-p} n^{-p/2}.$$

For d_j^2 we use Lemma 2(a) and Lemma 3 to get, for $p \geq 2$,

$$\mathbb{E}|v_n^{-12} d_j^2|^p \leq K_p v_n^{-13p} n^{-p/2} v_n^{-2p} = K_p \frac{1}{(n^{1/2} v_n^{15})^p},$$

so that for $\epsilon > 0$, $p \geq 2$

$$\mathbb{P}(\max_{j \leq N, x \in S_n} |d_j^2| v_n^{-12} > \epsilon) \leq K_p \epsilon^{-p} \frac{nN}{(n^{1/2} v_n^{15})^p}.$$

Using Lemma 1, Lemma 2(a), 2(b), and (10) we have,

$$|d_j^3 v_n^{-12}| \leq \frac{K}{v_n^{13}} \left(\frac{1}{nv_n^2} + \frac{1}{nv_n^6} \right) \leq K \frac{1}{nv_n^{19}}.$$

Therefore for any $p \geq 1$ and $\epsilon > 0$

$$\mathbb{P}(\max_{j \leq N, x \in S_n} |d_j^3| v_n^{-12} > \epsilon) \leq K_p \epsilon^{-p} \frac{nN}{(nv_n^{19})^p}.$$

Finally, for d_j^4, we use Lemma 1 and Lemma 2(a) to find:

$$|d_j^4 v_n^{-12}| \leq \frac{4}{v_n^{16}} (|(1/n) x_j^* A_n (C_{(j)} - zI)^{-1} x_j - (1/n) \mathrm{tr}\, A_n (C_{(j)} - zI)^{-1}| + (nv_n)^{-1}).$$

Therefore, by Lemma 3, for any $\epsilon > 0$, $p \geq 2$, we have

$$\begin{aligned} & \mathbb{P}(\max_{j \leq N, x \in S_n} |d_j^4| v_n^{-12} > \epsilon) \\ \leq & \sum_{j \leq N, x \in S_n} \left[\mathbb{P} \left(\frac{4}{v_n^{16}} |(1/n) x_j^* A_n (C_{(j)} - zI)^{-1} x_j - (1/n) \mathrm{tr}\, A_n (C_{(j)} - zI)^{-1}| > \frac{\epsilon}{2} \right) + K_p \epsilon^{-p} (nv_n^{11})^{-p} \right] \\ \leq & K_p \epsilon^{-p} \frac{nN}{(n^{1/2} v_n^{17})^p}, \end{aligned}$$

which, for $\delta \in [0, 1/4]$, can easily be verified to be the largest of the four bounds. Therefore, (11) holds.

2.2 Existence, convergence, and continuity of the solution

We can at this stage provide a proof of the existence of a unique e with nonnegative imaginary part satisfying (2) for any $z = x + iv$, $v > 0$, and the a.s. convergence in distribution of F^{C_n} to F. We also show the continuous dependence of e on F^A, F^B, and c. We see from (11) with $\delta = 0$, $\kappa = v$ and $t > 3$ we have

$$e_n(z) - \int \frac{a}{a \int \frac{b}{1 + c_n b e_n} dF^{B_n}(b) - z} dF^{A_n}(a) \quad \text{and} \quad m_n(z) - \int \frac{1}{a \int \frac{b}{1 + c_n b e_n} dF^{B_n}(b) - z} dF^{A_n}(a)$$

converge a.s. to zero. Consider a realization for which both convergences to zero occur on a subsequence $\{n_i\}$ for which e_n converges, say to e. Since $\Im(z e_n(z)) > 0$, we have for any $b \geq 0$,

$$\left| \frac{b}{z(1 + c_n b e_n)} \right| \leq \frac{1}{v}.$$

Therefore, by the Dominated Convergence Theorem (DCT)

$$p_n = \int \frac{-b}{z(1+c_n b e_n)} dF^{B_n}(b) \to \int \frac{-b}{z(1+cbe)} dF^B(b) \equiv p$$

along the subsequence. Since $\Im(zp_n(z)) > 0$, we also have

$$\left| \frac{a}{z(1+c_n a p_n)} \right| \leq \left| \frac{1}{z(1+c_n a p_n)} \right| \leq \frac{1}{v}.$$

Therefore, again by the DCT,

$$\begin{aligned} \int \frac{a}{a\int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z} dF^{A_n}(a) &= \int \frac{-a}{z(1+c_n a p_n)} dF^{A_n}(a) \\ &\to \int \frac{-a}{z(1+cap)} dF^A(a) \\ &= \int \frac{a}{a\int \frac{b}{1+cbe} dF^B(b) - z} dF^A(a) \end{aligned}$$

along the subsequence. Thus e is a solution to (2).

From uniqueness of e, proved below, we must have convergence of e_n to e on the whole sequence. Therefore, again by the DCT we have

$$m_n(z) \to \int \frac{1}{a\int \frac{b}{1+cbe} dF^B(b) - z} dF^A(a).$$

This event occurs with probability one, and for a countable number of v's with a limit point. Since the lim sup of the largest eigenvalue of C_n is a.s. bounded by $(1+\sqrt{c})^2$, the sequence $\{F^{C_n}\}$ is almost surely tight. Therefore, F^{C_n} converges in distribution to F a.s.

For probability distribution functions $F^{\underline{A}}$ and $F^{\underline{B}}$ on $[0,1]$ and $\underline{c} > 0$, let $\underline{e} = \underline{e}(z)$ be a solution to (2) with F^A, F^B, c replaced by $F^{\underline{A}}$, $F^{\underline{B}}$, and $\underline{c}$, respectively. Assume that $c \leq \underline{c}$. Then we have

$$\begin{aligned} e - \underline{e} &= \int \frac{a}{a\int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z} d(F^A(a) - F^{\underline{A}}(a)) \\ &+ \int \frac{a^2 \int \frac{b}{1+\underline{c}b\underline{e}} d(F^{\underline{B}}(b) - F^B(b))}{(a\int \frac{b}{1+cbe} dF^B(b) - z)(a\int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z)} dF^A(a) \\ &+ (c - \underline{c}) \int \underline{e} \frac{a^2 \int \frac{b^2}{(1+cbe)(1+\underline{c}b\underline{e})} dF^B(b)}{(a\int \frac{b}{1+cbe} dF^B(b) - z)(a\int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z)} dF^A(a) + \gamma(e - \underline{e}), \end{aligned} \quad (12)$$

where

$$\gamma = c \int \frac{a^2 \int \frac{b^2}{(1+cbe)(1+\underline{c}b\underline{e})} dF^B(b)}{(a\int \frac{b}{1+cbe} dF^B(b) - z)(a\int \frac{b}{1+\underline{c}b\underline{e}} dF^B(b) - z)} dF^A(a).$$

Notice that the first integrand in (12) is bounded in absolute value by $1/v$, the second by $|z|/v^3$, and the third by $|z|^2/v^5$. Let e_2 and $\underline{e}_2$ denote the imaginary parts of e and $\underline{e}$. Then we write

$$e_2 = \int \frac{a(ace_2 \int \frac{b^2}{|1+cbe|^2} dF^B(b) + v)}{|a\int \frac{b}{1+cbe} dF^B(b) - z|^2} dF^A(a) = e_2\alpha + v\beta,$$

and

$$\underline{e}_2 = \int \frac{a(a\underline{c}\ \underline{e}_2 \int \frac{b^2}{|1+\underline{c}b\underline{e}|^2} dF^{\underline{B}}(b) + v)}{|a \int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z|^2} dF^{\underline{A}}(a) = \underline{e}_2\underline{\alpha} + v\underline{\beta}.$$

We have used Cauchy-Schwartz inequality:

$$\begin{aligned} |\gamma| &\leq \int \left(\frac{ca^2 \int \frac{b^2}{|1+cbe_n|^2} dF^B(b)}{\left| a \int \frac{b}{1+cbe} dF^B(b) - z \right|^2} \right)^{1/2} \left(\frac{\underline{c}a^2 \int \frac{b^2}{|1+\underline{c}b\underline{e}|^2} dF^{\underline{B}}(b)}{\left| a \int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z \right|^2} \right)^{1/2} dF^A \\ &\leq \left(\int \frac{ca^2 \int \frac{b^2}{|1+cbe|^2} dF^B(b)}{\left| a \int \frac{b}{1+cbe} dF^B(b) - z \right|^2} dF^A(a) \right)^{1/2} \left(\int \frac{\underline{c}a^2 \int b^2 |1+\underline{c}b\underline{e}|^2 dF^{\underline{B}}(b)}{\left| a \int \frac{b}{1+\underline{c}b\underline{e}} dF^{\underline{B}}(b) - z \right|^2} dF^A(a) \right)^{1/2} \\ &= \left(\frac{e_2\alpha}{e_2\alpha + v\beta} \right)^{1/2} \left(\frac{\underline{e}_2\underline{\alpha}}{\underline{e}_2\underline{\alpha} + v\underline{\beta}} \right)^{1/2}. \end{aligned}$$

Notice that for v small we have by Lemma 2(a)

$$e_2\alpha/\beta \leq e_2 c \int \frac{b^2}{|1+cbe|^2} dF^B(b) = -\Im \int \frac{b}{1+cbe} dF^B \leq \frac{4c}{v}.$$

Therefore

$$\begin{aligned} \left(\int \frac{ca^2 \int \frac{b^2}{|1+cbe|^2} dF^B(b)}{\left| a \int \frac{b}{1+cbe} dF^B(b) - z \right|^2} dF^A(a) \right)^{1/2} &= \left(\frac{e_2\alpha}{e_2\alpha + v\beta} \right)^{1/2} \\ = \left(\frac{e_2\alpha/\beta}{v + e_2\alpha/\beta} \right)^{1/2} \leq \left(\frac{4c}{v^2 + 4c} \right)^{1/2} &\leq 1 - Kv^2 \end{aligned}$$

for some positive constant K. A corresponding bound obviously exists for the other factor making up the bound on γ, so we conclude that for v small

$$|\gamma| \leq 1 - Kv^2 \tag{13}$$

for some positive constant K.

We see then that (12) and (13) together reveal two things: uniqueness of solutions to (2) (with $F^A = F^{\underline{A}}$, $F^B = F^{\underline{B}}$, and $c = \underline{c}$), and continuous dependence of solutions to (2) on F^A, F^B (under the topology of weak convergence of probability measures from the DCT), and c.

2.3 Bound on the difference between Stieltjes transforms

We have then $e_n^0 = e_n^0(z)$, a unique solution to

$$e = \int \frac{a}{a \int \frac{b}{1+c_n be} dF^{B_n}(b) - z} dF^{A_n}(a). \tag{14}$$

Let

$$m_n^0 = m_n^0(z) = \int \frac{1}{a \int \frac{b}{1+c_n be_n^0} dF^{B_n}(b) - z} dF^{A_n}(a). \tag{15}$$

This of course is the Stieltjes transform of F^{c_n,A_n,B_n}.

Let e_2^0, e_2, m_2^0, m_2 denote the imaginary parts of e_n^0, e_n, m_n^0, m_n, respectively. Then

$$e_2^0 = \int \frac{a(ac_n e_2^0 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b) + v)}{|a \int \frac{b}{1+c_n b e_n^0} dF^{B_n}(b) - z|^2} dF^{A_n}(a),$$

$$e_2 = \int \frac{a(ac_n e_2 \int \frac{b^2}{|1+c_n b e_n|^2} dF^{B_n}(b) + v)}{|a \int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z|^2} dF^{A_n}(a) + \Im w_n^e,$$

$$m_2^0 = \int \frac{ac_n e_2^0 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b) + v}{|a \int \frac{b}{1+c_n b e_n^0} dF^{B_n}(b) - z|^2} dF^{A_n}(a),$$

$$m_2 = \int \frac{ac_n e_2 \int \frac{b^2}{|1+c_n b e_n|^2} dF^{B_n}(b) + v}{|a \int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z|^2} dF^{A_n}(a) + \Im w_n^m,$$

and as above we have

$$e_n - e_n^0 = (e_n - e_n^0)\underline{\gamma}_n + w_n^e,$$

where

$$|\underline{\gamma}_n| \le \left(\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) \right)^{1/2} \left(\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n^0} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) \right)^{1/2}.$$

Writing $e_2^0 = e_2^0 \alpha + v\beta$, we have again

$$e_2^0 \alpha / \beta \le \frac{4c_n}{v_n},$$

and subsequently

$$\left(\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n^0} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) \right)^{1/2} \le \left(\frac{4c_n}{v^2 + 4c_n} \right)^{1/2} \le 1 - Kv^2$$

for some positive constant K.

At this point on we assume that $\delta \in (0, 1/35]$, so that $v_n \to 0$ as $n \to \infty$. Let $\mu_n = n^{\delta/4}$. Therefore we have $v_n \mu_n^3 \to 0$.

Write $C_n = O\Lambda O^*$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, in its spectral decomposition. Let $\underline{A}_n = \{\underline{a}_{ij}\} = O^* A_n O$. Then

$$e_n = (1/n)\mathrm{tr}\, \underline{A}_n (\Lambda - zI)^{-1} = (1/n) \sum_{i=1}^{n} \frac{\underline{a}_{ii}}{\lambda_i - z}.$$

Let $\lambda_{\max}$ denote the largest eigenvalue of $(1/N)X_n X_n^*$, and let $K_1 > (1+\sqrt{c})^2$. Then, as in Bai and Silverstein (1998) p. 329, for all n large $|e_n| \ge \frac{1}{2}\mu_n^{-1} v_n^3 (1/n)\mathrm{tr}\, A_n$ whenever $|x| \le \mu_n v_n^{-3}$ and $\lambda_{\max} \le K_1$. We also have for all n large

$$|e_n - e_n^0| \le 3\mu_n^{-1} v_n^3$$

whenever $|x| > \mu_n v_n^{-3}$ and $\lambda_{\max} \leq K_1$. Notice that whenever $(1/n)\text{tr}\, A_n \leq v_n^4 \mu_n^{-1}$ we have

$$v_n^{-3}|e_n - e_n^0| \leq 2v_n^{-4}(1/n)\text{tr}\, A_n \leq 2\mu_n^{-1}.$$

Let α, β be such that $e_2 = e_2\alpha + v_n\beta + \Im w_n^e$. Then

$$\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) = \frac{e_2\alpha}{e_2\alpha + v_n\beta + \Im w_n^e}.$$

Using Cauchy-Schwartz we have

$$|e_n| \leq \beta^{1/2}((1/n)\text{tr}\, A_n)^{1/2} + |w_n^e|.$$

So, for all n large, whenever $|x| \leq \mu_n v_n^{-3}$, $|w_n^e| \leq v_n^{12}$, $\lambda_{\max} \leq K_1$, and $(1/n)\text{tr}\, A_n > v_n^4 \mu^{-1}$ we have

$$\frac{1}{2}\mu_n^{-1} v_n^3 (1/n)\text{tr}\, A_n \leq |e_n| \leq \beta^{1/2}((1/n)\text{tr}\, A_n)^{1/2} + |w_n^e| \leq \beta^{1/2}((1/n)\text{tr}\, A_n)^{1/2} + \mu_n v_n^8 (1/n)\text{tr}\, A_n.$$

Therefore

$$\frac{1}{3}\mu_n^{-1} v_n^3 (1/n)\text{tr}\, A_n \leq (1/n)\text{tr}\, A_n (\frac{1}{2}\mu_n^{-1} v_n^3 - \mu_n v_n^8) \leq \beta^{1/2}((1/n)\text{tr}\, A_n)^{1/2},$$

from which we get

$$\beta \geq \frac{1}{9} v_n^{10} \mu_n^{-3}.$$

Therefore

$$v_n\beta + \Im w_n^e \geq \frac{1}{9} v_n^{11} \mu_n^{-3} - v_n^{12} > 0,$$

and so

$$|e_n - e_n^0| \leq K^{-1} v_n^{-2} |w_n^e| \leq K^{-1} v_n^{10}.$$

Therefore, for all n large

$$\max_{x \in S_n} v_n^{-3} |e_n - e_n^0| \leq K^{-1} v_n^7 + 3\mu_n^{-1} + 2v_n^{-4} \max_{x \in S_n} (I_{[|w_n^e| > v_n^{12}]} + I_{[\lambda_{\max} > K_1]}).$$

Therefore, for any positive ϵ and t we have for all n large

$$\begin{aligned} & \mathbb{P}(\max_{x \in S_n} v_n^{-3} |e_n - e_n^0| > \epsilon) \\ \leq \quad & K_t \epsilon^{-t} \left(n^{-7\delta t} + n^{-\delta t/4} + v_n^{-4t} [\mathbb{P}(\max_{x \in S_n} |w_n^e| v_n^{-12} > 1) + \mathbb{P}(\lambda_{\max} > K_1)] \right) \\ \leq \quad & K_t \epsilon^{-t} n^{-\delta t/4}, \end{aligned} \tag{16}$$

where the last step follows by replacing t with

$$\frac{\frac{17}{4}\delta t + 3}{1 - 34\delta}$$

in (11) and t with $\frac{17}{4}\delta t$ in Lemma 4. Taking the difference between m_n and m_n^0 and using Cauchy-Schwartz, we get

$$
\begin{aligned}
&|m_n - m_n^0| \\
\leq\ & |e_n - e_n^0| \left(\int \frac{c_n \int \frac{b^2}{|1+c_n b e_n|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) \right)^{1/2} \left(\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b)}{\left| a \int \frac{b}{1+c_n b e_n^0} dF^{B_n}(b) - z \right|^2} dF^{A_n}(a) \right)^{1/2} \\
& + |w_n^m|.
\end{aligned}
$$

As before, we have the second factor on the right bounded above by 1, while the first factor is bounded above by $c_n^{1/2} v_n^{-2}$. Therefore, for any positive ϵ and t we get from (11) and (16),

$$
\begin{aligned}
\mathbb{P}(\max_{x \in S_n} v_n^{-1} |m_n - m_n^0| > \epsilon) &\leq \mathbb{P}(\max_{x \in S_n} v_n^{-3} |e_n - e_n^0| > \epsilon (2c_n^{1/2})^{-1}) + \mathbb{P}(\max_{x \in S_n} |w_n^m| v_n^{-1} > \epsilon/2) \\
&\leq K_t \epsilon^{-t} \max(n^{-\delta t/4}, n^{2-t(1/2-17\delta)}). \qquad (17)
\end{aligned}
$$

It is easy to verify from (17) that

$$
\mathbb{P}(\max_{x \in S_n} v_n^{-1} |m_n - m_n^0| > \epsilon) \leq K_t \epsilon^{-t} n^{-\delta t/4} \quad \text{for} \quad t \geq 280. \qquad (18)
$$

2.4 Bound on the number of eigenvalues falling outside the support

Suppose that the n elements in S_n are equally spaced between $-\sqrt{n}$ and $\sqrt{n}$. Since, for $|x_1 - x_2| \leq 2n^{-1/2}$,

$$
\begin{aligned}
|m_n(x_1 + iv_n) - m_n(x_2 + iv_n)| &\leq 2n^{-1/2} v_n^{-2} \\
|m_n^0(x_1 + iv_n) - m_n^0(x_2 + iv_n)| &\leq 2n^{-1/2} v_n^{-2},
\end{aligned}
$$

and when $|x| \geq \sqrt{n}$, for n large enough, for K as in Lemma 4,

$$
|m_n(x + iv_n)| \leq 2n^{-1/2} + v_n^{-1} I(\lambda_{\max} > K),
$$

and

$$
|m_n^0(x + iv_n)| \leq 2n^{-1/2},
$$

we conclude from (18) and Lemma 4, that for any $\epsilon > 0$ and $t \geq 280$, $0 < \delta \leq \frac{1}{35}$, for n large enough,

$$
\mathbb{P}(v_n^{-1} \sup_{x \in \mathbb{R}} |m_n - m_n^0| > \epsilon) \leq K_t \epsilon^{-t} n^{-\delta t/4}. \qquad (19)
$$

Let $\mathbb{E}_0$ denote the expectation, and $\mathbb{E}_k$ denote the conditional expectation with respect to the σ-field generated by $\{y_1, \ldots, y_k\}$. Since for any $r > 0$,

$$
\mathbb{E}_k(v_n^{-r} \sup_{x \in \mathbb{R}} |m_n(x + iv_n) - m_n^0(x + iv_n)|^r),
$$

for $k = 0, 1, \ldots, n$ forms a martingale, from *Jensen's inequality* it follows that for any $t \geq 1$,

$$
(\mathbb{E}_k(v_n^{-r} \sup_{x \in \mathbb{R}} |m_n(x + iv_n) - m_n^0(x + iv_n)|^r))^t
$$

for $k = 0, 1, \ldots, n$ forms a submartingale. Therefore, from Lemma 2.5 and Lemma 2.6 of Bai and Silverstein (1998), and (19), for any $\epsilon > 0$, $t \geq 1$, and $r > 0$, so that $2rt \geq 280$, we have,

$$\begin{aligned} & \mathbb{P}(\max_{k \leq N} \mathbb{E}_k(v_n^{-r} \sup_{x \in \mathbb{R}} |m_n(x+iv_n) - m_n^0(x+iv_n)|^r) > \epsilon) \\ \leq \; & \epsilon^{-t} \mathbb{E}(v_n^{-rt} \sup_{x \in \mathbb{R}} |m_n(x+iv_n) - m_n^0(x+iv_n)|^{rt}) \\ \leq \; & 2\epsilon^{-t} K_{rt}^{1/2} n^{-\delta rt/4}, \end{aligned} \tag{20}$$

whenever $\delta \in (0, 1/35]$. From this, it follows that with probability 1,

$$\max_{k \leq N} \mathbb{E}_k(v_n^{-r} \sup_{x \in \mathbb{R}} |m_n(x+iv_n) - m_n^0(x+iv_n)|^r) \to 0. \tag{21}$$

Let $\underline{\epsilon} > 0$ be such that $[a', b']$, with $a' = a - \underline{\epsilon}$ and $b' = b + \underline{\epsilon}$, also satisfy condition (f) of Theorem 1. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of $C_n := \frac{1}{N} A_n^{1/2} X_n B_n X_n^* A_n^{1/2}$, and write

$$m_{nj} = m_{nj}^{out} + m_{nj}^{in}, \qquad j = 1, 2,$$

where $j = 1$ refers to the real part of m_n and $j = 2$ refers to the imaginary part of m_n, so that

$$m_{n2}^{out}(x+iv_n) = \frac{1}{N} \sum_{\lambda_j \in [a', b']} \frac{v_n}{(x-\lambda_j)^2 + v_n^2} \ ,$$

and

$$m_{n1}^{out}(x+iv) = \frac{1}{N} \sum_{\lambda_j \in [a', b']} \frac{x-\lambda_j}{(x-\lambda_j)^2 + v_n^2}.$$

Similarly, write m_{n1}^0 and m_{n2}^0 to mean the real and imaginary parts of m_n^0 and define,

$$m_{02}^{out}(x+iv_n) = \int_{a'}^{b'} \frac{v_n}{(x-u)^2 + v_n^2} dF^{c_n, A_n, B_n}(u)$$

Note that for all $x \in [a, b]$, $m_{02}^{out}(x+iv_n) \equiv 0$, so that $m_{n2}^0(x+iv_n) \equiv m_{02}^{in}(x+iv_n)$, for large n. By (21), with probability 1,

$$\max_{k \leq N} \mathbb{E}_k(v_n^{-r} \sup_{x \in \mathbb{R}} |m_{n2}(x+iv_n) - m_{n2}^0(x+iv_n)|^r) \to 0. \tag{22}$$

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

$$G_{\sum_{j=1}^{n-1}(N(j)+1)+k}(x_1, x_2) = \mathbb{E}_k F^{C_n}(x_1) F^{C_n}(x_2),$$

for $k = 0, 1, 2, \ldots, N(n)$. Clearly, each G_q is a probability distribution function on $\mathbb{R}^2$. Also, for $q = \sum_{j=1}^{n-1}(N(j)+1) + k$, the two-dimensional Stieltjes transform, $m_q^{(G)}(x_1 + iv_1, x_2 + iv_2)$ of G_q is $\mathbb{E}_k m_n(x_1 + iv_1) m_n(x_2 + iv_2)$. Notice $x < 0$, $\lambda > 0$ implies that

$$\left| \frac{1}{\lambda - (x+iv)} - \frac{1}{\lambda - x} \right| \leq \frac{v}{x^2}.$$

Therefore, from (21), we have with probability 1,

$$|m_q^{(G)}(x_1, x_2) - m_n^0(x_1) m_n^0(x_2)| \to 0, \quad \text{as} \quad q \to \infty,$$

for countably many negative x_1 having a negative limit point, and countably many negative x_2 also having a negative limit point.

It is straightforward to show the following: Assume that $f(z_1, z_2)$ is a function of two complex variables, and analytic on a open rectangle $E \times F \subset \mathbb{C}^2$ (that is, for fixed $z_1 \in E$ $f(z_1, z_2)$ is analytic in z_2, and visa versa). Let $\{z_1^n\} \subset E$, $\{z_2^n\} \subset F$, where $\{z_1^n\}$ has a limit point in E, $\{z_2^n\}$ has a limit point in F. Then f is uniquely determined by the values it places on the set $\{(z_1, z_2) : z_1 \in \{z_1^n\}, z_2 \in \{z_2^n\}\}$. This, together with the a.s. tightness of G_q, gives us, with probability 1, $G_q(y_1, y_2)$ converging weakly to $F(y_1)F(y_2)$.

Notice that the integrands of

$$\int_{[a',b']^c \times [a',b']^c} \frac{d\mathbb{E}_k F^{C_n}(x_1) F^{C_n}(x_2)}{((x - x_1)^2 + v_n^2)((x - x_2)^2 + v_n^2)}$$

and

$$\int_{[a',b']^c} \frac{d\mathbb{E}_k F^{C_n}(x_1)}{(x - x_1)^2 + v_n^2}$$

on their respective domains are uniformly bounded and equicontinuous for $x \in [a, b]$. Therefore, from *Problem 8, p. 17*, in Billingsley (1968), the sequence $\{g_q\}_{q=1}^{\infty}$ defined through,

$$g_q = \sup_{x \in [a,b]} \mathbb{E}_k v_n^{-2} |m_{n2}^{in}(x + iv_n) - m_{n2}^0(x + iv_n)|^2$$

for $q = \sum_{j=1}^{n-1}(N(j) + 1) + k$ (with $0 \leq k \leq N(n)$), converges to 0 a.s. as $n \to \infty$. Thus we have, a.s., $\max_{n \geq n_0} \max_{1 \leq k \leq N(n)} g_{\sum_{j=1}^{n-1}(N(j)+1)+k} \to 0$, as $n_0 \to \infty$. This implies that

$$\max_{0 \leq k \leq N} \sup_{x \in [a,b]} \mathbb{E}_k v_n^{-2} |m_{n2}^{in}(x + iv_n) - m_{n2}^0(x + iv_n)|^2 \to 0 \quad \text{a.s.}$$

This, together with (22), and the fact that $m_{n2}^{out}(x + iv_n) \equiv 0$ for $x \in [a, b]$, implies that

$$\sup_{x \in [a,b]} \max_{k \leq N} v_n^{-2} \mathbb{E}_k (m_{n2}^{out}(x + iv_n))^2 \to 0. \tag{23}$$

Now, for any $x \in [a, b]$, we have

$$\begin{aligned} v_n^{-1} m_{n2}^{out}(x + iv_n) &\geq \int_a^b \frac{1}{(u - x)^2 + v_n^2} dF^{C_n}(u) \\ &\geq \int_{[a,b] \cap [x - v_n, x + v_n]} \frac{1}{(u - x)^2 + v_n^2} dF^{C_n}(u) \\ &\geq \frac{1}{2v_n^2} F^{C_n}([a, b] \cap [x - v_n, x + v_n]). \end{aligned}$$

We select $x_j \in [a, b]$, $j = 1, \ldots, J$, such that $v_n < x_j - x_{j-1}$, and $[a, b] \subset \cup_{j=1}^J [x_j - v_n, x_j + v_n]$. Notice that, as a consequence, $J \leq (b - a)v_n^{-1}$. Then from the inequality above, it follows that, with probability 1,

$$
\begin{aligned}
v_n^{-2}\max_{k\le N}\mathbb{E}_k(F^{C_n}([a,b]))^2 &\le v_n^{-2}\max_{k\le N}\mathbb{E}_k(\sum_{j=1}^{J}F^{C_n}([a,b]\cap[x_j-v_n,x_j+v_n]))^2\\
&\le v_n^{-2}\max_{k\le N}\mathbb{E}_k(\sum_{j=1}^{J}2v_n m_{n2}^{out}(x_j+iv_n))^2\\
&\le 4J\max_{k\le N}\sum_{j=1}^{J}\mathbb{E}_k(m_{n2}^{out}(x_j+iv_n))^2, \quad \text{by Hölder's inequality}\\
&\le 4J^2\max_{1\le j\le J}\max_{k\le N}\mathbb{E}_k(m_{n2}^{out}(x_j+iv_n))^2\\
&\le 4(b-a)^2v_n^{-2}\sup_{x\in[a,b]}\max_{k\le N}\mathbb{E}_k(m_{n2}^{out}(x+iv_n))^2\\
&\to 0,
\end{aligned}
$$

by (23).

This shows that,

$$\max_{k\le N}\mathbb{E}_k(F^{C_n}([a,b]))^2 = o(v_n^2), \quad \text{a.s.}$$

Clearly, the same argument holds for $[a',b']$ replacing $[a,b]$, and so we have

$$\max_{k\le N}\mathbb{E}_k(F^{C_n}([a',b']))^2 = o(v_n^2), \quad \text{a.s.} \tag{24}$$

Thus, taking $\delta = 1/35$, from (24) we get,

$$\max_{k\le N}\mathbb{E}_k(F^{C_n}([a',b']))^2 = o(N^{-2/35}) \quad \text{a.s.} \tag{25}$$

3 Convergence of $m_n - \mathbb{E}m_n$

Throughout the rest of the paper we take $v_n = \kappa n^{-\delta}$ with $\delta = \frac{1}{140}$, and some constant $\kappa > 0$. In this section, we verify (6). Since $|m_n(x_1+iv_n) - m_n(x_2+iv_n)| \le |x_1-x_2|v_n^{-2}$ (and from this, the same bound holds for $|\mathbb{E}m_n(x_1+iv_n) - \mathbb{E}m_n(x_2+iv_n)|$), we can prove (6) if we prove that

$$\max_{x\in S_n} Nv_n|m_n(x+iv_n) - \mathbb{E}m_n(x+iv_n)| \to 0, \qquad \text{a.s. as } n\to\infty, \tag{26}$$

for the set S_n consisting of n^2 points equally spaced in $[a,b]$.

Write $D = C_n - zI$, $D_j = D - b_j y_j y_j^*$ (where $y_j = \frac{1}{\sqrt{N}}A_n^{1/2}x_j$), and $D_{jj'} = D_j - b_{j'}y_{j'}y_{j'}^*$, for $j' \ne j$. Then, $m_n = \frac{1}{n}trD^{-1}$. Also, let $\overline{D} = C_n - \overline{z}I$, where $\overline{z}$ is the complex conjugate of z. Note that $\overline{D} = D^*$. Also, let

$$\alpha_j = y_j^* D_j^{-2} y_j - \frac{1}{N}\text{tr}\,(D_j^{-2}A_n), \qquad a_j = \frac{1}{N}\text{tr}\,(D_j^{-2}A_n),$$

$$\beta_j = \frac{1}{1 + b_j y_j^* D_j^{-1} y_j}, \qquad \hat{b}_j = \frac{1}{1 + b_j \frac{1}{N}\mathbb{E}[\text{tr}\,(D_j^{-1}A_n)]},$$

$$\gamma_j = y_j^* D_j^{-1} y_j - \frac{1}{N}\mathbb{E}[\text{tr}\,(D_j^{-1}A_n)], \qquad \hat{\gamma}_j = y_j^* D_j^{-1} y_j - \frac{1}{N}\text{tr}\,(D_j^{-1}A_n).$$

We first derive bounds for the moments of γ_j and $\hat{\gamma}_j$. Integrating first with respect to X_j, that is, conditioning on the set $\{X_i : j \neq i\}$, and using Lemma 3, for all $p \geq 2$,

$$\mathbb{E}|\hat{\gamma}_j|^p \leq K_p N^{-p} \mathbb{E}[\mathrm{tr}\,(A_n^{1/2} D_j^{-1} A_n \overline{D}_j^{-1} A_n^{1/2})]^{p/2} \leq K_p N^{-p/2} v_n^{-p}, \tag{27}$$

where the last step follows from the fact that $\| D_j^{-1} \| \leq v_n^{-1}$, and that $\| A_n \| \leq 1$.

Now, using the fact that $(\mathbb{E}_j - \mathbb{E}_{j-1})[f_n(X_1, \ldots, X_N)]$ (for any bounded f_N) forms a martingale difference sequence w.r.t. the sigma-fields $\mathcal{F}_{j-1}$ generated by columns $\{X_1, \ldots, X_{j-1}\}$, and that $\mathbb{E}_0[\mathrm{tr}\,(D_j^{-1} A_n)] = \mathbb{E}[\mathrm{tr}\,(D_j^{-1} A_n)]$, and $\mathbb{E}_N[\mathrm{tr}\,(D_j^{-1} A_n)] = \mathrm{tr}\,(D_j^{-1} A_n)$, from *Burkholder's inequality* (Lemma 2.2 in Bai and Silverstein (1998))

$$\begin{aligned}
\mathbb{E}|\gamma_j - \hat{\gamma}_j|^p &= N^{-p} \mathbb{E} \left| \sum_{j' \neq j}^{N} (\mathbb{E}_{j'} - \mathbb{E}_{j'-1})[\mathrm{tr}\,(D_j^{-1} A_n)] \right|^P \\
&= N^{-p} \mathbb{E} \left| \sum_{j' \neq j} \mathbb{E}_{j'}[\mathrm{tr}\,(D_j^{-1} - D_{jj'}^{-1}) A_n] - \mathbb{E}_{j'-1}[\mathrm{tr}\,(D_j^{-1} - D_{jj'}^{-1}) A_n] \right|^p \\
&= N^{-p} \mathbb{E} \left| \sum_{j' \neq j} (\mathbb{E}_{j'} - \mathbb{E}_{j'-1}) \left[\frac{b_j y_j^* D_{jj'}^{-1} A_n D_{jj'}^{-1} y_j}{1 + b_j y_j^* D_{jj'}^{-1} y_j} \right] \right|^p \\
&\leq K_p N^{-p} \mathbb{E} \left(\sum_{j' \neq j} \left| (\mathbb{E}_{j'} - \mathbb{E}_{j'-1}) \left[\frac{b_j y_j^* D_{jj'}^{-1} A_n D_{jj'}^{-1} y_j}{1 + b_j y_j^* D_{jj'}^{-1} y_j} \right] \right|^2 \right)^{p/2} \\
&\leq K_p N^{-p/2} v_n^{-p},
\end{aligned} \tag{28}$$

where in the last step we use Lemma 2.10 of Bai and Silverstein (1998) to bound the term within conditional expectations by $\| A_n \| \, v_n^{-1} \leq v_n^{-1}$.

Therefore, from (27) and (28) it follows that for any $p \geq 2$,

$$\mathbb{E}|\gamma_j|^p \leq K_p N^{-p/2} v_n^{-p}. \tag{29}$$

Next, we write

$$
\begin{aligned}
m_n - \mathbb{E}m_n &= \frac{1}{n}\sum_{j=1}^{N}(\mathbb{E}_j - \mathbb{E}_{j-1})[\mathrm{tr}\,(D^{-1})] \\
&= -\frac{1}{n}\sum_{j=1}^{N}(\mathbb{E}_j - \mathbb{E}_{j-1})\left[b_j\frac{y_j^*D_j^{-2}y_j}{1+b_jy_j^*D_j^{-1}y_j}\right] \quad (\text{since } \mathbb{E}_j\mathrm{tr}\,(D_j^{-1}) = \mathbb{E}_{j-1}\mathrm{tr}\,(D_j^{-1})) \\
&= -\frac{1}{n}\sum_{j=1}^{N}(\mathbb{E}_j - \mathbb{E}_{j-1})\left[b_j\frac{y_j^*D_j^{-2}y_j}{1+b_j\frac{1}{N}\mathbb{E}[\mathrm{tr}\,(D_j^{-1}A_n)]}\right] \\
&\quad +\frac{1}{n}\sum_{j=1}^{N}(\mathbb{E}_j - \mathbb{E}_{j-1})\left[b_j^2\frac{y_j^*D_j^{-2}y_j(y_j^*D_j^{-1}y_j - \frac{1}{N}\mathbb{E}[\mathrm{tr}\,(D_j^{-1}A_n)])}{(1+b_j\frac{1}{N}\mathbb{E}[\mathrm{tr}\,(D_j^{-1}A_n)])^2}\right] \\
&\quad -\frac{1}{n}\sum_{j=1}^{N}(\mathbb{E}_j - \mathbb{E}_{j-1})\left[b_j^3\frac{y_j^*D_j^{-2}y_j(y_j^*D_j^{-1}y_j - \frac{1}{N}\mathbb{E}[\mathrm{tr}\,(D_j^{-1}A_n)])^2}{(1+b_j\frac{1}{N}\mathbb{E}[\mathrm{tr}\,(D_j^{-1}A_n)])^2(1+b_jy_j^*D_j^{-1}y_j)}\right] \\
&= -\frac{1}{n}\sum_{j=1}^{N}b_j\hat{b}_j\mathbb{E}_j\alpha_j + \frac{1}{n}\sum_{j=1}^{N}b_j^2\hat{b}_j^2\mathbb{E}_ja_j\hat{\gamma}_j \\
&\quad +\frac{1}{n}\sum_{j=1}^{N}b_j^2\hat{b}_j^2(\mathbb{E}_j - \mathbb{E}_{j-1})[\alpha_j\gamma_j - b_jy_j^*D_j^{-2}y_j\beta_j\gamma_j^2] \\
&= W_1 + W_2 + W_3. \qquad (30)
\end{aligned}
$$

3.1 Boundedness of $\hat{b}_j$

Let

$$p_n^0 = -\frac{1}{z}\int\frac{b}{1+c_nbe_n^0}dF^{B_n}(b) \quad \text{and} \quad \hat{p}_n = -\frac{1}{z}\int\frac{b}{1+c_nb\mathbb{E}(e_n)}dF^{B_n}(b).$$

We have

$$m_n^0 = -\frac{1}{z}\int\frac{1}{ap_n^0+1}dF^{A_n}(a),$$

and

$$e_n^0 = -\frac{1}{z}\int\frac{a}{ap_n^0+1}dF^{A_n}(a),$$

We have then

$$e_n^0 = -\frac{1}{z}\frac{1}{p_n^0}\int\frac{ap_n^0+1-1}{ap_n^0+1}dF^{A_n}(a) = -\frac{1}{zp_n^0} - \frac{m_n^0}{p_n^0}.$$

Therefore

$$e_n^0p_n^0 = -\frac{1}{z} - m_n^0.$$

Suppose $z = z_j \in \mathbb{C}^+ \to x \in [a', b']$ as $j \to \infty$. Then

$$e_n^0(z)p_n^0(z) \to -\frac{1}{x} - m_n^0(x) \in \mathbb{R}.$$

We see that both $\{p_n^0(z_j)\}$ and $\{e_n^0(z_j)\}$ remain bounded, since if, say e_n^0 goes unbounded on some subsequence, p_n^0 would tend to zero on that subsequence, rendering e_n^0 converging to a finite number, a contradiction. Since $\Im m_n^0(x) = 0$, we must have $\lim_{j\to\infty} \Im p_n^0(z_j) = 0$. This in turn implies $\lim_{j\to\infty} \Im e_n^0(z_j) = 0$ as well. Therefore, the measures defining p_n^0 and e_n^0 have derivative 0 for each $x \in [a', b']$, so that (a', b') is outside the support of both these measures, which after considering a slightly larger $\underline{\epsilon}$, this statement extends to $[a', b']$.

From continuity, we have $e_n^0(z) \to e^0(z)$, and consequently, $m_n^0 \to m^0(z)$, and $p_n^0(z) \to p^0(z)$, e^0, m^0, p^0 defined for the limiting empirical distribution, for all $z \in \mathbb{C}^+ \cup [a', b']$. We must have $[-1/p_n^0(a'), -1/p_n^0(b')]$ not intersecting with any of the eigenvalues of A_n (respectively, $[-1/p^0(a'), -1/p^0(b')]$ not intersecting with the support of F^A). Therefore, since $p_n^0(x) \to p^0(x)$, for $x = a', a, b, b'$, and $p^0(a') < p^0(a) < p^0(b) < p^0(b')$, we must have $-1/p_n^0(z)$ uniformly bounded away from the eigenvalues of A_n for all $z = x + iv$, $x \in [a, b]$, and for $v \in [0, v_0]$ for some positive v_0.

Similarly, $-1/(c_n e_n^0)$ is uniformly bounded away from the eigenvalues of B_n for all $z = x + iv$, $x \in [a, b]$, $v \in [0, v_0]$. Therefore, using (16) and arguments analogous to those leading to (21) (now applied to e_n instead of m_n), we have, with $z = x + iv_n$

$$\begin{aligned}
\sup_{x\in[a,b]} |\hat{p}_n(z) - p_n^0(z)| &= \sup_{x\in[c,d]} |e_n^0(z) - \mathbb{E}(e_n(z))| \frac{c_n}{|z|} \left| \int \frac{b^2}{(1 + c_n b e_n^0)(1 + c_n b \mathbb{E}(e_n))} dF^{B_n}(b) \right| \\
\leq \sup_{x\in\mathbb{R}} \frac{K}{v_n} |e_n^0(z) - \mathbb{E}(e_n(z))| &\leq K\mathbb{E}(v_n^{-1} \sup_{x\in\mathbb{R}} |e_n - e_n^0|) \to 0
\end{aligned}$$

as $n \to \infty$. Thus we conclude

$$\sup_{x\in[a,b]} \|(I + \hat{p}_n(z)A_n)^{-1}\| \leq K, \tag{31}$$

and

$$\max_{j\leq N} \sup_{x\in[a,b]} \frac{1}{|(1 + c_n b_j \mathbb{E}(e_n))|} \leq K. \tag{32}$$

Let for $j \neq \underline{j} \leq N$,

$$\hat{b}_j = \frac{1}{1 + c_n b_j n^{-1} \mathbb{E}\mathrm{tr}\,(A_n D_j^{-1})} \quad \text{and} \quad \hat{b}_{j\,\underline{j}} = \frac{1}{1 + c_n b_{\underline{j}} n^{-1} \mathbb{E}(\mathrm{tr}\,(A_n D_{j\,\underline{j}}^{-1}))}.$$

From Lemma 1

$$|(1/n)\mathrm{tr}\,(A_n D_j^{-1}) - e_n| \leq (nv_n)^{-1} \quad \text{and} \quad |(1/n)\mathrm{tr}\,(A_n D_{j\,\underline{j}}^{-1}) - e_n| \leq 2(nv_n)^{-1},$$

so that from (32) we also have for all n large

$$\max_{j\leq N} \sup_{x\in[a,b]} (|\hat{b}_j|, \max_{\underline{j}\neq j} |\hat{b}_{j\,\underline{j}}|) \leq K. \tag{33}$$

3.2 Bounds on W_1, W_2, W_3

Let F_{nj} be the spectral distribution of the matrix $\sum_{k\neq j} b_k y_k y_k^*$. From Lemma 2.12 of Bai and Silverstein (1998), and (25), we get

$$\max_j \mathbb{E}_j(F_{nj}([a', b']))^2 = o(v_n^8) = o(N^{-2/35}), \quad \text{a.s.} \tag{34}$$

Define

$$\mathcal{B}_j = I_{[\mathbb{E}_{j-1} F_{nj}([a',b']) \leq v_n^4] \cap [\mathbb{E}_{j-1}(F_{nj}([a',b']))^2 \leq v_n^8]}\,.$$

Note that, $\mathcal{B}_j = I_{[\mathbb{E}_{j-1}F_{nj}([a',b'])\le v_n^4]\cap[\mathbb{E}_{j-1}(F_{nj}([a',b']))^2\le v_n^8]}$ a.s., and we have

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

Therefore, for any $\varepsilon > 0$,

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

where $\underline{\varepsilon} = \inf_n n\varepsilon/(N\max_{1\le j'\le n} b_{j'}|\hat{b}_{j'}|) > 0$, since $\max_{1\le j\le N}|b_j| \le 1$, and $\max_{1\le j\le N}\sup_{x\in[a,b]}|\hat{b}_j|$ is bounded for all n. Note that, for each $x\in\mathbb{R}$, $\{\mathbb{E}_j(\alpha_j)\mathcal{B}_j\}$ forms a martingale difference sequence.

By Lemma 2.1 in Bai and Silverstein (1998), and Lemma 3, for each $x\in[a,b]$, and $p\ge 2$,

$$\begin{aligned}
&\mathbb{E}\left|v_n\sum_{j=1}^N\mathbb{E}_j(\alpha_j)\mathcal{B}_j\right|^p\\
\le\ &K_p\left(\mathbb{E}\left(\sum_{j=1}^N\mathbb{E}_{j-1}|v_n\mathbb{E}_j(\alpha_j)\mathcal{B}_j|^2\right)^{p/2} + \sum_{j=1}^N\mathbb{E}|v_n\mathbb{E}_j(\alpha_j)\mathcal{B}_j|^p\right)\\
\le\ &K_p\left(\mathbb{E}\left(\sum_{j=1}^N\mathbb{E}_{j-1}v_n^2N^{-2}\mathcal{B}_j\ \mathrm{tr}\ \left(A_n^{1/2}D_j^{-2}A_n\overline{D}_j^{-2}A_n^{1/2}\right)\right)^{p/2} + v_n^p\sum_{j=1}^N\mathbb{E}|\alpha_j|^p\right)\\
\le\ &K_pv_n^pN^{-p}\mathbb{E}\left(\sum_{j=1}^N\mathcal{B}_j\mathbb{E}_{j-1}\mathrm{tr}\,(D_j^{-2}\overline{D}_j^{-2})\right)^{p/2} \qquad (\text{since } \parallel A_n\parallel\le 1)\\
&+\ K_pv_n^pN^{-p}\sum_{j=1}^N\mathbb{E}\left(\mathrm{tr}\,(A_n^{1/2}D_j^{-2}A_n\overline{D}_j^{-2}A_n^{1/2})\right)^{p/2} (\text{ by Lemma 3})\\
\le\ &K_p\left(v_n^pN^{-p}\mathbb{E}\left(\sum_{j=1}^N\mathcal{B}_j\mathbb{E}_{j-1}\mathrm{tr}\,(D_j^{-2}\overline{D}_j^{-2})\right)^{p/2} + v_n^{-p}N^{1-p/2}\right),
\end{aligned}$$

since $\max_j\parallel D_j^{-1}\parallel\le v_n^{-1}$) and $\parallel A_n\parallel\le 1$.

Let λ_{kj} denote the k-th smallest eigenvalue of $\sum_{k'\neq j} b_{k'} y_{k'} y_{k'}^*$. We have, for $x \in [a,b]$

$$
\begin{aligned}
& \sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1} \mathrm{tr}\,(D_j^{-2} \overline{D}_j^{-2}) \\
= \; & \sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1} \left[\sum_{\lambda_{kj} \notin [a',b']} \frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2} + \sum_{\lambda_{kj} \in [a',b']} \frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2} \right] \\
\leq \; & \sum_{j=1}^{N} (n\underline{\epsilon}^{-4} + \mathcal{B}_j v_n^{-4} n \mathbb{E}_{j-1} F_{nj}([a',b'])) \\
\leq \; & KN^2. \qquad (35)
\end{aligned}
$$

Here the last step follows from (34).

Therefore, for $p \geq \frac{70}{34}$,

$$
\mathbb{P}\left(\max_{x \in S_n} \left| v_n \sum_{j=1}^{N} \mathbb{E}_j(\alpha_j) \mathcal{B}_j \right| > \varepsilon \right) \leq K_{p,\varepsilon} n^2 N^{-p/140},
$$

which is summable when $p > 420$. Therefore, by Borel-Cantelli lemma,

$$
\max_{x \in S_n} |W_1| = o(1/Nv_n) \;\; \text{a.s.} \qquad (36)
$$

Next we prove

$$
\max_{x \in S_n} |W_2| = o(1/Nv_n) \;\; \text{a.s.} \qquad (37)
$$

by following similar arguments. First, observing that $\{\mathbb{E}_j(a_j\hat{\gamma}_j)\mathcal{B}_j\}$ forms a martingale difference sequence, and using Lemma 2.1 of Bai and Silverstein (1998), Lemmas 1 and 3, and the fact that $|a_j| \leq \frac{n}{N} v_n^{-2}$, we have,

$$
\begin{aligned}
& \mathbb{E}\left| v_n \sum_{j=1}^{N} \mathbb{E}_j(a_j\hat{\gamma}_j)\mathcal{B}_j \right|^p \\
\leq \; & K_p \left(\mathbb{E}\left(\sum_{j=1}^{N} \mathbb{E}_{j-1} |v_n \mathbb{E}_j(a_j\hat{\gamma}_j)\mathcal{B}_j|^2 \right)^{p/2} + \sum_{j=1}^{N} \mathbb{E}|v_n \mathbb{E}_j(a_j\hat{\gamma}_j)\mathcal{B}_j|^p \right) \\
\leq \; & K_p v_n^p N^{-p} \mathbb{E}\left(\sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1}(|a_j|^2 \mathrm{tr}\,(D_j^{-1}\overline{D}_j^{-1})) \right)^{p/2} \quad \text{(by Lemma 3)} \\
& + K_p v_n^{-p} N^{-p} \sum_{j=1}^{N} \left(\mathrm{tr}\,(D_j^{-1}\overline{D}_j^{-1}) \right)^{p/2} \quad \text{(by Lemma 3, and since } \max_j |a_j| \leq \frac{n}{N} v_n^{-2}) \\
\leq \; & K_p \left(v_n^p N^{-p} \mathbb{E}\left(\sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1}(|a_j|^2 \mathrm{tr}\,(D_j^{-1}\overline{D}_j^{-1})) \right)^{p/2} + v_n^{-2p} N^{1-p/2} \right),
\end{aligned}
$$

since $\max_j \parallel D_j^{-1} \parallel \le v_n^{-1}$, so that $\max_j \mathrm{tr}\,(D_j^{-1}\overline{D}_j^{-1}) \le n v_n^{-2}$.

Moreover, using same notation as before, the fact that $\parallel A_n \parallel \le 1$, and arguing as in the derivation of (35), we have

$$\begin{aligned}
&\sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1}(|a_j|^2 \mathrm{tr}\,(D_j \overline{D}_j^{-1})) \\
\le\ & \sum_{j=1}^{N} \mathcal{B}_j \mathbb{E}_{j-1} N^{-2} n \left[\sum_k \frac{1}{((\lambda_{kj}-x)^2+v_n^2)^2}\right]\left[\sum_k \frac{1}{(\lambda_{kj}-x)^2+v_n^2}\right], \\
& \qquad\qquad \text{(by Cauchy-Schwartz applied to } |a_j|^2) \\
\le\ & \sum_{j=1}^{N} \mathcal{B}_j N^{-2} n \mathbb{E}_{j-1}(n\underline{\epsilon}^{-4} + v_n^{-4} n F_{nj}([a',b']))(n\underline{\epsilon}^{-2} + v_n^{-2} n F_{nj}([a',b'])) \\
\le\ & K N^2.
\end{aligned}$$

Since $\max_j \max\{b_j, \sup_{x\in[a,b]} |\hat{b}_j|\}$ is bounded, for large enough n, we have (37) by arguments similar to the ones used already in the derivation of (36).

Note that, Lemma 1 implies that

$$\max_j \sup_{x\in[a,b]} |b_j y_j^* D_j^{-2} y_j \beta_j| = \max_j \sup_{x\in[a,b]} \left| \frac{b_j y_j^* D_j^{-2} y_j}{1 + b_j y_j^* D_j^{-1} y_j} \right| \le \frac{1}{v_n}. \tag{38}$$

Using Lemma 2.2 of Bai and Silverstein (1998) followed by Hölder's inequality, we have

$$\begin{aligned}
&\mathbb{E}\left| v_n \sum_{j=1}^{N} (\mathbb{E}_j - \mathbb{E}_{j-1})(\alpha_j\gamma_j - b_j y_j^* D_j^{-2} y_j \beta_j \gamma_j^2) \right|^p \\
\le\ & K_p v_n^p N^{p/2-1} \sum_{j=1}^{N} (\mathbb{E}|\alpha_j\gamma_j|^p + v_n^{-p}\mathbb{E}|\gamma_j|^{2p}) \quad \text{(by (38))} \\
\le\ & K_p v_n^p N^{p/2-1} \sum_{j=1}^{N} \left(N^{-p} \left(\mathbb{E}(\mathrm{tr}\,(D_j^{-2}\overline{D}_j^{-2}))^p \right)^{1/2} N^{-p/2} v_n^{-p} + v_n^{-3p} N^{-p} \right) \\
& \qquad \text{(by Cauchy Schwartz and Lemma 3, that } \parallel A_n \parallel \le 1, \text{ and (29))} \\
\le\ & K_p v_n^p N^{p/2} (N^{-p} N^{p/2} v_n^{-2p} N^{-p/2} v_n^{-p} + v_n^{-3p} N^{-p}) \quad (\text{since } \parallel D_j^{-1} \parallel \le v_n^{-1}) \\
\le\ & K_p v_n^{-2p} N^{-p/2}.
\end{aligned}$$

Thus, using arguments as in the proof of (36) and (37), we get

$$\max_{x\in S_n} |W_3| = o(1/N v_n). \tag{39}$$

Hence, (26), and consequently, (6), follow from (36), (37) and (39).

4 Convergence of expected value

In this section we prove (7). Let G_n^0, G^0 denote the distribution functions defining e_n^0, e^0. Then $G_n^0 \xrightarrow{\mathcal{D}} G^0$. We have

$$\int \frac{1}{(\lambda - x)^2} dG^0(\lambda) = \frac{d}{dx} e^0(x)$$

uniformly bounded for $x \in [a,b]$. For λ in either $(-\infty, a']$ or $[b', \infty)$, $\{(\lambda - x)^{-2} : x \in [c,d]\}$ form a uniformly bounded, equicontinuous family of functions in λ. From Billingsley (1968), *Problem 8, p. 17*, we have then

$$\lim_{n\to\infty} \sup_{x\in[a,b]} |\frac{d}{dx} e_n^0(x) - \frac{d}{dx} e^0(x)| = 0.$$

Since for all $x \in [a,b]$, $\lambda \in [a', b']^c$ and positive v

$$\left| \frac{1}{(\lambda - x)^2 + v^2} + \frac{1}{(\lambda - x)^2} \right| \le \frac{v^2}{\underline{\epsilon}^4},$$

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

$$\lim_{n\to\infty} \sup_{x\in[c,d]} \left| \frac{\Im e_n^0(x + iv_n')}{v_n'} - \frac{d}{dx} e_n^0(x) \right| = 0.$$

Therefore, we conclude that

$$\sup_{n, x\in[a,b]} \left| \frac{\Im e_n^0(x + iv_n)}{v_n} \right| \le K. \tag{40}$$

4.1 Martingale decompositions

Let $D = C_n - zI$, $D_j = C_{(j)} - zI$, and for $\underline{j} \ne j$ $D_{j\underline{j}} = C_n - b_j y_j y_j^* - b_{\underline{j}} y_{\underline{j}} y_{\underline{j}}^* - zI$. Let

$$\beta_{j\,\underline{j}} = \frac{1}{1 + b_{\underline{j}} y_{\underline{j}}^* D_{j\,\underline{j}}^{-1} y_{\underline{j}}}.$$

For $j \ne \underline{j} \le N$ let $\lambda_{kj\underline{j}}$ denote the k-th smallest eigenvalue of $C_n - b_j y_j y_j^* - b_{\underline{j}} y_{\underline{j}} y_{\underline{j}}^*$, and let $F_{nj\underline{j}}$ denote the empirical distribution function of this matrix. Using (34) and Lemma 2.12 of Bai and Silverstein (1998) we get

$$\max_{j\ne\underline{j}} \mathbb{E}(F_{nj\underline{j}}[a', b'])^2 = o(v_n^8).$$

Therefore

$$\begin{aligned}
\max_{j\le N} \sup_{x\in[a,b]} \mathbb{E}(\operatorname{tr} D_j^{-1} \overline{D}_j^{-1})^2 &= \max_{j\le N} \sup_{x\in[a,b]} \mathbb{E} \left(\sum_{\lambda_{kj}\notin[a',b']} \frac{1}{(\lambda_{kj} - x)^2 + v_n^2} + \sum_{\lambda_{kj}\in[a',b']} \frac{1}{(\lambda_{kj} - x)^2 + v_n^2} \right)^2 \\
&\le \max_{j\le N} \sup_{x\in[a,b]} \mathbb{E}(n\underline{\epsilon}^{-2} + v_n^{-2} n F_{nj}([a', b']))^2 \le Kn^2,
\end{aligned}$$

and, $\max_{j\le N} \sup_{x\in[c,d]} \mathbb{E}(\operatorname{tr} D_j^{-2} \overline{D}_j^{-2})^2 \le \max_{j\le N} \sup_{x\in[a,b]} \mathbb{E}(n\underline{\epsilon}^{-4} + v_n^{-4} n F_{nj}([a', b']))^2 \le Kn^2.$

The latter implies of course

$$\max_{j\le N} \sup_{x\in[c,d]} \mathbb{E}\operatorname{tr} D_j^{-2} \overline{D}_j^{-2} \le Kn.$$

Similarly,

$$\max_{j\ne\underline{j}} \sup_{x\in[a,b]} \mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-1} \bar{D}_{j\underline{j}}^{-1})^2 \le Kn^2, \quad \text{and} \quad \max_{j\ne\underline{j}} \sup_{x\in[a,b]} \mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-2} \overline{D}_{j\underline{j}}^{-2}) \le Kn.$$

Moreover

$$\max_{j\neq \underline{j}} \sup_{x\in[a,b]} \mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-1}\overline{D}_{j\underline{j}}^{-1})^4 \leq \mathbb{E}(n\underline{\epsilon}^{-2}+v_n^{-2}nF_{nj\underline{j}}([a',b']))^4 \leq Kn^4(\underline{\epsilon}^{-8}+v_n^{-8}\mathbb{E}(F_{nj\underline{j}}([a',b']))^2) \leq Kn^4.$$

Write

$$C_n - zI + zI + z\hat{p}_nA_n = \sum_{j=1}^N b_jy_jy_j^* + z\hat{p}_nA_n.$$

Taking first inverses and then expected values we have

$$\begin{aligned}
&\mathbb{E}(C_n-zI)^{-1}+(zI+z\hat{p}_nA_n)^{-1}\\
=\;& \mathbb{E}\left[\sum_{j=1}^N b_j(C_n-zI)^{-1}y_jy_j^*(zI+z\hat{p}_nA_n)^{-1}+z\hat{p}_nD^{-1}A_n(zI+z\hat{p}_nA_n)^{-1}\right]\\
=\;& \sum_{j=1}^N b_j\left[\mathbb{E}\frac{(C_{(j)}-zI)^{-1}y_jy_j^*(zI+z\hat{p}_nA_n)^{-1}}{1+b_jy_j^*D^{-1}y_j}-\frac{1}{z(1+c_nb_j\mathbb{E}(e_n))}(\mathbb{E}(C_n-zI)^{-1}A_n(I+\hat{p}_nA_n)^{-1}\right].
\end{aligned}$$

Taking the trace on both sides and dividing by n we have

$$\mathbb{E}(m_n(z))-\int\frac{1}{a\int\frac{b}{1+c_nb\mathbb{E}(e_n)}dF^{B_n}(b)-z}dF^{A_n}(a)=\frac{1}{zN}\sum_{j=1}^N b_j\hat{d}_j\equiv\hat{w}_n^m,$$

where

$$\hat{d}_j=\mathbb{E}[\beta_j(1/n)x_j^*A_n^{1/2}(I+\hat{p}_nA_n)^{-1}D_j^{-1}A_n^{1/2}x_j]-\frac{(1/n)\operatorname{tr}\mathbb{E}[D^{-1}]A_n(I+\hat{p}_nA_n)^{-1}}{(1+c_nb_j\mathbb{E}(e_n))}.$$

Multiplying both sides of the above matrix identity by A_n, and then taking traces and dividing by n, we find

$$\mathbb{E}(e_n(z))-\int\frac{a}{a\int\frac{b}{1+c_nb\mathbb{E}(e_n)}dF^{B_n}(b)-z}dF^{A_n}(a)=\frac{1}{zN}\sum_{j=1}^N b_j\hat{d}_j^e\equiv\hat{w}_n^e,$$

where

$$\hat{d}_j^e=\mathbb{E}[\beta_j(1/n)x_j^*A_n^{1/2}(I+\hat{p}_nA_n)^{-1}A_nD_j^{-1}A_n^{1/2}x_j]-\frac{(1/n)\operatorname{tr}A_n\mathbb{E}[D^{-1}]A_n(I+\hat{p}_nA_n)^{-1}}{(1+c_nb_j\mathbb{E}(e_n))}.$$

Again, we let E_n denote either A_n or I_n. We first show that

$$n^{-1}\max_{j\leq N}\sup_{x\in[a,b]}|(\operatorname{tr}E_n\mathbb{E}[D^{-1}]A_n(I+\hat{p}_nA_n)^{-1}-\operatorname{tr}E_n\mathbb{E}[D_j^{-1}]A_n(I+\hat{p}_nA_n)^{-1}|=O(n^{-1}). \qquad (41)$$

Using $\beta_j=\hat{b}_j-b_j\beta_j\hat{b}_j\gamma_j$, (3.3) of Bai and Silverstein (1998), (29), (31), and (33) we conclude that the left hand side of (41) becomes

$$
\begin{aligned}
&= n^{-1}\max_{j\le N} b_j \sup_{x\in[a,b]} |\mathbb{E}[\beta_j y_j^* D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} y_j]| \\
&\le n^{-1}\max_{j\le N} b_j \sup_{x\in[a,b]} (|\hat{b}_j||\mathbb{E}[y_j^* D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} y_j]| \\
&\quad + b_j|\hat{b}_j||\mathbb{E}[\beta_j\gamma_j y^* D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} y_j]|) \\
&\le K n^{-1}\max_{j\le N}\sup_{x\in[a,b]} (N^{-1}|\mathbb{E}[\operatorname{tr} A_n^{1/2} D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} A_n^{1/2}]| \\
&\quad + v_n^{-1}(\mathbb{E}|\gamma_j|^2)^{1/2}(\mathbb{E}|y_j^* D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} y_j|^2)^{1/2} \\
&\le K n^{-1}\max_{j\le N}\sup_{x\in[a,b]} (N^{-1}\mathbb{E}[\operatorname{tr} D_j^{-1}\overline{D}_j^{-1}]) \\
&\quad + v_n^{-1}N^{-1/2}v_n^{-1}N^{-1}(\mathbb{E}[\operatorname{tr} D_j^{-2}\overline{D}_j^{-2}] + \mathbb{E}(\operatorname{tr} D_j^{-1}\overline{D}_j^{-1})^2)^{1/2} \le K n^{-1}.
\end{aligned}
$$

Thus (41) holds.

From Lemma 3 and (31) we get

$$
\begin{aligned}
&\max_{j\le N}\sup_{x\in[a,b]} \mathbb{E}|(1/n)x_j^* A_n^{1/2}(I+\hat{p}_n A_n)^{-1} E_n D_j^{-1} A_n^{1/2} x_j - (1/n)\operatorname{tr} E_n D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1}|^2 \\
&\le K n^{-2}\max_{j\le N}\sup_{x\in[c,d]} \mathbb{E}[\operatorname{tr} D_j^{-1}\overline{D}_j^{-1}] \le K n^{-1}. \qquad (42)
\end{aligned}
$$

We next show

$$
\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\mathbb{E}|\operatorname{tr} E_n D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1} - \operatorname{tr} E_n \mathbb{E}[D_j^{-1}] A_n (I+\hat{p}_n A_n)^{-1}|^2 \le K n^{-1}. \qquad (43)
$$

Using (3.3) of Bai and Silverstein (1998), and the fact that $\beta_{j\underline{j}} = \hat{b}_{j\underline{j}} - b_{\underline{j}}\hat{b}_{j\underline{j}}\beta_{j\underline{j}}\gamma_{j\underline{j}}$, the hand left side of (43) becomes

$$
\begin{aligned}
&= \max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} \mathbb{E}|(\mathbb{E}_{\underline{j}} - \mathbb{E}_{\underline{j}-1})\operatorname{tr} E_n D_j^{-1} A_n (I+\hat{p}_n A_n)^{-1}|^2 \\
&\le 2\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} b_{\underline{j}}^2 \mathbb{E}|\beta_{j\underline{j}} y_{\underline{j}}^* D_{j\underline{j}}^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_{j\underline{j}}^{-1} y_{\underline{j}}|^2 \\
&= 2\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} b_{\underline{j}}^2 \mathbb{E}|(\hat{b}_{j\underline{j}} - b_{\underline{j}}\hat{b}_{j\underline{j}}\beta_{j\underline{j}}\gamma_{j\underline{j}}) y_{\underline{j}}^* D_{j\underline{j}}^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_{j\underline{j}}^{-1} y_{\underline{j}}|^2 \\
&\le K\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} \Big[\mathbb{E}|y_{\underline{j}}^* D_{j\underline{j}}^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_{j\underline{j}}^{-1} y_{\underline{j}}|^2 \\
&\qquad + v_n^{-2}(\mathbb{E}|\gamma_{j\underline{j}}|^4 \mathbb{E}|y_{\underline{j}}^* D_{j\underline{j}}^{-1} A_n (I+\hat{p}_n A_n)^{-1} E_n D_{j\underline{j}}^{-1} y_{\underline{j}}|^4)^{1/2}\Big] \\
&\le K\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} b_{\underline{j}}^2 n^{-2}\Big[\mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-2}\overline{D}_{j\underline{j}}^{-2}) + \mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-1}\overline{D}_{j\underline{j}}^{-1})^2 \\
&\qquad + v_n^{-2}n^{-1}v_n^{-2}(\mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-2}\overline{D}_{j\underline{j}}^{-2})^2 + \mathbb{E}(\operatorname{tr} D_{j\underline{j}}^{-1}\overline{D}_{j\underline{j}}^{-1})^4)^{1/2}\Big] \\
&\le K\max_{j\le N}\sup_{x\in[a,b]} n^{-2}\sum_{\underline{j}\ne j} n^{-2}(n + n^2 + v_n^{-4}n^{-1}(n^2+n^4)^{1/2}) \\
&\le K n^{-1}.
\end{aligned}
$$

So (43) is true.

We get the same bound when $(I + \hat{p}_n A_n)^{-1}$ is removed from the expressions, that is, we also have

$$\max_{j\leq N} \sup_{x\in[a,b]} \mathbb{E}|\gamma_j - \hat{\gamma}_j|^2 \leq Kn^{-1}.$$

Moreover, using (27),

$$\max_{j\leq N} \sup_{x\in[a,b]} \mathbb{E}|\hat{\gamma}_j|^2 \leq Kn^{-2}\mathbb{E}[\mathrm{tr}\, D_j^{-1}\overline{D}_j^{-1}] \leq Kn^{-1}.$$

Thus

$$\max_{j\leq N} \sup_{x\in[a,b]} \mathbb{E}|\gamma_j|^2 \leq Kn^{-1}.$$

Therefore, with $\hat{d}_j^{em}$ denoting either $\hat{d}_j$ or $\hat{d}_j^e$, we have, using Lemma 1, (41), (42), (29), and the fact that $\beta_j = \hat{b}_j - \hat{b}_j^2\gamma_j + \hat{b}_j^2\beta_j\gamma_j^2$,

$$\begin{aligned}
&\sup_{x\in[a,b]} \left|\frac{1}{zN}\sum_{j=1}^N b_j \hat{d}_j^{em}\right| \\
\leq\ & Kn^{-1} + \max_{j\leq N}\sup_{x\in[a,b]} \left|\mathbb{E}[\beta_j(1/n)x_j^* A_n^{1/2}(I+\hat{p}_n A_n)^{-1}E_n D_j^{-1}A_n^{1/2}x_j] - \frac{(1/n)\mathrm{tr}\, E_n\mathbb{E}[D_j^{-1}]A_n(I+\hat{p}_nA_n)^{-1}}{1+c_nb_j\mathbb{E}(e_n)}\right| \\
\leq\ & Kn^{-1} + K\max_{j\leq N}\sup_{x\in[a,b]} \left(\mathbb{E}|\beta_j - \hat{b}_j + \frac{c_nb_j\hat{b}_j}{1+c_nb_j\mathbb{E}(e_n)}\mathbb{E}(e_n - (1/n)\mathrm{tr}\, A_n^{1/2}D_j^{-1}A_n^{1/2})|(1/n)(\mathbb{E}(\mathrm{tr}\, D_j^{-1}\overline{D}_j^{-1}))^{1/2}\right. \\
& \left. + |\mathbb{E}[\beta_j((1/n)x_j^*A_n^{1/2}(I+\hat{p}_nA_n)^{-1}E_nD_j^{-1}A_n^{1/2}x_j - (1/n)\mathrm{tr}\, E_n\mathbb{E}(D_j^{-1})A_n(I+\hat{p}_nA_n)^{-1})]|\right) \\
\leq\ & Kn^{-1} + K\max_{j\leq N}\sup_{x\in[a,b]} \left(||\hat{b}_j^2|\mathbb{E}|\gamma_j - \beta_j\gamma_j^2| + v_n^{-1}n^{-1}|n^{-1/2}\right. \\
& \left. + |\hat{b}_j|^2|\mathbb{E}[(\gamma_j - \beta_j\gamma_j^2)((1/n)x_j^*A_n^{1/2}(I+\hat{p}_nA_n)^{-1}E_nD_j^{-1}A_n^{1/2}x_j - (1/n)\mathrm{tr}\, E_n\mathbb{E}(D_j^{-1})A_n(I+\hat{p}_nA_n)^{-1})]|\right) \\
\leq\ & K(n^{-1} + \max_{j\leq N}\sup_{x\in[a,b]} (\mathbb{E}|\gamma_j|^2 + v_n^{-2}\mathbb{E}|\gamma_j|^4)^{1/2})n^{-1/2}) \\
\leq\ & K(n^{-1} + (n^{-1} + v_n^{-2}n^{-2}v_n^{-4})^{1/2}n^{-1/2}) \\
\leq\ & Kn^{-1}.
\end{aligned}$$

As before, we have

$$\mathbb{E}(e_n) - e_n^0 = (\mathbb{E}(e_n) - e_n^0)\gamma_n + \hat{w}_n^e$$

where, after inserting $\hat{p}_n$ and p_n^0,

$$|\gamma_n| \leq \left(\int \frac{c_na^2\int \frac{b^2}{|1+c_nb\mathbb{E}(e_n)|^2}dF^{B_n}(b)}{|z|^2|a\hat{p}_n+1|^2}dF^{A_n}(a)\right)^{1/2}\left(\int \frac{c_na^2\int \frac{b^2}{|1+c_nbe_n^0|^2}dF^{B_n}(b)}{|z|^2|ap_n^0+1|^2}dF^{A_n}(a)\right)^{1/2}. \tag{44}$$

Writing again $e_2^0 = e_2^0\alpha + v_n\beta$ we have by (32) and (40)

$$\sup_{x\in[a,b]}\frac{e_2^0}{v_n\beta} \leq \sup_{x\in[a,b]}\frac{e_2^0}{v_n}c_n\int\frac{b^2}{|1+c_nbe_n^0|^2}dF^{B_n}(b) \leq K.$$

Therefore

$$\sup_{x\in[a,b]} \int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b)}{|z|^2|ap_n^0+1|^2} dF^{A_n}(a) = \sup_{x\in[a,b]} \frac{e_2^0/(v_n\beta)}{(e_2^0/(v_n\beta))+1}$$

is uniformly bounded away from 1. Moreover, from continuity and the uniform convergence of $\mathbb{E}(e_n)$ and $\hat{p}_n$, we must have that the supremum over all $x \in [a,b]$ of the first factor on the right hand side of (44) is also uniformly bounded away from 1 for all n large. We therefore have

$$\sup_{x\in[a,b]} |\mathbb{E}(e_n) - e_n^0| = O(n^{-1}).$$

Again

$$\begin{aligned} & |\mathbb{E}(m_n) - m_n^0| \\ \leq\ & |\mathbb{E}(e_n) - e_n^0| \left(\int \frac{c_n \int \frac{b^2}{|1+c_n b\mathbb{E}(e_n)|^2} dF^{B_n}(b)}{|z|^2|a\hat{p}_n+1|^2} dF^{A_n}(a) \right)^{1/2} \left(\int \frac{c_n a^2 \int \frac{b^2}{|1+c_n b e_n^0|^2} dF^{B_n}(b)}{|z|^2|a\hat{p}_n^0+1|^2} dF^{A_n}(a) \right)^{1/2} \\ & + |\hat{w}_n^m|. \end{aligned}$$

The second factor on the right is of course bounded by 1, and from (31) and (32), the first factor is bounded, uniformly for $x \in [a,b]$. We conclude that (7) holds

Thus combining the results of this section and the previous section, we arrive at (3), and along with section 6 of Bai and Silverstein (1998), this completes the proof of Theorem 1.

Appendix : mathematical tools

Lemma 1 (Lemma 2.6 of Silverstein and Bai (1995)): Let $z \in \mathbb{C}^+$ with $v = Im\, z$, A and B $n\times n$ with B Hermitian, and $r \in \mathbb{C}^n$. Then

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

Lemma 2 (Lemma 2.3 of Silverstein (1995)): For $z = x+iv \in \mathbb{C}^+$ l et $m_1(z)$, $m_2(z)$ be Stieltjes transforms of any two p.d.f.'s with respective total measure M_1, M_2, A, B, and C $n\times n$ with A Hermitian non-negative definite, and $r \in \mathbb{C}^n$. Then

(a)

$$\|(m_1(z)A+I)^{-1}\| \leq \max(4M_1\|A\|/v, 2)$$

(b)

$$\begin{aligned} & |\mathrm{tr}\, B((m_1(z)A+I)^{-1} - (m_2(z)A+I)^{-1})| \\ \leq\ & |m_2(z)-m_1(z)|n\|B\|\,\|A\| \max(4M_1\|A\|/v,2)\max(4M_2\|A\|/v,2) \end{aligned}$$

(c)

$$\begin{aligned} & |r^*B(m_1(z)A+I)^{-1}Cr - r^*B(m_2(z)A+I)^{-1}Cr \\ \leq\ & |m_2(z)-m_1(z)|\|r\|^2\|A\|\|B\| \max(4M_1\|A\|/v,2)\max(4M_2\|A\|/v,2) \end{aligned}$$

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

Lemma 3 (Lemma 2.7 of Bai and Silverstein (1998)) : For $X = (X_1, \ldots, X_n)^T$ i.i.d. standardized and bounded entries, C $n \times n$ matrix we have for any $p \geq 2$

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

where K_p depends on the distribution of $X_{\cdot 1}$.

Lemma 4 ((3.1) of Bai and Silverstein (1998)) : The largest eigenvalue of C_n, denoted by $\lambda_{\max}$, satisfies

$$\mathbb{P}(\lambda_{\max} > K) = o(n^{-t})$$

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

Reference

1. Bai, Z. D. and Silverstein, J. W. (1998) : No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices.

2. Baik, J. and Silverstein, J. W. (2006) : Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis* **97**, 1382-1408.

3. Baik, J., Ben Arous, G. and Péché, S. (2005) : Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *Annals of Probability* **35**, 1643-1697.

4. Billingsley, P. (1968) : *Convergence of Probability Measures.* Wiley, New York.

5. Burda, Z., Jurkiewicz, J. and Waclaw, B. (2005) : Spectral moments of correlated Wishart matrices. *Phiscal Review E* **71**, 026111.

6. de Monvel, A. B., Khorunzhy, A. and Vasilchuk, V. (1996) : Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Processes and Related Fields* **2**, 607-636.

7. Dozier, B. and Silverstein, J. W. (2007) Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. *Journal of Multivariate Analysis.* To appear.

8. El Karoui, N. (2007) : Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Annals of Probability* **35**, 663-714.

9. Krantz, S. G. (2001) : *Function Theory of Several Complex Variables, 2nd Ed.* American Mathematical Society.

10. Krein, M. K. and Nudelman, A. A. (1997) : *The Markov Moment Problem and Extremal Problems.* American Mathematical Society, Providence, RI.

11. Paul, D. (2007) : Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica.* To appear.

12. Silverstein, J. W. and Bai, Z. D. (1995) : On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis* **54**, 175-192.

13. Silverstein, J. W. and Choi, S. I. (1995) : Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis* **54**, 295-309.

14. Silverstein, J. W. and Combettes, P. L. (1992) : Signal detection via spectral theory of large dimensional random matrices. *IEEE Transactions on Signal Processing* **40**, 2100-2105.

15. von Storch, H. and Zwiers, F. W. (1999) : *Statistical Analysis in Climate Research.* Cambridge University Press.

16. Tulino, A. M. and Verdú, S. (2004) : Random matrix theory and wireless communications. *Foundations and Trends in Communication and Information Theory* **1**, 1-182.

17. Verdú, S. (2002) : Spectral Efficiency in the Wideband Regime. *IEEE Transactions on Information Theory* **48**, 1319-1343.

18. Yin, Y. Q., Bai, Z. D. and Krishnnaiah, P. R. (1988) : On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory and Related Fields* **78**, 509-521.

19. Zhang, L. (2006) : *Spectral Analysis of Large Dimensional Random Matrices.* Ph. D. Thesis. National University of Singapore.