Analysis of the limiting spectral distribution of large dimensional general information-plus-noise type matrices

Huanchao Zhou[a], Jiang Hu[b], Zhidong Bai[c], Jack W. Silverstein[d]

[a]KLASMOE and School of Mathematics and Statistics, Northeast Normal University, China
[b]KLASMOE and School of Mathematics and Statistics, Northeast Normal University, China
[c]KLASMOE and School of Mathematics and Statistics, Northeast Normal University, China
[d]Department of Mathematics, North Carolina State University, USA

Abstract

In this paper, we derive the analytical behavior of the limiting spectral distribution of non-central covariance matrices of the "general information-plus-noise" type, as studied in [14]. Through the equation defining its Stieltjes transform, it is shown that the limiting distribution has a continuous derivative away from zero, the derivative being analytic wherever it is positive, and we show the determination criterion for its support. We also extend the result in [14] to allow for all possible ratios of row to column of the underlying random matrix.

Keywords: Random matrix, LSD, Stieltjes transform, Information-plus-noise matrix.
2020 MSC: Primary 62H12, Secondary 62F12

1. Introduction

The key point considered in this paper is to characterize the limiting spectral distribution (LSD) of non-central covariance matrices of the form

$$B_n = \frac{1}{N}(R_n + T^{1/2}X_n)(R_n + T^{1/2}X_n)^*, \tag{1}$$

where * stands for the complex conjugate transpose, X_n is an $n \times N$ random matrix with independent standardized entries, $T_n^{1/2}$, $n \times n$, is the nonnegative definite square square root of the $n \times n$ nonrandom, nonnegative definite matrix T_n, R_n, $n \times N$ is nonrandom, and n, N are such that, as $\min(n, N) \to \infty$, $c_n \equiv n/N \to c > 0$.

We begin with some background information. It is well known that the LSD for random matrices has constituted a basic part of large dimensional random matrices theory (LDRMT). Since the famous semicircular law and M-P law were established in [12] and [8] respectively, many researchers have contributed to its subsequent development. One of the most extensively investigated in LDRMT is the so-called sample covariance type matrix taking the form $M_n = \frac{1}{N}T_n^{\frac{1}{2}}X_nX_n^*T_n^{\frac{1}{2}}$, where initially the entries of X_n are i.i.d Results on the LSD of the sample covariance type matrix, can be found in [10, 11, 13].

The majority of known results for the sample covariance type matrix are under the central condition, that is, the entries of X_n are zero mean (which can be extended to allowing the entries to have a common mean). Actually, the large non-central random matrices also have a significance that may be considered as an extension of non-central Wishart matrices, an important random matrix in multivariate linear regression under a non-null hypothesis. Due to its application in wireless communication and signal processing, it is also called an information-plus-noise matrix. The first work on this direction was done by Dozier and Silverstein [5]. The model they considered is

$$D_n = \frac{1}{N}(R_n + \sigma X_n)(R_n + \sigma X_n)^*,$$

representing the information, and the matrix σX_n is considered as additive noise. They studied the LSD of D_n and further investigated the analytical properties of the LSD, as well as characterization of its support set in [4]. Loubaton and Vallet [7] considered the model D_n. They assume R_n is of fixed rank and X_n with means zero and covariances $\frac{1}{N}$, and they characterized the a.s. limits of the few largest eigenvalues of D_n. Bai and Silverstein [1] showed that, with additional assumptons on R_n, in any closed interval outside the support of the LSD of D_n, with probability one, there is no eigenvalues of D_n falling in this interval for all large n. Capitaine [3] further proved the exact separation problem.

Inspired by the observations that the noise σX_n has covariance matrix $\sigma^2 I$ (I denoting the $n \times n$ identity matrix), by allowing not only a multiple of the identity but also other nonnegative definite matrices, extending the possibilities of the covariance matrix of the noise, Zhou and Bai et al [14] have further extended Dozier and Silverstein's model D_n to the information-plus-correlated-noise matrices

$$C_n = \frac{1}{N} T_n^{\frac{1}{2}} (R_n + X_n)(R_n + X_n)^* T_n^{\frac{1}{2}},$$

and proved that, almost surely, the empirical spectral distribution (ESD), F_n, of the matrix C_n, defined by $F_n(x) = (1/n)$(number of eigenvalues of $C_n \le x$) converges to a nonrandom LSD F, under certain conditions. Before continuing, it is necessary to introduce the major tool used in establishing almost sure weak convergence of ESD's. It is the *Stieltjes transform* of the empirical distribution of eigenvalues of a matrix. For any finite measure μ on $\mathbb{R}$ with distribution function F, the Stieltjes transform of μ is defined as

$$m_\mu(z) = m_F(z) = \int \frac{1}{\lambda - z} \mathrm{d}\mu(\lambda) = \int \frac{1}{\lambda - z} \mathrm{d}F(\lambda), \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}. \tag{2}$$

An important property is the inversion formula: F can be obtained by

$$F(b) - F(a) = \frac{1}{\pi} \lim_{v \to 0^+} \int_a^b \Im s_F(x + iv) \mathrm{d}x,$$

where a, b are continuity points of F.

The main reason Stieltjes transforms are used stem from the fact that for any $n \times n$ matrix A with real eigenvalues, with $m_A(z)$ denoting the Stieltjes transform of the ESD of A, we have

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

Some other important properties of Stieltjes transforms can be found in Lemma 2.2 of [9] and Theorems A.2, A.4, A.5 of [6]: If f is analytic on $\mathbb{C}^+$, both $f(z)$ and $zf(z)$ map $\mathbb{C}^+$ into $\mathbb{C}^+$, and there is a $\theta \in (0, \pi/2)$ for which $zf(z) \to c$, finite, as $z \to \infty$ restricted to $\{w \in \mathbb{C} : \theta < \arg w < \pi - \theta\}$, then c is real, $c < 0$ and f is the Stieltjes transform of a measure on the nonnegative reals with total mass $-c$. Notice that any Stieltjes transform maps $\mathbb{C}^+$ to $\mathbb{C}^+$.

Notice also that for any finite measure μ on the nonnegative reals, we also have $zm_\mu(z)$ maps $\mathbb{C}^+$ to $\mathbb{C}^+$ since with $z = x + iv$,

$$\Im z m_\mu(z) = \int \frac{\lambda v}{|\lambda - z|^2} \mathrm{d}\mu(\lambda) > 0.$$

For any sequence $\{\mu_n\}$ of probability measures, if it is shown that there is a countable number of $z_j \in \mathbb{C}^+$ possessing an accumulation point for which the Stieltjes transform

$$m_{\mu_n}(z) \equiv \int \frac{1}{\lambda - z} \mathrm{d}\mu_n(\lambda)$$

converges for each of these z_j, then the μ_n converge vaguely to a subprobability measure μ, and weakly if μ is a probability measure. Indeed, for any vaguely convergent subsequence μ_{n_i}, say, to $\hat{\mu}$, since the real and imaginary parts of $1/(\lambda - z)$, as functions of λ, are continuous and vanish at $\pm\infty$, we have for these z_j

$$\int \frac{1}{\lambda - z_j} \mathrm{d}\mu_{n_i} \to \int \frac{1}{\lambda - z_j} \mathrm{d}\hat{\mu} \quad \text{as } i \to \infty,$$

and this limiting analytic function is uniquely determined by the values it takes on the z_j. Therefore we have vague convergence of μ_n to a unique μ, and the convergence is weak if μ is proper.

We are now able to state the result in Zhou and Bai et al [14]. Under the assumptions:

(a) $c_n = n/N \to c > 0$ as $\min\{n, N\} \to \infty$.

(b) The entries of X_n are independent standardized random variables satisfying

$$\frac{1}{\eta^2 nN} \sum_{jk} \mathbf{E}(|x_{jk}^{(n)}|^2 I(|x_{jk}^{(n)}| \geq \eta \sqrt{n})) \to 0.$$

(c) T_n is $n \times n$ nonrandom nonnegative definite, R_n is $n \times N$ nonrandom, with T_n and $R_n R_n^*$ commutative.

(d) As $\min\{n, N\} \to \infty$, the two-dimensional distribution function $H_n(s,t)$, where $H_n(s,t) = n^{-1} \sum_{i=1}^n I(s_i \leq s, t_i \leq t)$ converges weakly to a nonrandom probability distribution $H(s,t)$ almost surely, where s_i, t_i are the paired eigenvalues of $\frac{1}{N} R_n R_n^*$ and T_n, respectively,

it is shown that, with probability one, for any $z \in \mathbb{C}^+$, any limit of a subsequence of the Stieltjes transforms $m^{C_n}(z)$ must converge to a number $m \in \mathbb{C}^+$ satisfying the equations

$$\begin{cases} m = \displaystyle\int \frac{\mathrm{d}H(s,t)}{\frac{st}{1+cg} - (1 + cmt)z + t(1-c)}, \\ g = \displaystyle\int \frac{t\mathrm{d}H(s,t)}{\frac{st}{1+cg} - (1 + cmt)z + t(1-c)}, \end{cases} \tag{3}$$

where $g \in \mathbb{C}^+$. We remark here that g is the limit (when it exists) of $g_n = (1/n)\mathrm{tr}(C_n - zI)^{-1} T_n$.

Moreover, it is shown that for $c \leq 1$ there is only one pair (m, g), each in $\mathbb{C}^+$ satisfying (3) for all $z \in \mathbb{C}^+$, which implies, with probability one, F_n converges vaguely to a distribution F having Stieltjes transform $m = m(z)$ satisfying (3) for all $z \in \mathbb{C}^+$.

We will presently prove that the convergence is actually weak, that is, F is a probability distribution function, a fact omitted in Zhou and Bai et al [14]. Let $U_n \Lambda_n U_n^*$ be the spectral decomposition of C_n. As mentioned previously, the quantity g is the limit of

$$g_n(z) = \frac{1}{n}\mathrm{tr}(C_n - zI)T_n = \frac{1}{n}\mathrm{tr} T_n^{1/2} U_n (\Lambda_n - zI)^{-1} U_n^* T_n^{1/2} = \frac{1}{n} \sum_{i=1}^n \frac{|T_n^{1/2} U_n|_{ii}^2}{\lambda_i - z},$$

where the λ_i are the eigenvalues of C_n. We see that g_n is the Stieltjes transform of a measure which places mass on the eigenvalues of C_n with total mass $(1/n)\mathrm{tr} T_n$. Therefore, as with $m(z)$, $g_n(z)$ is the Stieltjes transform of a finite measure on the nonnegative reals. Then necessarily, the limit $g(z)$, when it exists, is such that both $\Im g(z)$ and $\Im z g(z)$ are nonnegative.

It follows that, with $z = x + iv$, $v > 0$, the absolute value of the imaginary part of the denominator of the integrand in both equations in (3) is greater than v, so that for fixed x and all $v \geq v_0 > 0$ z times the integrand in the first equation in (3) is bounded. For each s, t as $v \uparrow \infty$ (x fixed), z times the integrand in the first equation of (3) converges to -1. Therefore, from the dominated convergence theorem we have $zm(z) \to -1$ as $v \uparrow \infty$, which implies the limiting F is a probability distribution function, so that the convergence is weak.

We turn now to the aim of this paper, namely to investigate the limiting properties of the ensemble (1). It is clear that C_n is of this form, so it seems B_n is a more general ensemble. However, it is shown in the appendix that the results on C_n can be used to show the limiting behavior of the ESD of the eigenvalues of B_n. Restricting T_n to remain nonsingular, we can infer the limiting behavior. Indeed, write

$$B_n = \frac{1}{N} T_n^{1/2} (T_n^{-1/2} R_n + X_n)(T_n^{-1/2} R_n + X_n)^* T^{1/2}, \tag{4}$$

which is in the form of C_n with the two dimensional distribution function

$$H_n'(s,t) = n^{-1} \sum_{i=1}^n I(t_i^{-1} s_i \leq s, t_i \leq t). \tag{5}$$

Writing $H_n(s,t) = n^{-1}\sum_{i=1}^n I(s_i \le s, t_i \le t)$, if H_n converges weakly to H and if $H(\infty, 0) = 0$, then H_n and H'_n can be interpreted in terms of distribution functions of random variables x_n, y_n where $(x_n, y_n) = (s_i, t_i)$ with probability $1/n$ and $H_n(s,t) = P(x_n \le s, y_n \le t)$, $H'_n(s,t) = P(y_n^{-1}x_n \le s, y_n \le t)$. We can think of x_n, y_n converging in distribution to random variables x, y, with joint distribution function H. We have then $y_n^{-1}x_n, y_n$ converging in distribution to random variables $y^{-1}x, y$, whose distribution function is the weak limit of H'_n, say H'. With H' placed in the equations in (3), the integrals can be interpreted as expected values of functions of x and y, resulting in the following pair:

$$\begin{cases} m = \displaystyle\int \frac{\mathrm{d}H(s,t)}{\frac{s}{1+cg} - (1+cmt)z + t(1-c)}, \\ g = \displaystyle\int \frac{t\mathrm{d}H(s,t)}{\frac{s}{1+cg} - (1+cmt)z + t(1-c)}. \end{cases} \tag{6}$$

Using truncation techniques, the appendix will show how results in Zhou and Bai et al [14] extend to B_n. It will also include a complete proof of almost sure weak convergence of the ESD for all posiive c.

Thus, under assumptions (a) – (d), for all $c > 0$, with probability one, the ESD of (1) converges weakly to F, nonrandom, with Stieltjes transform m_F satisfying (6).

This paper shows that, for $c \le 1$ and conditions on the limiting H, including the assumption that H has bounded support, the limiting distribution function F has, for $x \ne 0$ a continuous density, analytic in its support, along with a detailed analysis of how the support of F can be determined. This is shown in the next section.

2. Existence of a density

It is more convenient to consider the limiting ESD of the eigenvalues of the $N \times N$ matrix

$$\underline{B}_n = \frac{1}{n}(R_n + T_n^{\frac{1}{2}}X_n)^*(R_n + T_n^{\frac{1}{2}}X_n) \tag{7}$$

The eigenvalues of B_n are same as those of $\underline{B}_n$. except for $|n-N|$ zero eigenvalues. It follows that their ESDs and Stieltjes transforms have the following relations

$$\begin{aligned} F^{\underline{B}_n} &= \left(1 - \frac{n}{N}\right)I_{[0,\infty)} + \frac{n}{N}F^{B_n}. \\ m^{\underline{B}_n}(z) &= -\frac{\left(1-\frac{n}{N}\right)}{z} + \frac{n}{N}m^{B_n}. \end{aligned} \tag{8}$$

Then making a variable transformation

$$\begin{aligned} \underline{m}(z) &= -\frac{1-c}{z} + cm(z), \\ \underline{g}(z) &= -\frac{1}{z(1+cg(z))}, \end{aligned} \tag{9}$$

where $\underline{m}(z)$ is the Stieltjes transform of the LSD of $F^{\underline{B}_n}$. then the equations in (6) become

$$\begin{aligned} z &= -\frac{1-c}{\underline{m}} - \frac{c}{\underline{m}}\int \frac{\mathrm{d}H(s,t)}{1 + s\underline{g}(z) + t\underline{m}(z)}, \\ z &= -\frac{1}{\underline{g}} + c\int \frac{t\mathrm{d}H(s,t)}{1 + s\underline{g}(z) + t\underline{m}(z)}. \end{aligned} \tag{10}$$

Let $\underline{F}$ denote the almost sure limiting distribution function of the eigenvalues of $\underline{B}_n$, with Stieltjes transform

$$m_{\underline{F}}(z) = \int \frac{1}{\lambda - z}\mathrm{d}\underline{F}(\lambda), \quad z \in \mathbb{C}^+.$$

We begin with deriving an important identity. Write $\underline{m} = \underline{m}_1 + i\underline{m}_2$, $\underline{g} = \underline{g}_1 + i\underline{g}_2$, $z\underline{m} = (z\underline{m})_1 + i(z\underline{m})_2$, and $z\underline{g} = (z\underline{g})_1 + i(z\underline{g})_2$. Fix $z = x + iv \in \mathbb{C}^+$. Multiplying by $\underline{m}$ on both sides to the first equation of (10) and comparing the imaginary part of the resulting equation and that of the second equation, we obtain

$$(z\underline{m})_2 = cA_1\underline{g}_2 + cB_1\underline{m}_2$$

$$v = \frac{\underline{g}_2}{|\underline{g}|^2} - (cA_2\underline{g}_2 + cB_2\underline{m}_2).$$

where

$$\begin{aligned} A_j &= \int \frac{st^{j-1}\mathrm{d}H(s,t)}{|1+s\underline{g}+t\underline{m}|^2}, j = 1, 2, \\ B_j &= \int \frac{t^j\mathrm{d}H(s,t)}{|1+s\underline{g}+t\underline{m}|^2}, j = 0, 1, 2. \end{aligned} \tag{11}$$

From this, we have

$$V_0 \equiv |\underline{g}|^{-2} - cA_2 > 0, \tag{12}$$

$$\underline{g}_2 = V_0^{-1}(cB_2\underline{m}_2 + v),$$

and consequently

$$V\underline{m}_2 + V_0^{-1}cA_1 v = (z\underline{m})_2, \tag{13}$$

where

$$V \equiv c^2A_1B_2V_0^{-1} + cB_1. \tag{14}$$

Multiplying by $\underline{m}$ and dividing by z on both sides of the first equation of (10) and comparing the imaginary parts we obtain

$$\underline{m}_2|z|^2 = v(1-c) + cB_0v + cA_1(z\underline{g})_2 + cB_1(z\underline{m})_2. \tag{15}$$

Dividing z on both sides of the second equation of (10), we have

$$0 = (z\underline{g})_2|\underline{g}|^{-2} - cB_1v - cA_2(z\underline{g})_2 - cB_2(z\underline{m})_2. \tag{16}$$

From (16), we have

$$(z\underline{g})_2 = V_0^{-1}\left(cB_2(z\underline{m})_2 + cB_1v\right)$$

Then, from (15), we have

$$\underline{m}_2|z|^2 = V(z\underline{m})_2 + (1-c)v + cB_0v + c^2A_1B_1vV_0^{-1}.$$

Substituting this into (13) above, we obtain

$$\underline{m}_2|z|^2 = V^2\underline{m}_2 + (VcA_1 + c^2A_1B_1)vV_0^{-1} + (1-c)v + cB_0v, \tag{17}$$

We see, then, for $c \leq 1$

$$V = c^2A_1B_2V_0^{-1} + cB_1 < |z| \tag{18}$$

In order to obtain the properties of the LSD more conveniently, we give a preliminary result. It is stated under conditions sufficient for this paper.

Theorem 1. *We assume $c \leq 1$. Under the assumptions on H and the additional condition that $\lambda_j \leq K$, $j = -1, 1$ for some constant K, where $\lambda_j = \left(\int t^j\mathrm{d}H(s,t)\right)^{1/2}$, we have*

(a) *For* $0 < m \le |z| \le M < \infty \quad z \in \mathbb{C}^+$, *there exist* $0 < \delta < \Delta < \infty$ *such that*

$$|\underline{m}(z)| \le \Delta \quad \text{and} \quad \delta \le |\underline{g}(z)| \le \Delta.$$

For the existence of the lower bound of $\underline{m}(z)$, *we need one more condition that* $\int (s/t)\mathrm{d}H(s,t) \le \frac{1}{c}$.

(b) *All the four quantities* A_j, B_j, $j = 1, 2$ *are bounded.*

(c) *If* $\lim_{z_n \in \mathbb{C}^+} z_n \to z$, $\underline{m}(z_n) \to \tilde{m}(z)$ *and* $\underline{g}(z_n) \to \tilde{g}(z)$, *then* $(z, \tilde{m}(z), \tilde{g}(z))$ *satisfies equations* (10).

We call any triple $(z, \tilde{m}(z), \tilde{g}(z))$ *formed in this way an "extended solution".*

(d) *For each* $z \neq 0$, *the extended solutions are unique.*

Remark 1. The condition $\int s/t\mathrm{d}H(s,t) \le \frac{1}{c}$ is one related to the ratio of signal to noise (RSN). We have difficulty proving the lower bound without this condition. This condition means that the average RSN of each sample vector is not more than 1. That is

$$1 \ge \frac{1}{N}\sum_{k=1}^{N} T_n^{-\frac{1}{2}}(N^{-\frac{1}{2}}r_k)(N^{-\frac{1}{2}}r_k)^* T_n^{-\frac{1}{2}} = \frac{n}{N}\frac{1}{n}\mathrm{tr}(\frac{1}{N}R_nR_n^*)T_n^{-1} \to c\int s/t\mathrm{d}H(s,t).$$

Thus, this condition seems reasonable. However, we hope the lower bound for $m(z)$ can be proved without this additional condition.

Proof. If (a) is not true for the upper bound, then there exists $\underline{m}(z_n) \to \infty$ or $\underline{g}(z_n) \to \infty$. Noting that, from (18) $cB_1 \le |z| \le M$, from the Cauchy-Schwarz inequality we find

$$\begin{aligned}\int \frac{\mathrm{d}H(s,t)}{|1 + s\underline{g}(z_n) + t\underline{m}(z_n)|} &\le \lambda_{-1}\sqrt{B_1} \le M_1,\\ \int \frac{t\mathrm{d}H(s,t)}{|1 + s\underline{g}(z_n) + t\underline{m}(z_n)|} &\le \lambda_1\sqrt{B_1} \le M_1.\end{aligned} \tag{19}$$

for some positive M_1. The first estimate above, together with the assumption $\underline{m}(z_n) \to \infty$ implies the right hand side of the first equation of (10) tends to zero while the left hand side does not. The second estimate above, together with the assumption $\underline{g}(z_n) \to \infty$ implies the right hand side of the second equation of (10) tends to zero while the left hand side does not. These contradictions show that both $\underline{s}(z)$ and $\underline{g}(z)$ are bounded from above. Now, suppose there is a sequence $\{z_n\}$ such that $\underline{g}(z_n) \to 0$. By the second equation of (10), we have

$$z_n\underline{g}(z_n) = -1 + c\int \frac{t\underline{g}(z_n)}{1 + s\underline{g}(z_n) + t\underline{m}(z_n)}\mathrm{d}H(s,t).$$

The second estimate in (19) shows that the second term on the right side of the above and the left hand side tend to zero, a contradiction.

Finally, we show that $\underline{m}(z)$ is bounded from zero. Now, suppose there is a bounded sequence $\{z_n\}$ such that $\underline{m}(z_n) \to 0$. By the first equation of (10), we have

$$z_n\underline{m}(z_n) = -1 + c - c\int \frac{1}{1 + s\underline{g}(z_n) + t\underline{m}(z_n)}\mathrm{d}H(s,t). \tag{20}$$

Since it has been proven that $\{z_n, \underline{g}(z_n)\}$ is bounded, there is a subsequence n' such that $z_{n'} \to z_0$ and $\underline{g}(z_{n'}) \to \underline{g}_0$. The first estimation in (19) together with Fatou's Lemma shows that

$$\int \frac{1}{|1 + s\underline{g}_0|}\mathrm{d}H(s,t) \le M_1.$$

Consequently, $1/(1+s\underline{g}_0)$ is integrable with respect to H, hence,

$$
\begin{aligned}
&\left|\int \frac{1}{1+s\underline{g}(z_{n'})+t\underline{m}(z_{n'})}\mathrm{d}H(s,t) - \int \frac{1}{1+s\underline{g}_0}\mathrm{d}H(s,t)\right| \\
=&\left|\int \frac{s(\underline{g}_0-\underline{g}(z_{n'}))-t\underline{m}(z_{n'})}{(1+s\underline{g}(z_{n'})+t\underline{m}(z_{n'}))(1+s\underline{g}_0)}\mathrm{d}H(s,t)\right| \\
\le&|\underline{g}_0-\underline{g}(z_{n'})|\sqrt{A_1(n')A_1(0)} + |\underline{m}(z_{n'})||z_{n'}|\sqrt{B_1(n')B_1(0)},
\end{aligned}
$$

where $A_1(n')$ ($B_1(n')$) and $A_1(0)$ ($B_1(0)$) are the A (B) functions defined by extended solutions $(z_{n'}, g(z_{n'}), m(z_{n'}))$ and $(z_0, g_0, 0)$, respectively. In part (b), it will be proven that both the A and B functions are bounded, hence, the right hand side of the inequality above tends to zero. Therefore, from (20), we obtain

$$
0 = -1 + c - c\int \frac{1}{1+s\underline{g}_0}\mathrm{d}H(s,t) = -1 + c\int \frac{s\underline{g}_0}{1+s\underline{g}_0}\mathrm{d}H(s,t). \tag{21}
$$

If the imaginary of $\underline{g}_0$ is positive, the equality above could not be true (including the case $H(\{s=0\})=1$). By Cauchy-Schwarz, we have

$$
\begin{aligned}
&\int \frac{|s\underline{g}_0|}{|1+u\underline{g}_0|}\mathrm{d}H(s,t) \le \left(\int (s/t)\mathrm{d}H(s,t)\int \frac{st|\underline{g}_0|^2}{|1+s\underline{g}_0|^2}\mathrm{d}H(s,t)\right)^{\frac{1}{2}} \\
=&\sqrt{\int (s/t)\mathrm{d}H(s,t)|\underline{g}_0^2|A_2} < \frac{1}{c}.
\end{aligned}
$$

Here the last inequality follows from the additional condition and (12). We reach a contradiction to (21) and the proof of (a) is complete.

Proof of (b). From (18) we have seen that $B_1 \le |z|/c \le M/c$ by assumption. Also, we have from (12) that $A_2 \le c^{-1}|\underline{g}|^{-2} \le M_2$ for some positive M_2, since $|\underline{g}| \ge \delta > 0$ proven in (a). Next, by (18) we have

$$
A_1B_2 \le |z||\underline{g}|^{-2}/c^2 < M_3 \tag{22}
$$

for some M_3. To show A_1 is bounded from above, we claim that there is a positive t_0 such that

$$
\int \frac{I_{\{t>t_0\}}s\mathrm{d}H(s,t)}{|1+s\underline{g}+t\underline{m}|^2} \ge 0.5A_1.
$$

If not, then for $t_0 = 0$,

$$
\int \frac{I_{\{t>0\}}s\mathrm{d}H(s,t)}{|1+u\underline{g}+t\underline{m}|^2} \le 0.5A_1
$$

which implies either $A_1 = 0$ or $H(\{t=0\}) > 0$ which violate the assumption that $\lambda_{-1} \le K$. Consequently, we have

$$
\begin{aligned}
A_1 &\le 2\int \frac{I_{\{t>t_0\}}s\mathrm{d}H(s,t)}{|1+s\underline{g}+t\underline{m}|^2} \\
&\le \frac{2}{t_0}\int \frac{I_{\{t>t_0\}}st\mathrm{d}H(s,t)}{|1+s\underline{g}+t\underline{m}|^2} \le \frac{2}{t_0}A_2 \le M.
\end{aligned}
$$

Then, we conclude that A_1 is bounded from above.

If B_2 is not bounded for bounded z, then there exists a $z_0 \in \mathbb{C}^+$ such that $B_2(z_0) = \infty$. By (22), we have $A_1(z_0) = 0$.

This implies that the marginal for s of H is concentrated at zero. Therefore,

$$B_2 = \int \frac{t^2 \mathrm{d}H(s,t)}{|1+t\underline{m}(z_0)|^2} = \infty.$$

This equation implies that $\underline{m}(z_0)$ is real and not positive. By part (a), $\underline{m}(z_0) \le -\delta$ for some $\delta > 0$. If $t > 2/\delta$, then $|1+t\underline{m}(z_0)| > t\delta - 1 > 0.5t\delta$, therefore.

$$\int \frac{I_{\{t \le 2/\delta\}} t^2 \mathrm{d}H(s,t)}{|1+t\underline{m}(z_0)|^2} = \infty.$$

This is impossible because

$$\int \frac{I_{\{t \le 2/\delta\}} t^2 \mathrm{d}H(s,t)}{|1+t\underline{m}(z_0)|^2} \le \frac{2}{\delta} \int \frac{I_{\{t \le 2/\delta\}} t \mathrm{d}H(s,t)}{|1+t\underline{m}(z_0)|^2} \le \frac{2}{\delta} B_1 \le M.$$

Thus the contradiction proves that B_2 is bounded from above.

For (c), by applying Fatou's Lemma when the limiting z is real, the same bounds on A_j, B_j hold true for extended solutions to equation (10). Also, from (19), using Fatou's Lemma, the integrals on the right side of (10) also exist and are bounded.

We have

$$\begin{aligned}
&\int \frac{\mathrm{d}H(s,t)}{1+s\underline{g}(z_n)+t\underline{m}(z_n)} - \int \frac{\mathrm{d}H(s,t)}{1+s\tilde{g}+t\tilde{m}} \\
=&(\tilde{g}-\underline{g}(z_n)) \int \frac{s\mathrm{d}H(s,t)}{(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})} \\
&+ (\tilde{m}-\underline{m}(z_n)) \int \frac{t\mathrm{d}H(s,t)}{(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})} \to 0, \\
&\int \frac{t\mathrm{d}H(s,t)}{1+s\underline{g}(z_n)+t\underline{s}(z_n)} - \int \frac{tdH(s,t)}{1+s\tilde{g}+t\tilde{m}} \\
=&(\tilde{g}-\underline{g}(z_n)) \int \frac{st\mathrm{d}H(s,t)}{(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+u\tilde{g}+t\tilde{m})} \\
&+ (\tilde{m}-\underline{m}(z_n)) \int \frac{t^2\mathrm{d}H(s,t)}{(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+u\tilde{g}+t\tilde{m})} \to 0,
\end{aligned}$$

since, with $A_j(n), B_j(n), A_j, B_j$ having their obvious meanings, for some positive M.

$$\begin{aligned}
\int \frac{s\mathrm{d}H(s,t)}{|(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})|} &\le \sqrt{A_1(n)A_1} \le M, \\
\int \frac{t\mathrm{d}H(s,t)}{|(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})|} &\le \sqrt{B_1(n)B_1} \le M, \\
\int \frac{st\mathrm{d}H(s,t)}{|(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})|} &\le \sqrt{A_2(n)A_2} \le M, \\
\int \frac{t^2\mathrm{d}H(u,t)}{|(1+s\underline{g}(z_n)+t\underline{m}(z_n))(1+s\tilde{g}+t\tilde{m})|} &\le \sqrt{B_2(n)B_2} \le M.
\end{aligned}$$

Proof of (d). Notice in (b) we have both A_1 and B_2 both bounded from below. Therefore we have from (18) V_0^{-1} is bounded. Hence, from Fatou's Lemma and the fact that $\underline{g}$ is bounded from below we have from (18)

$$V = c^2 A_1 B_2 V_0^{-1} + cB_1 \le |z| \tag{23}$$

holding for all z with $\Im z \geq 0$.

Suppose for $z \neq 0$ with $\Im z \geq 0$ there are two sets of extended solutions $(\underline{m}_{(j)}, \underline{g}_{(j)})$, $j = 1, 2$. Multiplying $\underline{m}_{(j)}$ on both sides of the first equation of (10) and taking the difference on both sides, we obtain

$$\begin{aligned} z(\underline{m}_{(1)} - \underline{m}_{(2)}) &= c\tilde{A}_1(\underline{g}_{(1)} - \underline{g}_{(2)}) + c\tilde{B}_1(\underline{m}_{(1)} - \underline{m}_{(2)}), \\ 0 &= \frac{\underline{g}_{(1)} - \underline{g}_{(2)}}{\underline{g}_{(1)}\underline{g}_{(2)}} - c\tilde{A}_2(\underline{g}_{(1)} - \underline{g}_{(2)}) - c\tilde{B}_2(\underline{m}_{(1)} - \underline{m}_{(2)}), \end{aligned} \tag{24}$$

where

$$\tilde{A}_j = \int \frac{st^{j-1}\mathrm{d}H(s,t)}{(1 + s\underline{g}_{(1)} + t\underline{m}_{(1)})(1 + s\underline{g}_{(2)} + t\underline{m}_{(2)})}, \quad j = 1, 2$$

$$\tilde{B}_j = \int \frac{t^j\mathrm{d}H(s,t)}{(1 + s\underline{g}_{(1)} + t\underline{m}_{(1)})(1 + s\underline{g}_{(2)} + t\underline{m}_{(2)})}, \quad j = 0, 1, 2.$$

Let $A_j^{(i)}$, $B_j^{(i)}$, $i, j = 1, 2$ be the A_j, B_j functions defined in (11) by the i-th extended solutions. Then by Cauchy-Schwarz we have

$$|c\tilde{A}_2| \leq c(A_2^{(1)})^{1/2}(A_2^{(2)})^{1/2} < \frac{1}{|\underline{g}_{(2)}||\underline{g}_{(1)}|},$$

so we have

$$\left|\frac{1}{\underline{g}_{(2)}\underline{g}_{(1)}} - c\tilde{A}_2\right| \geq \left|\frac{1}{|\underline{g}_{(2)}||\underline{g}_{(1)}|} - |c\tilde{A}_2|\right| \geq \frac{1}{|\underline{g}_{(2)}||\underline{g}_{(1)}|} - c\sqrt{A_2^{(1)}A_2^{(2)}}^{1/2} > 0.$$

We see then, from the second equation in (24), that $\underline{m}_{(1)} = \underline{m}_{(2)}$ if and only if $\underline{g}_{(1)} = \underline{g}_{(2)}$

Suppose $\underline{m}_{(1)} \neq \underline{m}_{(2)}$. We can then write

$$\frac{\underline{g}_{(1)} - \underline{g}_{(2)}}{\underline{m}_{(1)} - \underline{m}_{(2)}} = \frac{c\tilde{B}_2}{\frac{1}{\underline{g}_{(2)}\underline{g}_{(1)}} - c\tilde{A}_2}.$$

From (24) and (23), applying Cauchy-Schwarz, and using the inequality

$$(ab - cd)^2 \geq (a^2 - c^2)(b^2 - d^2), \text{ for real } a, b, c, d$$

we have

$$\begin{aligned} |z| &= \left|\frac{c\tilde{B}_2 c\tilde{A}_1}{\left(\frac{1}{\underline{g}_{(2)}\underline{g}_{(1)}}\right) - c\tilde{A}_2} + c\tilde{B}_1\right| \leq \frac{c^2\sqrt{B_2^{(1)}B_2^{(2)}A_1^{(1)}A_1^{(2)}}}{\left(\frac{1}{|\underline{g}_{(2)}||\underline{g}_{(1)}|}\right) - c\sqrt{A_2^{(1)}A_2^{(2)}}} + c\sqrt{B_1^{(1)}B_1^{(2)}} \\ &\leq \left(\left(\frac{c^2B_2^{(1)}A_1^{(1)}}{\left(\frac{1}{|\underline{g}_{(1)}|^2} - cA_2^{(1)}\right)} + cB_1^{(1)}\right)\left(\frac{c^2B_2^{(2)}A_1^{(2)}}{\left(\frac{1}{|\underline{g}_{(2)}|^2} - cA_2^{(2)}\right)} + cB_1^{(2)}\right)\right)^{\frac{1}{2}} \leq |z|. \end{aligned} \tag{25}$$

We see that when $v > 0$, from (18) we get an immediate contradiction. Therefore for $z \in \mathbb{C}^+$ there is only one $\underline{m}$ which satisfy (10), namely $m_{\underline{F}}(z)$

It amounts to show uniqueness when $v = 0$, that is $z = x \in \mathbb{R}$, $x \neq 0$. Then we must have (23) with an equal sign. When at least one $\Im m_{(j)} > 0$ we arrive at a contradiction from the fact that $\widetilde{B}_1 < \sqrt{B_1^{(1)}B_1^{(2)}}$, yielding the first inequality in (25) a strict one. Indeed, in deriving the Cauchy-Schwarz inequality $|\mathbf{E}(X\hat{Y})| \leq \sqrt{\mathbf{E}(|X|^2)\mathbf{E}(|Y|^2)}$, one finds the $a = re^{i\theta}$ which minimizes $\mathbf{E}|X - aY|^2$. Equality only occurs when $X = aY$ with probability one. Here we can take

$$X = \frac{t^{1/2}}{1 + s\underline{g}_{(1)} + t\underline{m}_{(1)}} \quad \text{and} \quad Y = \frac{t^{1/2}}{1 + s\overline{\underline{g}}_{(2)} + t\overline{\underline{m}}_{(2)}},$$

so clearly one cannot be a multiple of the other.

Which leaves the case $\Im m_{(1)} = \Im m_{(2)} = 0$. In order to have the second inequality in (25) to be an equality we need to have the vectors

$$\left(\sqrt{\frac{c^2 A_1^{(1)} B_2^{(1)}}{\frac{1}{|\underline{g}_{(1)}|^2} - cA_2^{(1)}}}, \sqrt{cB_1^{(1)}}\right), \left(\sqrt{\frac{c^2 A_1^{(2)} B_2^{(2)}}{\frac{1}{|\underline{g}_{(2)}|^2} - cA_2^{(2)}}}, \sqrt{cB_1^{(2)}}\right)$$

proportional to each other, say the first vector is k times the second vector, necessarily k is positive, and if the last inequality is to be an equality, from (23), it must be 1. Therefore we have $B_1^{(1)} = B_2^{(2)}$.

If the first inequality in (25) is to be an equality, then it must be that $\tilde{B}_1 > 0$, $\sqrt{\tilde{B}_1} = B_1^{(1)} = B_1^{(2)}$, and there is a $k > 0$ such that, with probability one with respect to $H(s,t)$

$$\frac{1}{1 + s\underline{g}_{(1)} + t\underline{m}_{(1)}} = k \frac{1}{1 + s\underline{g}_{(2)} + t\underline{m}_{(2)}}.$$

Therefore, we must have $k = 1$, and we see from the first equation in (10) that the integrals are identical for $j = 1, 2$, from which we conclude that $\underline{m}_{(1)} = \underline{m}_{(2)}$, which implies $\underline{g}_{(1)} = \underline{g}_{(2)}$, This completes the proof of the theorem. $\square$

Dozier and Silverstein [4] extended the analysis for LSD of generalized M-P law to the LSD of the non-central sample covariance matrix. Theorem 2.1 in [4] can be extended to the solution for model (7). We shall prove the following theorem under the assumptions $c \leq 1$ and the conditions imposed on the limiting H in Theorem 1.

Theorem 2. *Assume $c \leq 1$ and the conditions imposed in the limiting H in Theorem 1. Suppose $\underline{m}(x)$ is the solution to* (10) *for $x \neq 0$. Then $x \in \mathbb{R}\backslash\{0\}$, $\lim_{z \in \mathbb{C}^+ \to x} m_{\underline{F}}(z) \equiv \underline{m}(x)$ exists. The function $\underline{m}$ is continuous on $\mathbb{R}\backslash\{0\}$, and F has a continuous derivative f on $\mathbb{R}\backslash\{0\}$ given by $f(x) = \frac{1}{\pi}\Im\underline{m}(x)$. Furthermore, if $\Im\underline{m}(x) > 0$ ($f(x) > 0$) for $x \in \mathbb{R}^+$, then the density f is analytic about x.*

Proof of Theorem 2. The first conclusion of Theorem 2 is a special case of Theorem 1. Because of (a), (c), and (d), as $z \in \mathbb{C}^+ \to x \neq 0$ $\underline{m}(z) \to \underline{m}(x)$, solution to (10). Thus, the Stieltjes transform extends uniquely on $\mathbb{R}/\{0\}$.

Then, due to Theorems 1.1 and 1.2 of [4] we have that for all $x \neq 0$ $\underline{F}$ is continuously differentiable and the density is given by $f(x) = \frac{1}{\pi}\Im(\underline{m}(x)) = \frac{1}{\pi}\lim_{v \downarrow 0} \Im(\underline{m}(x + iv))$.

Next we show that the density $f(x)$ is analytic when $\Im(\underline{m}(x)) > 0$. Therefore, the denominator of the fractions in (10) cannot be zero at the point $(m(x), g(x), x)$. Hence, the two integrals in (10) are analytic in some neighborhood D_x of $(\underline{m}(x), \underline{g}(x), x)$. We shall employ the implicit function theorem to show that $\underline{m}(x)$ is analytic in some neighborhood D_x.

Rewrite the the two equations in (10) as $G_{\underline{m}}(\underline{m}, \underline{g}, z) = 0$ and $G_{\underline{g}}(\underline{m}, \underline{g}, z) = 0$, where

$$\begin{aligned} G_{\underline{m}} &= z\underline{m} + (1 - c) + c \int \frac{\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}}, \\ G_{\underline{g}} &= z + \frac{1}{\underline{g}} - c \int \frac{t\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}}. \end{aligned} \tag{26}$$

Then

$$\frac{\partial}{\partial \underline{m}} G_{\underline{m}}(\underline{m}, \underline{g}, x) = x - \hat{B}_1 c,$$
$$\frac{\partial}{\partial \underline{g}} G_{\underline{m}}(\underline{m}, \underline{g}, x) = -\hat{A}_1 c,$$
$$\frac{\partial}{\partial \underline{g}} G_{\underline{g}}(\underline{m}, \underline{g}, x) = -\frac{1}{\underline{g}^2} + \hat{A}_2 c,$$
$$\frac{\partial}{\partial \underline{m}} G_{\underline{g}}(\underline{m}, \underline{g}, x) = \hat{B}_2 c.$$

where

$$\hat{A}_1 = \int \frac{s\mathrm{d}dH(s,t)}{(1 + s\underline{g} + t\underline{m})^2},$$
$$\hat{B}_1 = \int \frac{t\mathrm{d}H(s,t)}{(1 + s\underline{g} + t\underline{m})^2},$$
$$\hat{A}_2 = \int \frac{st\mathrm{d}H(s,t)}{(1 + s\underline{g} + t\underline{m})^2},$$
$$\hat{B}_2 = \int \frac{t^2\mathrm{d}H(s,t)}{(1 + s\underline{g} + t\underline{m})^2}.$$

By the implicit function theorem, $\underline{m}$ and $\underline{g}$ are analytic in some neighborhood D_x of x if the absolute value of the below determinant is greater than zero. We have for $x > 0$

$$\begin{aligned}
&\left\| \begin{matrix} \frac{\partial}{\partial \underline{m}} G_{\underline{m}}(\underline{m}, \underline{g}, x) & \frac{\partial}{\partial \underline{g}} G_{\underline{m}}(\underline{m}, \underline{g}, x) \\ \frac{\partial}{\partial \underline{m}} G_{\underline{g}}(\underline{m}, \underline{g}, x) & \frac{\partial}{\partial \underline{g}} G_{\underline{g}}(\underline{m}, \underline{g}, x) \end{matrix} \right\| > |(x - \hat{B}_1 c)(\frac{1}{\underline{g}^2} - \hat{A}_2 c) - c^2 \hat{A}_1 \hat{B}_2| \\
\geq &(x - B_1 c)(\frac{1}{|\underline{g}|^2} - A_2 c) - c^2 A_1 B_2 \\
= &(\frac{1}{|\underline{g}|^2} - A_2 c)(|x| - cB_1 - c^2 A_1 B_2 (\frac{1}{|\underline{g}|^2} - A_2 c)^{-1}) = 0
\end{aligned}$$

since, from the same argument given above, the ">" comes from the fact that the absolute values of the $\hat{A}_i$, $\hat{B}_i$ are strictly less than their real counterparts, and "=" is due to (25). The proof is complete. □

3. The Support of F

In this section we present results on the support of the limiting distribution F. The support of a distribution function F is the set of all points x satisfying $F(x + \varepsilon) - F(x - \varepsilon) > 0$ for all $\varepsilon > 0$. Let S_F and S_H denote the support of F and H, respectively. Clearly, by definition of F and H, we have $S_F \subset [0, \infty)$ and $S_H \subset [0, \infty)$.

We will concentrate our investigation of the support, S_F of F by the support, $S_{\underline{F}}$ of $\underline{F}$. The latter can be found by determining those values of the real line which are in its complement, $S^c_{\underline{F}}$.

Corresponding to [4], we begin our analysis of the support of the LSD F with the following result.

Theorem 3. *When $c \leq 1$, and assuming the conditions imposed on H in Theorem 1, the LSD F determined by* (6) *has no mass at zero.*

Proof. We use the formula $G(\{0\}) = \lim_{v\downarrow 0}(-ivS_G(iv))$ for any probability measure G. We have by the relation between $m(z)$ and $\underline{m}(z)$, and (10)

$$F(\{0\}) = \lim_{v\downarrow 0}(-ivs(iv)) = \lim_{v\downarrow 0} \int \frac{\mathrm{d}H(u,t)}{1 + u\underline{g}(iv) + t\underline{m}(iv)}$$
$$= \lim_{v\downarrow 0} \int \frac{-iv\mathrm{d}H(u,t)}{-iv - uiv\underline{g}(iv) - tiv\underline{m}(iv)} = \lim_{v\downarrow 0} \int \frac{-iv\mathrm{d}H(u,t)}{-uiv\underline{g}(iv) + t\underline{F}(\{0\})}.$$

where $\underline{F}(\{0\}) = F(\{0\}) \geq 0$ if $c = 1$ and $\underline{F}(\{0\}) \geq (1-c)$ if $c < 1$. In the Appendix it is argued that when H has bounded support, $\underline{g}(z)$ is the Stieltjes transform of $P_{\underline{g}}$, a probability measure, and therefore $\lim(-iv\underline{g}(iv)) \geq 0$. Thus $F(\{0\}) > 0$ will lead to a contradiction because the left hand side would be positive while the right hand would necessarily be zero.

$\square$

To find points, x, in $S^c_{\underline{F}}$, It is sufficient to consider only $x \geq 0$, and when $c \leq 1$, due to the above theorem, $x > 0$, since if $(0,\epsilon) \subset S^c_{\underline{F}}$, then necessarily $\underline{F}$ has mass $1 - c$ at 0.

We next claim that the supports of $\underline{F}$ and $P_{\underline{g}}$ are identical. Notice that $x \in \mathbb{R}$ outside the support of a measure implies the Stieltjes transform of the measure is real at x. If $(x, \underline{m}, \underline{g})$, with x real, is an extended solution, and if $\underline{m}$ is real, then from the first equation in (10) we must have $\underline{g}$ real. If $\underline{g}$ is real, then the second equation in (10) requires $\underline{m}$ to be real. This proves the claim. It will be useful to consider $S^c_{P_{\underline{g}}}$ in the arguments below.

The following presents a scheme for compute solutions $z, \underline{m}, \underline{g}$ of (10). Equating the two equations in (10) we get

$$-\frac{1-c}{\underline{m}} - \frac{c}{\underline{m}} \int \frac{\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}} = -\frac{1}{\underline{g}} + c \int \frac{t\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}}.$$

This is equivalent to

$$c\underline{g}^2 \int \frac{s\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}} + \underline{m} - \underline{g} = 0. \tag{27}$$

When $(z, \underline{m}, \underline{g})$ is an extended solution to (10), we claim there is only one $\underline{m}$ for every $u\underline{g}$ solving this equation. If not, suppose there are two different $\underline{m}_1$ and $\underline{m}_2$ satisfying the equation. Taking the difference between the two equations we obtain,

$$c(\underline{m}_2 - \underline{m}_1)\underline{g}^2 \int \frac{st\mathrm{d}H(s,t)}{(1 + s\underline{g} + t\underline{m}_1)(1 + u\underline{g} + t\underline{m}_2)} + \underline{m}_1 - \underline{m}_2 = 0.$$

Consequently, we have

$$c \int \frac{st\mathrm{d}H(u,t)}{(1 + s\underline{g} + t\underline{m}_1)(1 + s\underline{g} + t\underline{m}_2)} = \frac{1}{\underline{g}^2}.$$

Therefore, by Cauchy-Schwarz

$$\frac{1}{|\underline{g}|^2} \leq \left(c \int \frac{st\mathrm{d}H(s,t)}{|1 + s\underline{g} + t\underline{s}_1|^2} \cdot c \int \frac{st\mathrm{d}H(s,t)}{|1 + s\underline{g} + t\underline{m}_2|^2} \right)^{1/2} < \frac{1}{|\underline{g}|^2}.$$

Here, the last inequality follows from (12) (as mentioned above, true for real x). The contradiction proves our assertion.

Theorem 4. *Assume $c \leq 1$ and the conditions imposed on H in Theorem 1 hold. Let $x_0 \in S^c_{\underline{F}} \cap \mathbb{R}^+$*

(a) *Then $\underline{m}(z) = \int (t-z)^{-1}\mathrm{d}\underline{F}(t)$ is analytic in a neighborhood D_{x_0} of x_0 and there exists a co-solution $\underline{g}(z)$ which is also analytic in D_{x_0}. The triple $(x, \underline{m}(x), \underline{g}(x))$, $x \in D_{x_0} \cap \mathbb{R}^+$ is an extended solution to* (10) *with $V < x$.*

(b) *For any support point (s,t) of H, $s\underline{g}(x_0)+t\underline{m}(x_0)\neq -1$.*

On the other hand, if $x_0, \underline{m}_0, \underline{g}_0$, with $x_0>0$, form a real extended solution to (10) *satisfying (b), then from* (27)*, there exists a real analytic function $x=x(\underline{g})$, defined in an interval containing $\underline{g}_0$ which satisfy* (10)*, and if $x'(\underline{g}_0)\neq 0$, then $x_0\in S_{\underline{F}}^c$*

Proof. . When x_0 is outside the support of a probability distribution, its Stieltjes transform exists at x_0, is analytic in a real neighborhood of x_0, and has a positive derivative at each point in this neigborhood. So for $x_0\in S_{\underline{F}}^c\cap\mathbb{R}^+$, there is a constant $\varepsilon\in(0,x_0)$ such that $U_{x_0}=(x_0-\varepsilon, x_0+\varepsilon)\subset S_{\underline{F}}^c$. Let D_{x_0} denote the ball in the complex plane centered at x_0 with radius ε. Then, for any $z\in D_{x_0}$, $\underline{\tilde{m}}(z)=\int(t-z)^{-1}\mathrm{d}\underline{F}(t)$ is well defined and analytic.

On the other hand, for each $z=x+iv\in D_{x_0}\cap\mathbb{C}^+$, there are $\underline{m}(z)$ and $\underline{g}(z)$ well defined solutions of (10) and analytic. Then $\underline{m}(x+iv)=\int(t-(x+iv))^{-1}\mathrm{d}\underline{F}(t)$ and $\underline{\tilde{m}}(z)$ are identical. Let $v\to 0$, i.e., $z=x+iv\xrightarrow{\mathbb{C}^+}x$. By Theorem 2, $\underline{m}(x)$, $\underline{g}(x)$ exist and satisfy (10). By the unique extension theorem of analytic functions $\underline{m}(x)=\underline{\tilde{m}}(x)$ for all $x\in U_{x_0}$. In the following, we will identify $\underline{\tilde{m}}$ with $\underline{m}$.

The same argument applies for $P_{\underline{g}}$, so $\underline{g}(z)$ is also analytic in a neighborhood of x_0. Let D_0 and U_0 be the smaller of the two respective sets associated with $\underline{m}$ and $\underline{g}$.

Suppose there is a support point (s,t) of $H(s,t)$ such that $s\underline{g}(x_0)+t\underline{m}(x_0)+1=0$. By Lemma 1 which will be given later, there is a point x^* arbitrarily close to x_0 such that $\int(1+s\underline{g}(x^*)+t\underline{m}(x^*))^{-2}\mathrm{d}H(s,t)=\infty$, which contradicts $\underline{m}(x)$ being analytic in a neighborhood of x_0. Thus, (b) holds.

Since $\underline{m}(x)$ for $x\in U_{x_0}$ is real with $\underline{m}'(x)>0$, by the inverse function theorem, there is a unique analytic function $z(\underline{m})$ defined on some neighborhood $D_{\underline{m}_0}$ in the $\underline{m}$ plane and such that $x=z(\underline{m}(x))$, $x\in U_{x_0}$. Since $z'(\underline{m}(x))=1/\underline{m}'(x)$, we conclude that $z'(\underline{m}(x))>0$.

We can compute the derivative of $\underline{m}(x)$ as follows. Recalling the definitions of $G_{\underline{m}}$ and $G_{\underline{g}}$, we have

$$\frac{\partial}{\partial z}G_{\underline{m}}+\underline{m}'\frac{\partial}{\partial \underline{m}}G_{\underline{m}}+\underline{g}'\frac{\partial}{\partial \underline{g}}G_{\underline{m}}=\underline{m}+\underline{m}'x-\underline{m}'cB_1-c\underline{g}'A_1=0,$$

$$\frac{\partial}{\partial z}G_{\underline{g}}+\underline{m}'\frac{\partial}{\partial \underline{m}}G_{\underline{g}}+\underline{g}'\frac{\partial}{\partial \underline{g}}G_{\underline{g}}=1+\underline{m}'cB_2-\frac{\underline{g}'}{\underline{g}^2}+c\underline{g}'A_2=0.$$

From the second equation, we get

$$\underline{g}'=(1+\underline{m}'cB_2)/(\underline{g}^{-2}-cA_2)$$

and then substituting it to the first equation

$$\underline{m}'=[(-\underline{m}+cA_1)/(\underline{g}^{-2}-cA_2)](x-V)^{-1}, \tag{28}$$

Therefore

$$z'=z'(\underline{m})=(x-V)[(-\underline{m}+cA_1)/(\underline{g}^{-2}-cA_2)]^{-1}. \tag{29}$$

Conversely, suppose $x_0, \underline{m}_0, \underline{g}_0$, with $x_0\neq 0$, form a real extended solution to (10) satisfying (b). Since $\inf\{|s\underline{g}_0+t\underline{m}_0+1| : (s,t)\in S_H\}>0$ (since S_H is a closed set), the integrals in (10) as well as the integral in (27) are analytic functions of $\underline{m}$ and $\underline{g}$ in a neighborhood of $(\underline{m}_0,\underline{g}_0)$. Notice that the partial of (27) with respect to $\underline{m}$ is

$$-c\underline{g}^2A_2+1=\underline{g}^2V_0>0.$$

Therefore, by the implicit function theorem, $\underline{m}=\underline{m}(\underline{g})$ is uniquely defined and is analytic in a neighborhood of $\underline{g}_0$ and with either of the two equations in (10) we determine z. So we see that in a neighborhood of $\underline{g}_0$, $z=z(\underline{g})$ is analytic in a neighborhood of $\underline{g}_0$. which are solutions to (10). Because of uniqueness we have, when $z\in\mathbb{C}^+$, $m_{\underline{F}}(z)=\underline{m}(\underline{g})\in\mathbb{C}^+$ and $m_{\underline{g}}(z)=\underline{g}\in\mathbb{C}^+$. We also see that (27) yields real $\underline{m}$ for each real $\underline{g}$, and so $x=x(\underline{g})$ is a real valued function for $\underline{g}$ lying in an interval containing $\underline{g}_0$.

If $x'(\underline{g}_0) \neq 0$, then by the inverse function theorem there exists an analytic function $\underline{g} = \underline{g}(z)$ for $z \in \mathbb{C}$ lying in a neighborhood of x_0 which is the inverse of $z(\underline{g})$, so that $z, \underline{m}(\underline{g}), \underline{g}$ solve (10). Since $x_0, \underline{m}_0, \underline{g}_0$ is an extended solution, there exists a sequence $z_n \in \mathbb{C}^+ \to x_0$ and $m_{\underline{F}}(z_n), m_{\underline{g}}(z_n) \to \underline{m}_0, \underline{g}_0$. Therefore $\underline{g}(z)$ is the analytic extension of $m_{\underline{g}}$ onto an interval of the real line containing x_0, real valued when z is real, which implies that the density of $P_{\underline{g}}$ exists and is zero in an interval containing x_0, so that necessarily x_0 is outside the support of $P_{\underline{g}}$ and $\underline{F}$. Since x_0 is outside the support of $P_{\underline{g}}$ $m_{\underline{g}}(z_0)$ and $x'(\underline{g}_0)$ are necessarily positive. □

Lemma 1. *Suppose $H(s,t)$ is a measure supported by a closed subset of the first quadrant excluding the origin. If there is a support point (s_0, t_0) such that $g_0 s_0 + m_0 t_0 + c = 0$, where $c > 0$ is a constant, $g_0 = g(w_0)$ and $m_0 = m(w_0)$ are functions defined on the interval $U = (w_0 - \eta, w_0 + \eta)$ whose derivative are not less than $k \in (0, m)$. Then there exists a $w^* \in U$ such that $\int |sg(w^*) + tm(w^*) + c|^{-2} \mathrm{d}H(s,t) = \infty$.*

Proof. By assumptions, for any support point (s,t) such that $s + t > \phi > 0$ where ϕ is the distance of the support og H to the origin, we have

$$\int_0^{\eta} (sg'(w) + tm'(w))\mathrm{d}w \geq k\phi\eta > 0 \tag{30}$$

and similiarly

$$\int_0^{-\eta} (sg'(w) + tm'(w))\mathrm{d}w \leq -k\phi\eta < 0$$

Without loss of generality, we may assume that $c = 1$. Since (s_0, t_0) is a support point of H, which is not the origin, let the constant $d > 0$ be such that $d < k\phi\eta / \sqrt{2(g_0^2 + m_0^2)}$, and define R_0 to be a square contained in the first quadrant, containing the point (s_0, t_0), the edge length $d > 0$ and does not cover the origin. There is a positive constant $\alpha > 0$ such that $H(R_0) = \alpha$.

Remark 2. If (s_0, t_0) is on one coordinate axis, we can take (s_0, t_0) as an inner point of one edge of R_0. If (s_0, t_0) is an inner point of the first quadrant, then we can take (s_0, t_0) as an inner point of R_0. In either way, we may guarantee $H(R_0) > 0$.

Split the square into four squares by equally dividing each edge into two. Denote one with H measure no less than $\alpha/4$ by R_1. Inductively, split R_n similarly into four small squares and denote the one with H measure no less than $\alpha 4^{-n-1}$ by R_{n+1}. Note that for any n, $H(R_n) \geq 4^{-n}\alpha$. Write the center of R_n as (s_n, t_n) for all $n \geq 1$. By the nested interval theorem, (s_n, t_n) tends to a limit $(s^*, t^*) \in R_n$, for all n. Also, by the construction of the squares, we know that $\|(s_{n+1} - s_n, t_{n+1} - t_n)\| = \sqrt{2} 2^{-n-1} d$.

Write $Q_n(w) = s_n g(w) + t_n m(w) + 1$.

$$\begin{aligned} Q_n(w) =& Q_n(w_0) + \int_{w_0}^{w} Q_n'(w)\mathrm{d}w \\ =& ((s_n - s_0)g_0 + (t_n - t_0)m_0 + \int_{w_0}^{w} Q_n'(w)\mathrm{d}w) \\ & \begin{cases} \geq & -d\sqrt{2(g_0^2 + m_0^2)} + k\phi\eta > 0 \text{ when } w = w_0 + \eta, \\ \leq & d\sqrt{2(g_0^2 + m_0^2)} - k\phi\eta < 0 \text{ when } w = w_0 - \eta. \end{cases} \end{aligned}$$

Therefore, we may select $w_n \in U$ such that $s_n g(w_n) + t_n m(w_n) + 1 = 0$. Similarly, we may select $w^* \in U$such that $u^* g(w^*) + t^* s(w^*) + 1 = 0$.

For each $(s,t) \in R_n$, $\|(s - s^*, t - t^*)\| \leq \sqrt{2} 2^{-n} d$. Thus, we have

$$|sg^* + tm^* + 1| = |(s - s^*)g^* + (t - t^*)m^*| \leq \sqrt{2} 2^{-n} d \sqrt{(g^*)^2 + (m^*)^2} := M2^{-n}.$$

If there are infinitely many n such that $H(R_n) \geq H(R_{n-1})/2$ and suppose n_k is the kth such n, then as $k \to \infty$

$$\begin{aligned} &\int (sg^* + tm^* + 1)^{-2} \mathrm{d}H(s,t) \geq \int_{R_{n_k}} (sg^* + tm^* + 1)^{-2} \mathrm{d}H(u,t) \\ &\geq H(R_{n_k}) M^2 2^{2n_k} \geq 2^k 4^{-n_k} \alpha M^2 2^{2n_k} \geq M^2 \alpha 2^k \to \infty. \end{aligned}$$

Otherwise, there is an N_0 such that for all $n \geq N_0$, $H(R_{n-1} - R_n) \geq H(R_n)$. Therefore,

$$\int (sg^* + tm^* + 1)^{-2} \mathrm{d}H(s,t) \geq \sum_{n=1}^{\infty} \int_{R_{n-1}-R_n} (sg^* + tm^* + 1)^{-2} \mathrm{d}H(s,t)$$
$$\geq \sum_{n=1}^{\infty} H(R_{n-1} - R_n) M^2 2^{2n} \geq M^2 \sum_{n=N_0+1}^{\infty} H(R_n) 2^{2n} \geq M^2 \sum_{n=N_0+1}^{\infty} 4^{-n} \alpha 2^{2n} = \infty.$$

The proof is complete. □

4. Appendix

In this section we do not assume any restrictions on H, and c imposed in Theorem 1.

We begin with establishing some results on sequences of probability distribution functions. It centers on the Lévy distance $L(F,G)$ between two distribution functions defined as

$$L(F,G) = \inf\{\delta : G(x-\delta) - \delta \leq F(x) \leq G(x+\delta) + \delta \quad \text{for all x}\}.$$

It is a metric on the set of all distribution functions yielding weak convergence: G_n converging weakly to G if and only if $d(G_n, G) \to 0$. It follows that

$$L(F,G) \leq \|F - G\|,$$

where $\|\cdot\|$ is the sup norm on functions. Also, from Corollary A.42 and Theorem A.44 of [2] we have, if A and B are both $n \times N$, then

$$L^4(F^{AA^*}, F^{BB^*}) \leq \frac{2}{n^2}(\mathrm{tr}(AA^* + BB^*))(\mathrm{tr}[(A-B)(A-B)^*]) \tag{31}$$

and

$$\|F^{AA^*} - F^{BB^*}\| \leq \frac{1}{n}\mathrm{rank}(A-B). \tag{32}$$

Since the rank of a matrix A is equal to the dimension of its row space, we have

$$\mathrm{rank}(A) \leq \text{ the number of nonzero entries of } A. \tag{33}$$

The following rank inequalities are well-known: For matrices A B of the same dimensions

$$\mathrm{rank}(A+B) \leq \mathrm{rank}(A) + \mathrm{rank}(B). \tag{34}$$

For A B for which AB is defined

$$\mathrm{rank}(AB) \leq \min(\mathrm{rank}(A), \mathrm{rank}(B)). \tag{35}$$

We will also need the fact that, for Hermitian A, B, with B nonnegative definite

$$|\mathrm{tr}AB| \leq \|A\|\mathrm{tr}B, \tag{36}$$

where $\|\cdot\|$ is the spectral norm.

The following extends Lemmas 4.7 and 4.8 of [2].

Lemma 2. *Let $\{F_n\}$ be a sequence of distribution functions, and for each $\epsilon > 0$ there exists a tight sequence of distribution functions $\{F_{n,\epsilon}\}$ such that for each $\delta > 0$ there exists $\epsilon = \epsilon(\delta)$ for which*

$$\limsup_n L(F_{n,\epsilon}, F_n) < \delta. \tag{37}$$

Then the sequence $\{F_n\}$ is tight. Moreover, if for each ϵ $F_{n,\epsilon}$ converges weakly to F_ϵ, then F_n converges weakly to F with $F_\epsilon = F_{\epsilon(\delta)}$ converging weakly to F as $\delta \to 0$.

Proof. . For any δ we have $L(F_{n,\epsilon(\delta/2)}, F_n) < \delta/2$ for all n large. Choose x so that $F_{n,\epsilon(\delta/2)}(x) > 1-\delta/2$ for all n. Then for all n large

$$1-\delta/2 < F_{n,\epsilon(\delta/2)}(x) \le F_n(x+\delta/2)+\delta/2.$$

Therefore $1-\delta < F_n(x+\delta/2)$, and since δ is arbiitrary we see that $\{F_n\}$ is tight.

If $F_{n,\epsilon}$ converges weakly to F_ϵ, suppose F_1 and F_2 are two distributions, each being weak limits of F_n along two different subsequences $\{n'\}$, $\{n''\}$.. Then for any ϵ

$$\begin{aligned} L(F_1,F_2) &= \lim_{n',n"\to\infty} L(F_{n'},F_{n''}) \le \limsup_{n'\to\infty} L(F_{n'},F_{n',\epsilon}) + \limsup_{n''\to\infty} L(F_{n''},F_{n'',\epsilon}) \\ &\quad + \lim_{n'\to\infty} L(F_\epsilon,F_{n',\epsilon}) + \limsup_{n''\to\infty} L(F_\epsilon,F_{n'',\epsilon}) = \limsup_{n'\to\infty} L(F_{n'},F_{n',\epsilon}) + \limsup_{n''\to\infty} L(F_{n''},F_{n'',\epsilon}), \end{aligned} \quad (38)$$

so we see that $L(F_1,F_2)$ can be made arbitrarily small. Therefore we have weak convergence of F_n to some distribution function F. Finally we see that

$$L(F,F_\epsilon) = \lim_{n\to\infty} L(F_n,F_\epsilon) \le \lim_{n\to\infty} L(F_n,F_{n,\epsilon}) + \lim_{n\to\infty} L(F_{n,\epsilon},F_\epsilon) = \lim_{n\to\infty} L(F_n,F_{n,\epsilon}),$$

which can be made arbitrarily small. □

We begin by truncating R_n and T_n to matrices of bounded norm for all n.

Since $(1/N)R_nR_n^*$ and T_n commute, there exists unitary $\mathbf{U}_n$ which simultaneously diagonalizes these two matrices:

$$(1/N)R_nR_n^* = \mathbf{U}_n\text{diag})(s_1,\ldots,s_n)\mathbf{U}_n^*, \quad T_n = \mathbf{U}_n\text{diag}(t_1,\ldots,t_n)\mathbf{U}_n^*.$$

Write $R_n = \sqrt{N}\mathbf{U}_n\text{diag}(s_1^{1/2},\ldots,s_n^{1/2})\mathbf{V}_n$, where $\mathbf{V}_n$ is $N\times N$ unitary, and $\text{diag}(s_1^{1/2},\ldots,s_n^{1/2})$ is $n\times N$ diagonal. Since H_n converges weakly to H, for any $\varepsilon>0$, there exists a constant $\tau=\tau_\epsilon>0$ such that the number of eigenvalues of $(1/N)R_nR_n^*$ larger than τ and the number eigenvalues of $\mathbf{T}_n$ larger than τ are both less than $n\varepsilon/6$. Let

$$R_{n,\varepsilon} = \sqrt{N}\mathbf{U}_n\text{diag}((\min(s_1,\tau))^{1/2},\ldots,(\min(s_n,\tau))^{1/2})\mathbf{V}_n,$$
$$T_{n,\varepsilon}^{1/2} = \mathbf{U}_n\text{diag}((\min(t_1,\tau))^{1/2},\ldots,(\min(t_n,\tau))^{1/2})\mathbf{U}_n^*,$$

and define

$$C_{n,\varepsilon} = (1/N)T_{n,\varepsilon}^{1/2}(R_{n,\varepsilon}+X_n)(R_{n,\varepsilon}+X_n)^*T_{n,\varepsilon}^{1/2},$$
$$B_{n,\varepsilon} = (1/N)(R_{n,\varepsilon}+T_{n,\varepsilon}^{1/2}X_n)(R_{n,\varepsilon}+T_{n,\varepsilon}^{1/2}X_n)^*.$$

Then, using (32), (34), and (35) we have

$$\max(\|F^{C_n}-F^{C_{n,\varepsilon}}\|,\|F^{B_n}-F^{B_{n,\varepsilon}}\|) \le \frac{1}{n}(\text{rank}(R_n-R_{n,\varepsilon})+\text{rank}(T_n-T_{n,\varepsilon})) \le \varepsilon/3 \quad (39)$$

We turn now to truncating and centralizing the entries of X_n. Since the Lindeberg condition hold for any $\eta>0$, we can find a sequence $\eta_n\to 0$ for which $\eta_n\sqrt{n}\to\infty$ and

$$\frac{1}{nN\eta_n^2}\sum_{i=1}^n\sum_{j=1}^N \mathbf{E}|x_{ij}^2|I(|x_{ij}|>\eta_n\sqrt{n}) \to 0. \quad (40)$$

Let $\hat{x}_{ij} = x_{ij}I(|x_{ij}|<\eta_n\sqrt{n})$ and construct $\widehat{C}_{n,\varepsilon}$, $\widehat{B}_{n,\varepsilon}$ similarly as $C_{n,\varepsilon}$ $B_{n,\varepsilon}$ with x_{ij} replaced by $\hat{x}_{ij}$.

Using (32), (34), and (35) we have

$$\max(\|F^{\widehat{C}_{n,\varepsilon}}-F^{C_{n,\varepsilon}}\|,|F^{\widehat{B}_{n,\varepsilon}}-F^{B_{n,\varepsilon}}\|) \le \frac{1}{n}\text{rank}(\mathbf{X}_n-\widehat{\mathbf{X}}_n) \le \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^N I(|x_{ij}|\ge\eta_n\sqrt{n})$$

Note that

$$\mathrm{Var}\left(\sum_{i=1}^{n}\sum_{j=1}^{N} I(|x_{ij}| \geq \eta_n \sqrt{n})\right) \leq \mathbf{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{N} I(|x_{ij}| \geq \eta_n \sqrt{n})\right)$$

$$\leq \frac{1}{n\eta_n^2}\sum_{i=1}^{p}\sum_{j=1}^{n} \mathbf{E}|x_{ij}|^2 |I(|x_{ij} \geq \eta_n \sqrt{n}) = o(n).$$

Therefore, by Bernstein inequality (p. 21 of [2]) we have

$$\Pr\left(\max(\|F^{\widehat{C}_{n,\varepsilon}} - F^{C_{n,\varepsilon}}\|\|, F^{\widehat{B}_{n,\varepsilon}} - F^{B_{n,\varepsilon}}\|) \geq \varepsilon/6\right) \leq 2\exp\left(\frac{-(\varepsilon n/12)^2}{o(n) + \varepsilon n/12}\right) \leq 2e^{-\delta n}, \tag{41}$$

where δ is some positive constant. The right hand side is summable, hence,

$$\max(\|F^{\widehat{C}_{n,\varepsilon}} - F^{C_{n,\varepsilon}}\|, \|F^{\widehat{B}_{n,\varepsilon}} - F^{B_{n,\varepsilon}}\|) \leq \varepsilon/6, a.s. \text{ for all large } n. \tag{42}$$

Define

$$\tilde{x}_{ij} = \begin{cases} (\hat{x}_{ij} - \mathbf{E}\hat{x}_{ij})/\sigma_{ij}, & \text{if } \sigma_{ij}^2 = \mathbf{E}|\hat{x}_{ij} - \mathbf{E}\hat{x}_{ij}|^2 \geq 1/2, \\ y_{ij}, & \text{otherwise,} \end{cases}$$

where y_{ij} are iid. random variables taking values ± 1 with equal probabilities. Define $\widetilde{C}_n$ and $\widetilde{B}_n$ with $\tilde{x}_{ij}$ replacing $\hat{x}_{ij}$. Then, from (31) and (36) we have

$$L^4(F^{\widehat{C}_{n,\varepsilon}}, F^{\widetilde{C}_{n,\varepsilon}}) \leq \frac{2\tau}{n^2 N}(\mathrm{tr}\widehat{C}_{n,\varepsilon} + \mathrm{tr}\widetilde{C}_{n,\varepsilon})\mathrm{tr}(\widehat{X}_n - \widetilde{X}_n)(\widehat{\mathbf{X}}_n - \widetilde{\mathbf{X}}_n)^* \tag{43}$$

$$L^4(F^{\widehat{B}_{n,\varepsilon}}, F^{\widetilde{B}_{n,\varepsilon}}) \leq \frac{2\tau}{n^2 N}(\mathrm{tr}\widehat{B}_{n,\varepsilon} + \mathrm{tr}\widetilde{B}_{n,\varepsilon})\mathrm{tr}(\widehat{X}_n - \widetilde{X}_n)(\widehat{X}_n - \widetilde{X}_n)^*.$$

Using (36) and the fact that $(a+b)^2 \leq 2(a^2+b^2)$ we have

$$n^{-1}\mathrm{tr}\widehat{C}_{n,\varepsilon} \leq 2\tau(nN)^{-1}\mathrm{tr}(R_{n,\varepsilon}R_{n,\varepsilon}^* + \widehat{X}_n\widehat{X}_{n,}^*)$$

$$n^{-1}\mathrm{tr}\widetilde{C}_{n,\varepsilon} \leq 2(nN)^{-1}\mathrm{tr}(R_{n,\varepsilon}R_{n,\varepsilon}^* + \widetilde{X}_n\widetilde{X}_n^*)$$

$$n^{-1}\mathrm{tr}\widehat{B}_{n,\varepsilon} \leq 2(nN)^{-1}\mathrm{tr}(R_{n,\varepsilon}R_{n,\varepsilon}^* + \tau\widehat{X}_n\widehat{X}_n^*)$$

$$n^{-1}\mathrm{tr}\widetilde{B}_{n,\varepsilon} \leq 2(nN)^{-1}\mathrm{tr}(R_{n,\varepsilon}R_{n,\varepsilon}^* + \widetilde{X}_n\widetilde{X}_n^*).$$

We have $(nN)^{-1}\mathrm{tr}R_{n,\varepsilon}R_{n,\varepsilon}^* \leq \tau$. We claim that both $(nN)^{-1}\mathrm{tr}\widehat{X}_n\widehat{X}_n^*$ and $(nN)^{-1}\mathrm{tr}\widetilde{X}_n\widetilde{X}_n^*$ converge a.s. to 1. We have

$$(nN)^{-1}\mathbf{E}\mathrm{tr}\widehat{X}_n\widehat{X}_n^* = 1 - (nN)^{-1}\sum_{ij}\mathbf{E}|x_{ij}|^2 I(|x_{ij}| \geq \eta_n\sqrt{n}) \to 1.$$

Since $\mathbf{E}|x_{ij}|^k I(|x_{ij}| < \delta_n\sqrt{n}) \leq n^{\frac{k-2}{2}}$ for $k \geq 2$, the fourth central moment of each quantity is bounded by

$$Kn^{-8}(n^2n^3 + n^4n^2),$$

which is summable. The claim is proven.

We have

$$\frac{1}{nN}\mathrm{tr}(\widehat{\mathbf{X}}_n-\widetilde{\mathbf{X}}_n)(\widehat{\mathbf{X}}_n-\widetilde{\mathbf{X}}_n)^*=\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}|\hat{x}_{ij}-\tilde{x}_{ij}|^2:=I_1+I_2,$$

where

$$I_1=\frac{1}{nN}\sum_{\sigma_{ij}^2\geq 1/2}|\hat{x}_{ij}-\tilde{x}_{ij}|^2$$
$$I_2=\frac{1}{nN}\sum_{\sigma_{ij}^2<1/2}|\hat{x}_{ij}-y_{ij}|^2$$

We see that

$$\begin{aligned}
\mathbf{E}I_1&=\frac{1}{nN}\sum_{\sigma_{ij}^2\geq 1/2}\mathbf{E}\left|\frac{\mathbf{E}x_{ij}I(|x_{ij}|\geq\eta_n\sqrt{n})}{\sigma_{ij}}-\hat{x}_{ij}\frac{\sigma_{ij}-1}{\sigma_{ij}}\right|^2\leq\frac{2}{nN}\sum_{\sigma_{ij}^2\geq 1/2}\left(\frac{1}{\eta_n^2 n}+\left(\sigma_{ij}-1\right)^2\right)\\
&\leq\frac{2}{nN}\sum_{\sigma_{ij}^2\geq 1/2}\left(\frac{1}{\eta_n^2 n}+\left(\mathbf{E}|x_{ij}|^2I(|x_{ij}|\leq\eta_n\sqrt{n})-1-|\mathbf{E}x_{ij}I(|x_{ij}|\leq\eta_n\sqrt{n})|^2\right)^2\right)\\
&=\frac{2}{nN}\sum_{\sigma_{ij}^2\geq 1/2}\left(\frac{1}{\eta_n^2 n}+\left(\mathbf{E}|x_{ij}|^2I(|x_{ij}|>\eta_n\sqrt{n})+|\mathbf{E}x_{ij}I(|x_{ij}|>\eta_n\sqrt{n})|^2\right)^2\right)\\
&\leq\frac{2}{nN}\sum_{\sigma_{ij}^2\geq 1/2}\left(\frac{1}{\eta_n^2 n}+4\mathbf{E}|x_{ij}|^2I(|x_{ij}|>\eta_n\sqrt{n})\right)\to 0
\end{aligned}$$

We have

$$\begin{aligned}
\mathbf{E}|I_1-\mathbf{E}I_1|^4&\leq\frac{K}{n^8}\Big(\sum_{ij}\left||\hat{x}_{ij}-\tilde{x}_{ij}|^2-\mathbf{E}|\hat{x}_{ij}-\tilde{x}_{ij}|^2\right|^4\\
&\quad+\sum_{\{i,j\}\neq\{i',j'\}}\mathbf{E}\left||\hat{x}_{ij}-\tilde{x}_{ij}|^2-\mathbf{E}|\hat{x}_{ij}-\tilde{x}_{ij}|^2\right|^2\mathbf{E}\left||\hat{x}_{i'j'}-\tilde{x}_{i'j'}|^2-\mathbf{E}|\hat{x}_{i'j'}-\tilde{x}_{i,j'}|^2\right|^2\Big)\\
&\leq\frac{K}{n^8}(nN\times n^3+(nN)^2n^2),
\end{aligned}$$

which is summable and hence $I_1\to 0, a.s.$ When $\sigma_{ij}^2<1/2$, we have

$$1/2>\mathbf{E}|x_{ij}|^2I(|x_{ij}|\leq\eta_n\sqrt{n})-|\mathbf{E}x_{ij}I(|x_{ij}|\leq\eta_n\sqrt{n})|^2=1-\mathbf{E}|x_{ij}|^2I(|x_{ij}|>\eta_n\sqrt{n})-|\mathbf{E}x_{ij}I(|x_{ij}|>\eta_n\sqrt{n})|^2,$$

which implies

$$\mathbf{E}|x_{ij}^2|I(|x_{ij}|\geq\eta_n\sqrt{n})\geq\frac{1}{2}-o(1)\geq\frac{1}{3}$$

for all n large Denote by N the number of pairs (i,j) such that $\sigma_{ij}^2<1/2$. When $\sigma_{ij}^2<1/2$, we have

$$\sum_{\sigma_{ij}^2<1/2}\mathbf{E}|x_{ij}^2|I(|x_{ij}|\geq\eta_n\sqrt{n})\geq N/3.$$

Combining this with

$$\sum_{\sigma_{ij}^2<1/2}\mathbf{E}|x_{ij}^2|I(|x_{ij}|\geq\eta_n\sqrt{n})\leq o(nN),$$

we obtain $N = o(nN)$. We have

$$I_2 \le \frac{2}{nN} \sum_{\sigma_{ij}^2<1/2} |\hat{x}_{ij}|^2 + \frac{o(nN)}{nN}.$$

As above, the fourth central moment of $\frac{1}{nN}\sum_{\sigma_{ij}^2<1/2} |\hat{x}_{ij}|^2$ is summable, but now

$$\frac{1}{nN}\sum_{\sigma_{ij}^2<1/2} \mathbf{E}|\hat{x}_{ij}|^2 = \frac{1}{nN}\Big(o(nN) - \sum_{\sigma_{ij}^2<1/2} \mathbf{E}|\hat{x}_{ij}|^2 I(|x_{ij}| \ge \eta_n\sqrt{n})\Big) \to 0.$$

We have at this stage that for any $\epsilon > 0$, almost surely

$$\limsup_n \max(L(C_{n,\epsilon}, \widetilde{C}_{n,\epsilon}), L(B_{n,\epsilon}, \widetilde{B}_{n,\epsilon})) \le \epsilon/3.$$

The results in [14] can now be applied to C_n and $C_{n,\epsilon}$. Let $g_{n,\epsilon}(z) = (1/n)\mathrm{tr}(C_{n,\epsilon} - zI)^{-1}T_{n,\epsilon}$. Assume (x, y) is a random vector with joint distribution function $H(s,t)$, Denote by H^ϵ the joint distribution function of $(\min(x,\tau), \min(y,\tau))$. Then from [14] we have, for all positive c and $z \in \mathbb{C}^+$, almost surely, for any converging subsequence of $(m^{C_{n,\epsilon}}(z), g_{n,\epsilon}(z))$ to (m, g), this pair must satisfy (3) with H replaced by H^ϵ. It amounts to verify that there is only one m satisfying the equations for an infinite number of $z \in \mathbb{C}^+$ with an accumulation point.

We turn our attention to B_n, its definition also including C_n. We have the transition from (3) to (6) when T_n remains invertible. To this end we define (without loss of generality we can assume $\epsilon < \tau$)

$$\widetilde{T}_{n,\varepsilon}^{1/2} = \mathbf{U}_n \mathrm{diag}((\max(\epsilon, \min(t_1,\tau)))^{1/2}, \ldots, (\max(\epsilon, \min(t_n,\tau)))^{1/2})\mathbf{U}_n^*,$$

and define

$$\widetilde{\widetilde{B}}_{n,\varepsilon} = (1/N)(R_{n,\varepsilon} + \widetilde{T}_{n,\varepsilon}^{1/2}\widetilde{X}_n)(R_{n,\varepsilon} + \widetilde{T}_{n,\varepsilon}^{1/2}\widetilde{X}_n)^*.$$

Then from (31) and (36) we have

$$L^4(F^{\widetilde{\widetilde{B}}_{n,\varepsilon}}, F^{\widetilde{B}_{n,\varepsilon}}) \le \frac{2}{Nn^2}(\mathrm{tr}\widetilde{\widetilde{B}}_{n,\varepsilon} + \mathrm{tr}\widetilde{B}_{n,\varepsilon})\mathrm{tr}(T_{n,\varepsilon}^{1/2} - \widetilde{T}_{n,\varepsilon}^{1/2})\widetilde{X}_n\widetilde{X}_n^*(T_{n,\varepsilon}^{1/2} - \widetilde{T}_{n,\varepsilon}^{1/2}) \le \frac{K\epsilon}{nN}\mathrm{tr}\widetilde{\mathbf{X}}_n\widetilde{\mathbf{X}}_n^* \to K\epsilon, a.s.$$

We see that, almost surely, the assumption (37) in Lemma 2 is met.

To show uniqueness we switch our concentration to the matrix $\underline{B}_n$ given in (7). As previously mentioned, the eigenvalues B_n are same as those of $\underline{B}_n$. except for $|n - N|$ zero eigenvalues. The relation between their ESDs and Stieltjes transforms are given in (8). After making the variable transformations in (9) we arrive at the equations in (10).

Fix $z = x+iv$, $v > 0$. Notice that, when (m, g) are finite limits of $(m^{B_n}(z), g_n(z))$ then necessarily $\Im\underline{m}$, $\Im\underline{g}$, $\Im(z\underline{m})$, $\Im(z\underline{g})$ are all nonnegative.

Suppose $(\underline{m}, \underline{g})$ is such a set of solutions to (10). Recall the identity (17) in section 2. In the case where H has bounded support, we see that we can make B_0 arbitrarily close to 1, and therefore the quantity $(1-c)v + cB_0v$ positive, for suitably small $\underline{m}$ and $\underline{g}$. Therefore for these values of $\underline{m}$ and $\underline{g}$ we have (18)

Redefining $B_{n,\epsilon}$ to be $B_{n,\epsilon} = (1/N)(R_{n,\epsilon} + \widetilde{T}_{n,\epsilon}X_n)(R_{n,\epsilon} + \widetilde{T}_{n,\epsilon}X_n)^*$, we use the results on $C_{n,\epsilon}$ to conclude that, almost surely, for any weakly convergent subsequence of $F^{B_{n,\epsilon}}$, the corresponding m will satisfy (6) for some g, and therefore the corresponding $(\underline{m}, \underline{g})$ will satisfy (10).

Since $\widetilde{T}_{n,\epsilon}$ has bounded spectral norm for all n, we see that, with $g_{n,\epsilon}$ now equal to $(1/n)\mathrm{tr}(B_{n,\epsilon} - zI)^{-1}\widetilde{T}_{n,\epsilon}$,

$$\frac{1}{\frac{1}{n}\mathrm{tr}\widetilde{T}_{n,\epsilon}} g_{n,\epsilon}(z)$$

is the Stieltjes transform of a probability measure. Consequently, from (9)

$$\underline{g}_{n,\epsilon}(z) \equiv -\frac{1}{z(1 + c_n g_{n,\epsilon}(z))}$$

satisfies the conditions of being the Stieltjes transform of a probability measure $P_{n,\epsilon}$ with mass on the nonnegative reals. Any vaguely convergent subsequence of $P_{n,\epsilon}$ will have the limiting Stieltjes transform also of this form, which is the Stieltjes transform of a probability measure concentrated on the nonnegative reals. Therefore the convergence is weak.

Suppose, with probability one, there is a subsequence for which both $F^{B_{n,\epsilon}}$ and $P_{n,\epsilon}$ converge weakly, with resulting limiting Stieltjes transforms $\underline{m}(z)$ and $g(z)$, which satisfy (10) for all $z = x + iv \in \mathbb{C}^+$. Since Stieltjes transforms of probability measures are bounded by $1/v$, we can find v suitably large so that (18) holds.

Suppose for one of these z there are two sets of solutions $(\underline{m}_{(i)}, \underline{g}_{(i)})$, $i = 1, 2$, resulting. almost surely, from two weakly converging subsequences of $F^{B_{n,\epsilon}}$ and $P_{n,\epsilon}$ with $\underline{m}_{(1)} \neq \underline{m}_{(2)}$. We have then (25), except there is a strict inequality at the last step, resulting in a contradiction.

Thus we have unique Stieltjes transforms for any a.s. weakly converging subsequence of $F^{B_{n,\epsilon}}$ and $P_{n,\epsilon}$, so that, with probability one, $F^{B_{n,\epsilon}}$ and $P_{n,\epsilon}$ converge weakly to a nonrandom probability distribution function F^ϵ and a nonrandom probabilty measure P^ϵ with limiting Stieltjes transforms satisfying (6). From Lemma 2 we have almost surely F^{B_n} converging weakly to a nonrandom distribution function F, which implies that F^{C_n} also converges a.s. weakly to a nonrandom distribution function.

It amounts to showing that this limiting distribution satisfies the equations. With x, y denoting random variables with joint distribution function $H(s,t)$, we let $H^\epsilon(s,t)$ denote the joint distribution function of the random variables $\min(x, \tau_\epsilon)$,$\max(\epsilon, \min(y, \tau_\epsilon))$. Let F^ϵ be the distribution function associated with H^ϵ. Then as $\epsilon \to 0$, F^ϵ and H^ϵ converge in distribution to F and H, respectively. Fix $z = x + iv \in \mathbb{C}^+$. With $m^\epsilon = m^\epsilon(z)$, $\underline{m}^\epsilon = \underline{m}^\epsilon(z)$, $g^\epsilon = g^\epsilon(z)$, $\underline{g}^\epsilon = \underline{g}^\epsilon(z)$ denoting the values in equations (6) and (10), since F^ϵ converges in distribution to F, we immediately get m^ϵ and $\underline{m}^\epsilon$ converging to m, the Stieltjes transform of F at z, and to $\underline{m} \equiv -\frac{1-c}{z} + cm$, respectively.

We claim that g^ϵ remains bounded as $\epsilon \to 0$. We have $\underline{m}_2$ the imaginary part of $\underline{m}$ is positive, since $\underline{m}$ is the value of a Stieljes transform for $z \in \mathbb{C}^+$. Let $\delta > 0$ be a lower bound on $\underline{m}_2^\epsilon$. Then the integrand in the second equation in (10) satisfies

$$\left| \frac{t}{1 + s\underline{g} + t\underline{m}} \right| = \left| \frac{t}{1 + s\underline{g} + t\underline{m}_1 + it\underline{m}_2} \right| \leq \frac{1}{\delta}.$$

Suppose on a sequence ϵ_n g^{ϵ_n} goes unbounded. Then necessarily $\underline{g}^{\epsilon_n} \to 0$. But from the second equation of (10), we see the right side goes unbounded, while the left side remains at $\overline{z}$, a contradiction.

On a sequence ϵ_n let $g = \lim_{n\to\infty} g^{\epsilon_n}$. Notice the integrand in the first equation in (6) is bounded in absolute value by $1/v$. We have

$$\begin{aligned}
&\int \frac{\mathrm{d}H^{\epsilon_n}(s,t)}{\frac{s}{1+cg^{\epsilon_n}} - (1 + cm^{\epsilon_n})z + t(1-c)} - \int \frac{\mathrm{d}H(s,t)}{\frac{s}{1+cg} - (1 + cm)z + t(1-c)} \\
=& \int \frac{\left(\frac{s}{1+cg} - \frac{s}{1+cg^{\epsilon_n}} + m^{\epsilon_n} - m\right) \mathrm{d}H^{\epsilon_n}(s,t)}{\left(\frac{s}{1+cg^{\epsilon_n}} - (1 + cm^{\epsilon_n})z + t(1-c)\right)\left(\frac{s}{1+cg} - (1 + cm)z + t(1-c)\right)} \\
&+ \int \frac{\mathrm{d}H^{\epsilon_n}(s,t)}{\frac{s}{1+cg} - (1 + cm)z + t(1-c)} - \int \frac{\mathrm{d}H(s,t)}{\frac{s}{1+cg} - (1 + cm)z + t(1-c)} \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

On the same sequence ϵ_n let $\underline{g} = \lim_{n\to\infty} \underline{g}^{\epsilon_n}$.. Since the relationship between g and $\underline{g}$ in (9) still holds, we see that, since the absolute value of the denominator in the right side is at least v, we must have $\underline{g}_2 > 0$. Let now $\delta > 0$ be a lower bound on $\underline{m}_2^\epsilon$ and $\underline{g}_2^{\epsilon_n}$. For the second equation in (10) we have

$$\begin{aligned}
&\limsup_n \left| \int \frac{t\mathrm{d}H^{\epsilon_n}(s,t)}{1 + s\underline{g}^{\epsilon_n} + t\underline{m}^{\epsilon_n}} - \int \frac{t\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}} \right| \\
\leq& \limsup_n \left| \int \frac{(s(\underline{g} - \underline{g}^{\epsilon_n}) + t(\underline{m} - \underline{m}^{\epsilon_n}))t\mathrm{d}H^{\epsilon_n}(s,t)}{\left(1 + s\underline{g}^{\epsilon_n} + t\underline{m}^{\epsilon}\right)\left(1 + s\underline{g} + t\underline{m}\right)} \right| + \limsup_n \left| \int \frac{t\mathrm{d}H^{\epsilon}(s,t)}{1 + s\underline{g} + t\underline{m}} - \int \frac{t\mathrm{d}H(s,t)}{1 + s\underline{g} + t\underline{m}} \right| \\
\leq& \frac{1}{\delta^2} \limsup_n (|\underline{g} - \underline{g}^{\epsilon_n}| + |\underline{m} - \underline{m}^{\epsilon_n}|) = 0.
\end{aligned}$$

We conclude that the limiting F has, for each $z \in \mathbb{C}^+$, its Stieltjes transform $m = m(z)$ satisfying (6) for some g with $\Im g \geq 0$, with corresponding $\underline{m}$, $\underline{g}$ satisfying (10).

Acknowledgments

J. Hu was supported by NSFC (No. 12171078, 11971097) and National Key R & D Program of China (No. 2020YFA0714102). Z. D. Bai was partially supported by NSFC Grant 12171198 and Team Project of Jilin Provincial Department of Science and Technology (No.20210101147JC).

[1] Z. Bai, J. Silverstein, No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices, Random Matrices: Theory and Applications 1 (2012) 1150004.

[2] Z. Bai, J. W. Silverstein, Spectral analysis of large dimensional random matrices, volume 20, Springer, 2010.

[3] M. Capitaine, Exact separation phenomenon for the eigenvalues of large information-plus-noise type matrices, and an application to spiked models, Indiana University Mathematics Journal (2014) 1875–1910.

[4] R. Dozier, J. Silverstein, Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices, Journal of Multivariate Analysis 98 (2007) 1099–1122.

[5] R. Dozier, J. Silverstein, On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices, Journal of Multivariate Analysis 98 (2007) 678–694.

[6] M. K. Krein, A. A. Nudelman, The Markov Moment Problem and Extremal Problems, in: [6], 1997.

[7] P. Loubaton, P. Vallet, Almost sure localization of the eigenvalues in a gaussian information plus noise model. application to the spiked models., Electronic Journal of Probability 16 (2011) 1934–1959.

[8] V. Marčenko, L. Pastur, The eigenvalue distribution in some ensembles of random matrices, Math. USSR Sbornik 1 (1967) 457–483.

[9] J. A. Shohat, J. D. Tamarkin, The Problem of Moments, in: [6], 1970.

[10] J. Silverstein, Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices, Journal of Multivariate Analysis 55 (1995) 331–339.

[11] J. Silverstein, Choi, Analysis of the limiting spectral distribution of large dimensional random matrices, Journal of Multivariate Analysis 54 (1995) 295–309.

[12] E. Wigner, On the distribution of the roots of certain symmetric matrices, Annals of Mathematics (1958) 325–327.

[13] Y. Yin, Limiting spectral distribution for a class of random matrices, Journal of multivariate analysis 20 (1986) 50–68.

[14] H. Zhou, Z. Bai, J. Hu, The limiting spectral distribution of large dimensional general information-plus-noise type matrices, Journal of Theoretical Probability (2022).