Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2012, Vol. 48, No. 2, 444–476
DOI: 10.1214/11-AIHP414

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Central limit theorems for linear spectral statistics of large dimensional F-matrices

Shurong Zheng

KLAS and School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin Province, 130024, P.R. China.
E-mail: zhengsr@nenu.edu.cn; zhengsr1993@gmail.com

Received 24 June 2010; revised 7 October 2010; accepted 18 January 2011

Abstract. In many applications, one needs to make statistical inference on the parameters defined by the limiting spectral distribution of an F matrix, the product of a sample covariance matrix from the independent variable array $(X_{jk})_{p\times n_1}$ and the inverse of another covariance matrix from the independent variable array $(Y_{jk})_{p\times n_2}$. Here, the two variable arrays are assumed to either both real or both complex. It helps to find the asymptotic distribution of the relevant parameter estimators associated with the F matrix. In this paper, we establish the central limit theorems with explicit expressions of means and covariance functions for the linear spectral statistics of the large dimensional F matrix, where the dimension p of the two samples tends to infinity proportionally to the sample sizes (n_1, n_2). Moreover, the assumptions of the i.i.d. structures of arrays $(X_{jk})_{p\times n_1}, (Y_{jk})_{p\times n_2}$ and the restriction of the fourth moments equaling 2 or 3 made in Bai and Silverstein (*Ann. Probab.* **32** (2004) 553–605) are relaxed to that arrays $(X_{jk})_{p\times n_1}$ and $(Y_{jk})_{p\times n_2}$ are independent respectively but not necessarily identically distributed except for a common fourth moment for each array. As a consequence, we obtain the central limit theorems for the linear spectral statistics of the beta matrix that is of the form $(\mathbf{I} + d\cdot F \text{ matrix})^{-1}$, where d is a constant and $\mathbf{I}$ is an identity matrix.

Résumé. Dans beaucoup d'applications, on cherche à effectuer une inférence statistique sur des paramètres définis à partir de la mesure spectrale d'une F-matrice, matrice obtenue comme le produit d'une matrice de covariance du tableau de variables indépendantes $(X_{jk})_{p\times n_1}$ et de l'inverse d'une autre matrice de covariance $(Y_{jk})_{p\times n_2}$. Les variables sont soient toutes réelles soient complexes. Il est donc utile d'étudier les distributions asymptotiques des estimateurs de ces paramètres associés à la F-matrice. Dans cet article, nous établissons des théorèmes centraux limites pour les statistiques linéaires du spectre de la F-matrice dans la limite où p, n_1, n_2 tendent vers l'infini en restant de même ordre, et donnons des formules exactes pour leurs moyennes et covariances. De plus, l'hypothèse que les variables $(X_{jk})_{p\times n_1}$ et $(Y_{jk})_{p\times n_2}$ sont i.i.d. et la restriction que le quatrième moment est égal à 2 ou 3 comme dans Bai et Silverstein (*Ann. Probab.* **32** (2004) 553–605) sont affaiblies de la manière suivante; les coefficients $(X_{jk})_{p\times n_1}$ et $(Y_{jk})_{p\times n_2}$ sont indépendants mais non nécessairement équidistribués, pourvu qu'ils aient le même quatrième moment dans chaque tableau. Par conséquent, nous obtenons le théorème de la limite centrale pour les statistiques linéaires de la matrice beta qui est de la forme $(\mathbf{I} + d\cdot F \text{ matrix})^{-1}$, où d est une constante et $\mathbf{I}$ la matrice identité.

MSC: Primary 15A52; 60F05; secondary 62H10

Keywords: Linear spectral statistics; Central limit theorem; F-matrix; Beta matrix

1. Introduction

The central limit theorems (CLT) for linear spectral statistics (LSS) of large dimensional random matrices have a long history, and received considerable attention in recent years. They have important applications in various domains like number theory, high-dimensional multivariate statistics or wireless communication networks; see a survey by Johnstone [25]. A CLT is proposed by Jonsson [26] for $(\operatorname{tr}(A_n), \ldots, \operatorname{tr}(A_n^k))$ for a sequence of Wishart matrices (A_n), where k is a given fixed order of moments, and the dimension p of the matrices grows to infinity proportionally to the sample size n. Sinaĭ and Soshnikov [36,37] consider the same problem for the Wigner matrices where the order k can grow at the rate $o(n^{2/3})$ and where more general functions can be employed in place of the simple trace function. Johansson [24] establishes a CLT for a large class of random Hermitian matrices that includes the Gaussian unitary ensemble. Diaconis and Evans [13] obtain the CLT for the Haar matrices using the method of moments. Bai and Yao [7] establish the CLT for LSS of Wigner matrices using the Stieltjes transformation method. Other CLTs can be found in Girko [15] and Khorunzhy et al. [28]. The method of stochastic calculus is introduced in Cabanal-Duvillard [10]. Guionnet [16] obtains CLTs for non-commutative functionals of Gaussian large random matrices. Anderson and Zeitouni [1] establish CLTs for a band matrix model. Actually, there is a vast literature on CLTs for the eigenvalues of random matrices of various types, see e.g., Costin and Lebowtz [12], Boutet De Monvel et al. [9], Johansson [23], Keating and Snaith [27], Hughes et al. [20], Israelson [21], Soshnikov [38], Wieand [39], Dumitriu and Edelman [14], Mingo and Speicher [29], Ridelury and Silverstein [32], Hachem et al. [17,18], Rider and Virág [33], Chatterjee [11] and Jiang [22].

In this article, we consider CLTs for a specific matrix ensemble called *Fisher matrices*, or simply F-matrices. Let $\{X_{jk}, j,k=1,2,\dots\}$ and $\{Y_{jk}, j,k=1,2,\dots\}$ be either both real or both complex random variable arrays. For $p\ge 1$, $n_1\ge 1$ and $n_2\ge 1$, let

$$\mathbf{X}=(X_{jk}\colon 1\le j\le p, 1\le k\le n_1)=(\mathbf{X}_{\cdot 1},\dots,\mathbf{X}_{\cdot n_1}),$$
$$\mathbf{Y}=(Y_{jk}\colon 1\le j\le p, 1\le k\le n_2)=(\mathbf{Y}_{\cdot 1},\dots,\mathbf{Y}_{\cdot n_2})$$

be the upper-left $p\times n_1$ and $p\times n_2$ sections of the above arrays, with column vectors $(\mathbf{X}_{\cdot k})$ and $(\mathbf{Y}_{\cdot k})$, respectively. The reason for considering such sub-arrays is that $(\mathbf{X}_{\cdot 1},\dots,\mathbf{X}_{\cdot n_1})$ and $(\mathbf{Y}_{\cdot 1},\dots,\mathbf{Y}_{\cdot n_2})$ are two independent samples of p-dimensional observations of sizes n_1 and n_2, respectively. Note that in this statistical interpretation, it is common to assume that the p-dimensional vectors $(\mathbf{X}_{\cdot k})_{1\le k\le n_1}$ as well as the vectors $(\mathbf{Y}_{\cdot k})_{1\le k\le n_2}$ are identically distributed. However, the CLTs developed in this paper do not need this equal distribution property.

Let us introduce the following sample covariance matrices:

$$\mathbf{S}_1=\frac{1}{n_1}\sum_{k=1}^{n_1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*=\frac{1}{n_1}\mathbf{X}\mathbf{X}^*, \tag{1.1}$$

$$\mathbf{S}_2=\frac{1}{n_2}\sum_{k=1}^{n_2}\mathbf{Y}_{\cdot k}\mathbf{Y}_{\cdot k}^*=\frac{1}{n_2}\mathbf{Y}\mathbf{Y}^*, \tag{1.2}$$

where * stands for complex conjugate and transpose. These two matrices are both of size $p\times p$. Then the *Fisher matrix* (or F-matrix) is defined as

$$\mathbf{F}:=\mathbf{S}_1\mathbf{S}_2^{-1}. \tag{1.3}$$

When considering the asymptotic limits as $p\wedge n_1\wedge n_2\to\infty$, it will be necessary to ensure that almost surely, the matrix $\mathbf{S}_2$ is invertible for all large enough p, n_1 and n_2. In particular, we will assume that $n_2\ge p$ for large enough p and n_2.

Let us take a moment to explain the importance of the F-matrices in multivariate statistical analysis. Assume that in the two-sample problem introduced above, $\operatorname{cov}(\mathbf{X}_{\cdot k})=\mathbf{\Sigma}_1$ and $\operatorname{cov}(\mathbf{Y}_{\cdot k})=\mathbf{\Sigma}_2$. Clearly, $E\mathbf{S}_j=\mathbf{\Sigma}_j$, $j=1,2$. Following the pioneering work of R. Fisher, we can use the F-matrices $\mathbf{F}=\mathbf{S}_1\mathbf{S}_2^{-1}$ to test the hypothesis "$\mathbf{\Sigma}_1=\mathbf{\Sigma}_2$." Intuitively speaking, one would reject this equality hypothesis when $\mathbf{F}$ deviates significantly from the identity matrix. More precisely, assuming that the two samples are both Gaussian, the likelihood ratio statistic equals

$$W_{\mathbf{n}}=\sum_{i=1}^{p}\log\left(\frac{p}{n_1}+\frac{p}{n_2}\lambda_i^{\mathbf{F}}\right)-\frac{n_1}{n_1+n_2}\sum_{i=1}^{p}\log\bigl(\lambda_i^{\mathbf{F}}\bigr),$$

where $(\lambda_i^{\mathbf{F}})$ denote the eigenvalues of $\mathbf{F}$. The random variable $W_{\mathbf{n}}$ is a special instance of the so-called *linear spectral statistics* (LSS) of the random matrices $\mathbf{F}$, because it can be rewritten in the form

$$W_{\mathbf{n}} = p \int f_{\mathbf{n}}(x)\,\mathrm{d}F_{\mathbf{n}}(x), \tag{1.4}$$

where

$$f_{\mathbf{n}}(x) = \log\left(\frac{p}{n_1} + \frac{p}{n_2}x\right) - \frac{n_1}{n_1+n_2}\log(x) \tag{1.5}$$

and

$$F_{\mathbf{n}} = \frac{1}{p}\sum_{i=1}^{p} \delta_{\lambda_i^{\mathbf{F}}}$$

denotes the *empirical spectral distribution* (ESD) of $\mathbf{F}$.

The goal of this paper is to find when $p \wedge n_1 \wedge n_2 \to \infty$, $p/n_1 \to y_1 \in (0, +\infty)$ and $p/n_2 \to y_2 \in (0, 1)$, whether $r_{\mathbf{n}}(W_{\mathbf{n}} - \alpha_{\mathbf{n}})$ has a limiting distribution with some suitable centering and scaling parameters $r_{\mathbf{n}}$ and $\alpha_{\mathbf{n}}$. It is worth mentioning that $f_{\mathbf{n}}$ in $W_{\mathbf{n}}$ can be a general analytic function and not only the function defined in (1.5).

For earlier references on the F-matrices, we would like to mention that their spectral properties are studied in Yin et al. [40], Silverstein [34], Bai [2], Bai et al. [8], Pillai [30] and Pillai and Flury [31]. A closely related piece of work is that of Bai and Silverstein [5] and the work establishes a *CLT* for the LSS of a general sample covariance matrices of the form $\mathbf{S}_1\mathbf{T}_p$, where $(\mathbf{T}_p)$ is a sequence of non-negative definite Hermitian matrices. Explicit formula are provided for the mean and covariance parameters of the limiting Gaussian distributions. From one point of view, this result covers the LSS of the F-matrices as a special case because $\mathbf{S}_2^{-1}$ can be viewed as an instance of $\mathbf{T}_p$. However, in such an application of [5], the centering term α_n in $r_{\mathbf{n}}(W_{\mathbf{n}} - \alpha_{\mathbf{n}})$ depends on $\mathbf{S}_2^{-1}$, so it is a random variable. In the applications of the CLT, the normalizing center should be related to the population property, which should be non-random. For example, in the statistical inference, the normalizing center used in the test should be a functional of the null hypothesis, which is non-random. A non-random centering constant is also required for constructing confidence intervals for certain population parameters. Hence, their results will be of limited use in inference. Moreover, Bai and Silverstein [5] require X_{ij} to be i.i.d., $EX_{ij}^4 = 3$ when X_{ij} is real, and $EX_{ij}^2 = 0$, $E|X_{ij}|^4 = 2$ when X_{ij} is complex. In this paper, X_{ij} is required to be independent with finite fourth moments. Bai and Silverstein's [5] CLT is not directly applicable to this case. Finally, assuming that all the entries of $\mathbf{S}_1$ and $\mathbf{S}_2$ are Gaussian, one could use a result from Chatterjee [11] to obtain conditions ensuring a Gaussian limit for the LSS of the F-matrices. However, the method of Chatterjee [11] does not provide explicit formula for the mean and covariance parameters of the asymptotic distribution, hence will be of limited practical interest for applications.

The paper is organized as follows. In Section 2 we recall some useful background about the spectral theory of the F-matrices. Section 3 presents the main results of the paper about the CLTs for LSS of large dimensional F matrices and beta matrices. In Section 4, we present several applications of these CLTs. Section 5 gives some comments and conclusions. The proofs of the main results are then given in Section 6.

2. Limits of the spectral distribution of an F-matrix

Before formulating the CLT for the LSS of an F-matrix, we introduce some basic concepts and notations. Recall that given a real or complex-valued matrix $\mathbf{A}$ of size $p \times p$ with eigenvalues denoted by $\{\lambda_j^{\mathbf{A}}\}$, the distribution $\frac{1}{p}\sum_{i=1}^{p} \delta_{\lambda_j^{\mathbf{A}}}$ is called the empirical spectral distribution (ESD) of $\mathbf{A}$. The Stieltjes transform of a cumulative distribution function (c.d.f.) G on the real line is defined as

$$m_G(z) \equiv \int \frac{1}{\lambda - z}\,\mathrm{d}G(\lambda), \quad z \in \mathbb{C}^+ = \{z\colon z \in \mathbb{C}, \Im(z) > 0\}. \tag{2.1}$$

This definition has a natural extension to the lower-half of the complex space by letting

$$m_G(z) = \overline{m}_G(\bar{z}) \quad \text{for } z \in \mathbb{C}^- = \{z\colon z \in \mathbb{C}, \Im(z) < 0\}.$$

Throughout the paper, we will use the following assumptions.

Assumption [A]. *Either all real or all complex random variables* $\{X_{ki}, i, k = 1, 2, \ldots\}$ *and* $\{Y_{ki}, i, k = 1, 2, \ldots\}$ *are independent but not necessarily identically distributed, and with mean* 0 *and variance* 1.

Moreover, for any fixed $\eta > 0$ and when $n_1, n_2 \to \infty$,

$$\frac{1}{n_1 p} \sum_{j=1}^{p} \sum_{k=1}^{n_1} E|X_{jk}|^4 \mathbf{1}_{\{|X_{jk}| \geq \eta\sqrt{n_1}\}} \to 0, \qquad \frac{1}{n_2 p} \sum_{j=1}^{p} \sum_{k=1}^{n_2} E|Y_{jk}|^4 \mathbf{1}_{\{|Y_{jk}| \geq \eta\sqrt{n_2}\}} \to 0. \tag{2.2}$$

Assumption [B]. *The sample sizes* n_1, n_2 *and the dimension* p *grow to infinity in such a related way that*

$$y_{n_1} := p/n_1 \to y_1 \in (0, +\infty), \qquad y_{n_2} := p/n_2 \to y_2 \in (0, 1). \tag{2.3}$$

The assumption (2.2) is a standard one of Lindeberg type, which is necessary as the columns vectors $\{\mathbf{X}_{\cdot k}\}$ and $\{\mathbf{Y}_{\cdot k}\}$ are not necessarily identically distributed. Moreover, it allows a suitable truncation of these random variables without modifying limiting results.

Let

$$\mathbf{n} = (n_1, n_2), \qquad \mathbf{y_n} = (y_{n_1}, y_{n_2}), \qquad \mathbf{y} = (y_1, y_2).$$

Subsequently, the limiting scheme (2.3) will be simply refereed as $\mathbf{n} \to \infty$.

It is well-known (e.g., [6], p. 79) that under Assumptions [A]–[B], when $\mathbf{n} \to \infty$, almost surely the random ESD $f_\mathbf{n}$ of the F-matrix $\mathbf{S}_1\mathbf{S}_2^{-1}$ converges to a distribution

$$F_\mathbf{y}(\mathrm{d}x) = g_\mathbf{y}(x)\mathbf{1}_{[a,b]}(x)\,\mathrm{d}x + (1 - 1/y_1)\mathbf{1}_{\{y_1>1\}}\delta_0(\mathrm{d}x), \tag{2.4}$$

where

$$\begin{aligned} &h = \sqrt{y_1 + y_2 - y_1 y_2}, \qquad a = \frac{(1-h)^2}{(1-y_2)^2}, \qquad b = \frac{(1+h)^2}{(1-y_2)^2}, \\ &g_\mathbf{y}(x) = \frac{(1-y_2)}{2\pi x(y_1 + y_2 x)}\sqrt{(b-x)(x-a)}, \quad a < x < b. \end{aligned} \tag{2.5}$$

In other words, this so-called *limiting spectral distribution* (LSD) has an absolutely continuous component on $[a, b]$ and a point mass at the origin if $y_1 > 1$. Furthermore, under Assumptions [A]–[B], the Stieltjes transform $m_\mathbf{y}(z)$ of $F_\mathbf{y}$ equals

$$m_\mathbf{y}(z) = \frac{1}{zy_1} - \frac{1}{z} - \frac{y_1[z(1-y_2) + 1 - y_1] + 2zy_2 + y_1\sqrt{[(1-y_1) + z(1-y_2)]^2 - 4z}}{2zy_1(y_1 + zy_2)}, \quad z \in \mathbb{C}^+ \tag{2.6}$$

(see [6], p. 79). Let us denote the Stieltjes transform of the ESD $f_\mathbf{n}$ by $m_\mathbf{n}(z)$ which will converge almost surely and pointwisely to $m_\mathbf{y}(z)$.

For technical reasons, we will also need two Stieltjes transforms of the ESD of the random matrix $\mathbf{X}^*\mathbf{S}_2^{-1}\mathbf{X}$ and its limit when $\mathbf{n} \to \infty$. Their Stieltjes transforms will be denoted by $\underline{m}_\mathbf{n}(z)$ and $\underline{m}_\mathbf{y}(z)$. Because the spectra of $\mathbf{X}^*\mathbf{S}_2^{-1}\mathbf{X}$ is different from that of $\mathbf{S}_1\mathbf{S}_2^{-1}$ by $|n_1 - p|$ zeros, then from the definition (2.1) of the Stieltjes transform we have

$$\underline{m}_\mathbf{n}(z) = -\frac{1 - y_{n_1}}{z} + y_{n_1} m_\mathbf{n}(z), \tag{2.7}$$

$$\underline{m}_\mathbf{y}(z) = -\frac{1 - y_1}{z} + y_1 m_\mathbf{y}(z). \tag{2.8}$$

So from (2.6) and (2.8) we obtain

$$\underline{m}_{\mathbf{y}}(z) = -\frac{y_1[z(1-y_2)+1-y_1]+2zy_2+y_1\sqrt{[(1-y_1)+z(1-y_2)]^2-4z}}{2z(y_1+zy_2)}, \quad z \in \mathbb{C}^+. \tag{2.9}$$

In the i.i.d. case, it is proved by [35] that $\underline{m}_{\mathbf{y}}(z)$ satisfies the following important equation

$$z = -\frac{1}{\underline{m}_{\mathbf{y}}(z)} + \int \frac{y_1\,\mathrm{d}F_{y_2}(x)}{x+\underline{m}_{\mathbf{y}}(z)} \quad \text{or} \quad z = -\frac{1}{\underline{m}_{\mathbf{y_n}}(z)} + \int \frac{y_{n_1}\,\mathrm{d}F_{y_{n_2}}(x)}{x+\underline{m}_{\mathbf{y_n}}(z)}, \tag{2.10}$$

where $F_{y_2}(x)$ denotes the LSD of $\mathbf{S}_2$. Recall that in the proof of (2.10), after the truncation, all arguments remain true provided that the second moments of the truncated variables tend to 1 uniformly, which is a straightforward consequence of Assumption [A]. Since $F_{n_2} \to F_{y_2}$, *a.s.* under Assumption [A], the equation remains true under Assumption [A].

In the remainder of the paper, for brevity, $m_{\mathbf{y}}(z)$ and $\underline{m}_{\mathbf{y}}(z)$ will be simplified as $m(z)$ and $\underline{m}(z)$ or even more simpler as m and $\underline{m}$, if no confusion would arise.

With a slight abuse of notation, let $m_{n_1}(z)$ denote the Stieltjes transform of the ESD $F_{n_1}(x)$ of $\mathbf{S}_1$ and let $m_{y_1}(z)$ denote the Stieltjes transform of the LSD $F_{y_1}(x)$ with density $f_{y_1(x)}$. Note that F_{y_1} is simply the well-known Marčenko–Pastur distribution of index y_1. And similar to Eqs (2.7)–(2.8), the following Stieltjes transforms for some distributions

$$\underline{m}_{n_1}(z) = -\frac{1-y_{n_1}}{z} + y_{n_1} m_{n_1}(z), \tag{2.11}$$

$$\underline{m}_{y_1}(z) = -\frac{1-y_1}{z} + y_1 m_{y_1}(z) \tag{2.12}$$

are needed. It follows from all these definitions that

$$f_{y_1} = g_{(y_1,0)}, \qquad m_{y_1} = m_{(y_1,0)}, \qquad \underline{m}_{y_1} = \underline{m}_{(y_1,0)}.$$

Replacing $\mathbf{S}_1$ by $\mathbf{S}_2$, we get a similar ESD F_{n_2} and its LSD F_{y_2} with density f_{y_2} with respect to the Stieltjes transforms m_{n_2} and m_{y_2}. Definitions of $\underline{m}_{n_2}$ and $\underline{m}_{y_2}$ are similar to (2.11) and (2.12).

Furthermore, let $H_{n_2}(x)$ and $H_{y_2}(x)$ denote the ESD and LSD of $\mathbf{S}_2^{-1}$, respectively. Clearly if λ is a positive eigenvalue of $\mathbf{S}_2^{-1}$, then $1/\lambda$ is a positive eigenvalue of $\mathbf{S}_2$. Therefore, we have for all $x > 0$,

$$H_{n_2}(x_-) = 1 - F_{n_2}(1/x), \qquad H_{y_2}(x_-) = 1 - F_{y_2}(1/x). \tag{2.13}$$

3. Main results

As explained in the introduction, Bai and Silverstein [5] establish the CLT for the LSS of a general sample covariance matrix of the form $\mathbf{B}_p = \mathbf{S}_1\mathbf{T}_p$, where $\mathbf{S}_1$ is a $p \times p$ sample covariance matrix generated by i.i.d. entries and $\mathbf{T}_p$ is a non-negative definite matrix in two cases: either both X_{11} and $\mathbf{T}_p$ are real with $E|X_{11}|^4 = 3$ (referred to as the real case) or both X_{11} and $\mathbf{T}_p$ are complex with $EX_{11}^2 = 0$ and $E|X_{11}|^4 = 2$ (referred to as the complex case). This paper will establish the CLTs for the LSS of F-matrices with two additional improvements: first, the initial arrays of random vectors $\{\mathbf{X}_{\cdot k}\}$ and $\{\mathbf{Y}_{\cdot k}\}$ are not necessarily identically distributed; second, a common value will be assumed for the forth moment of all the variables $\{X_{jk}\}$ as well as for the variables $\{Y_{jk}\}$, but these two common values can be arbitrary instead of fixed constants used as in [5].

For a given function f and the associated LSS $\int f(x)\,\mathrm{d}f_{\mathbf{n}}(x)$, we will consider the following centered and scaled variables

$$p\int f(x)\,\mathrm{d}\big[f_{\mathbf{n}}(x) - F_{\mathbf{y_n}}(x)\big] =: \int f(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x) \tag{3.1}$$

with

$$\widetilde{G}_n = p\big[f_{\mathbf{n}}(x) - F_{\mathbf{y_n}}\big].$$

Due to the exact separation theorem (see [4]), for large enough n_j and p, the discrete part, i.e., the mass at the origin of $f_{\mathbf{n}}$ will coincide exactly with that of $F_{\mathbf{y_n}}$, so that in (3.1) we can reduce the integral to the continuous part only. Therefore, we have

$$\begin{aligned}\int f(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x) &= \sum_{j=1}^{p} f\big(\lambda_j^{\mathbf{F}}\big) - p\int f(x)\,\mathrm{d}F_{\mathbf{y_n}}(x)\\ &= \sum_{j=1}^{p} f\big(\lambda_j^{\mathbf{F}}\big)\mathbf{1}_{(\lambda_j^{\mathbf{F}}>0)} - p\int_{a_p}^{b_p} f(x) g_{\mathbf{y_n}}(x)\,\mathrm{d}x,\end{aligned} \tag{3.2}$$

where $g_{\mathbf{y_n}}$ is the density function defined in (2.4) with the substitution of $\mathbf{y_n} = (y_{n_1}, y_{n_2})$ for $\mathbf{y} = (y_1, y_2)$ and the associated constants (h_p, a_p, b_p) for (h, a, b), namely

$$h_p = \sqrt{y_{n_1} + y_{n_2} - y_{n_1}y_{n_2}}, \qquad a_p = \frac{(1-h_p)^2}{(1-y_{n_2})^2}, \qquad b_p = \frac{(1+h_p)^2}{(1-y_{n_2})^2}.$$

The main results of the paper are the following. Let

$$m_0(z) = \underline{m}_{y_2}\big(-\underline{m}(z)\big), \quad z \in \mathbb{C}^+, \tag{3.3}$$

and

$$\kappa = \begin{cases} 1, & \text{if all the } \mathbf{X}\text{- and } \mathbf{Y}\text{-variables are complex,}\\ 2, & \text{if all the } \mathbf{X}\text{- and } \mathbf{Y}\text{-variables are real.}\end{cases}$$

Theorem 3.1. *Assume that*

1. *Assumptions* [A]–[B] *are satisfied.*
2. *For all* j, k, $E|X_{jk}|^4 = 1+\kappa$, $E|Y_{jk}|^4 = 1+\kappa$. *If both* $\mathbf{X}$ *and* $\mathbf{Y}$ *are complex valued, then* $EX_{jk}^2 = EY_{jk}^2 = 0$.

Let $f_1, \ldots, f_s$ (*s is a fixed integer*) *be functions analytic in an open region in the complex plane containing the interval* $[a, b]$ *which is the support of the continuous part of the LSD* $F_{\mathbf{y}}$ *defined in* (2.4).

Then, as $\mathbf{n} \to \infty$, *the random vector*

$$\left[\int f_j(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)\right], \quad 1 \le j \le s,$$

converges weakly to a Gaussian vector $(X_{f_1}, \ldots, X_{f_s})$ *with means*

$$\begin{aligned}EX_{f_i} &= \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\left(\frac{(1-y_2)m_0^2(z) + 2m_0(z) + 1 - y_1}{(1-y_2)m_0^2(z) + 2m_0(z) + 1}\right)\\ &\quad + \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\big(1 - y_2 m_0^2(z)\big(1+m_0(z)\big)^{-2}\big)\end{aligned} \tag{3.4}$$

and covariance functions

$$\operatorname{cov}(X_{f_i}, X_{f_j}) = -\frac{\kappa}{4\pi^2}\oint\oint \frac{f_i(z_1) f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1) - m_0(z_2))^2}. \tag{3.5}$$

Here all the contour integrals can be evaluated on any contour enclosing $[a, b]$.

Remark 3.1. *An interesting example can be found on p.* 513 *of* [6] *that different distributions have identical moments for all orders. Moreover, as demonstrated by the examples developed later, to get more explicit expressions for the asymptotic means and covariances, it is necessary to evaluate the involved contour integrals on a contour approaching the interval* $[a,b]$. *If* $y_1 \le 1$, *then* $[a,b]$ *is exactly the limiting support of the LSD* $F_{\mathbf{y}}$, *and we can choose the contours very close to* $[a,b]$. *If* $y_1 > 1$, $F_{\mathbf{y}}$ *has a positive mass at the origin and a priori, the contours should enclose the whole support* $\{0\} \cup [a,b]$ *of* $F_{\mathbf{y}}$. *However, as explained earlier, due to the exact separation, we can restrict the integrals* $\int f_j(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)$ *to positive eigenvalues of the* F*-matrix* $\mathbf{F}$ *and the continuous part of the distributions* $F_{\mathbf{y_n}}$. *Consequently, in Theorem* 3.1, *we can choose contours close to* $[a,b]$.

Generally, it is difficult to compute the asymptotic means and covariances by using the expressions given in Theorem 3.1. In the following corollary, we convert the integrals into another form of contour integrations for computing the means and covariance functions.

Corollary 3.1. *Under the assumptions of Theorem* 3.1, *the asymptotic means and covariances of the limiting random vector can be computed as follows*

$$EX_{f_i} = \lim_{r\downarrow 1} \frac{\kappa-1}{4\pi \mathrm{i}} \oint_{|\xi|=1} f_i\left(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\right)\left[\frac{1}{\xi - r^{-1}} + \frac{1}{\xi + r^{-1}} - \frac{2}{\xi + y_2/h}\right]\mathrm{d}\xi$$

and

$$\begin{aligned}&\operatorname{cov}(X_{f_i}, X_{f_j})\\&\quad = \lim_{r\downarrow 1} -\frac{\kappa}{4\pi^2}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1} \frac{f_i((1+h^2+2h\Re(\xi_1))/(1-y_2)^2) f_j((1+h^2+2h\Re(\xi_2))/(1-y_2)^2)}{(\xi_1 - r\xi_2)^2}\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2,\end{aligned}$$

where $\Re(\cdot)$ *denotes the real part of a complex number and* $r \downarrow 1$ *means that "*r *approaches* 1 *from above."*

Next, we extend Theorem 3.1 to the case where the common values of the 4th moments of $\mathbf{X}$ and $\mathbf{Y}$ are arbitrary.

Theorem 3.2. *Assume that*

1. *Assumptions* [A]–[B] *are satisfied.*
2. *For all* j,k, $E|X_{jk}|^4 = \beta_x + 1 + \kappa$, $E|Y_{jk}|^4 = \beta_y + 1 + \kappa$, *and if both* $\mathbf{X}$ *and* $\mathbf{Y}$ *are complex valued, then* $EX_{jk}^2 = EY_{jk}^2 = 0$.

Let $f_1, \ldots, f_s$ *(*s *is a fixed integer) be functions analytic in an open region in the complex plane containing the interval* $[a,b]$, *which is the support of the continuous part of the LSD* $F_{\mathbf{y}}$ *defined in* (2.4).

Then, as $\mathbf{n} \to \infty$, *the random vector*

$$\left[\int f_j(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)\right], \quad 1 \le j \le s,$$

converges weakly to a Gaussian vector $(X_{f_1}, \ldots, X_{f_k})$ *with means*

$$\begin{aligned}EX_{f_i} &= \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\left(\frac{(1-y_2)m_0^2(z)+2m_0(z)+1-y_1}{(1-y_2)m_0^2(z)+2m_0(z)+1}\right)\\&\quad + \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)\\&\quad + \frac{\beta_x\cdot y_1}{2\pi\mathrm{i}}\oint f_i(z)\bigl(m_0(z)+1\bigr)^{-3}\,\mathrm{d}m_0(z)\\&\quad + \frac{\beta_y}{4\pi\mathrm{i}}\oint f_i(z)\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)\,\mathrm{d}\log\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr) \qquad (3.6)\end{aligned}$$

and covariance functions

$$\operatorname{cov}(X_{f_i}, X_{f_j}) = -\frac{\kappa}{4\pi^2} \oint \oint \frac{f_i(z_1) f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1) - m_0(z_2))^2} - \frac{(\beta_x y_1 + \beta_y y_2)}{4\pi^2} \oint \oint \frac{f_i(z_1) f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)+1)^2 (m_0(z_2)+1)^2}. \tag{3.7}$$

Similar to Corollary 3.1, the following corollary helps the evaluation of the asymptotic means and covariance functions in Theorem 3.2.

Corollary 3.2. *Under the assumptions of Theorem 3.2, the asymptotic means and covariance functions of the limiting Gaussian vector can be written as*

$$\begin{aligned} EX_{f_i} &= \lim_{r\downarrow 1} \frac{\kappa - 1}{4\pi \mathrm{i}} \oint_{|\xi|=1} f_i\left(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\right)\left[\frac{1}{\xi - r^{-1}} + \frac{1}{\xi + r^{-1}} - \frac{2}{\xi + y_2/h}\right]\mathrm{d}\xi \\ &\quad + \frac{\beta_x \cdot y_1(1-y_2)^2}{2\pi \mathrm{i}\cdot h^2} \oint_{|\xi|=1} f_i\left(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\right)\frac{1}{(\xi + y_2/h)^3}\,\mathrm{d}\xi \\ &\quad + \frac{\beta_y \cdot (1-y_2)}{4\pi \mathrm{i}} \oint_{|\xi|=1} f_i\left(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\right)\frac{\xi^2 - y_2/h^2}{(\xi + y_2/h)^2} \\ &\quad \times \left[\frac{1}{\xi - \sqrt{y_2}/h} + \frac{1}{\xi + \sqrt{y_2}/h} - \frac{2}{y_2/h}\right]\mathrm{d}\xi \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} &\operatorname{cov}(X_{f_i}, X_{f_j}) \\ &\quad = -\lim_{r\downarrow 1} \frac{\kappa}{4\pi^2} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} \frac{f_i((1+h^2+2h\Re(\xi_1))/(1-y_2)^2) f_j((1+h^2+2h\Re(\xi_2))/(1-y_2)^2)}{(\xi_1 - r\xi_2)^2}\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2 \\ &\qquad - \frac{(\beta_x y_1 + \beta_y y_2)(1-y_2)^2}{4\pi^2 h^2} \oint_{|\xi_1|=1} \frac{f_i((1+h^2+2h\Re(\xi_1))/(1-y_2)^2)}{(\xi_1 + y_2/h)^2}\,\mathrm{d}\xi_1 \\ &\qquad \times \oint_{|\xi_2|=1} \frac{f_j((1+h^2+2h\Re(\xi_2))/(1-y_2)^2)}{(\xi_2 + y_2/h)^2}\,\mathrm{d}\xi_2. \end{aligned} \tag{3.9}$$

4. Applications

4.1. *Beta-matrices*

In multivariate statistical analysis, many commonly used matrices are beta matrices. Because the beta matrix is a functional of the F matrix, we can use the result on the F matrix to get asymptotic results for the beta matrix. Because of its common use, we give the CLT of the beta matrix here so that the result can be directly used in later applications. We first give an application of our CLTs to the set of so-called *beta-matrices*. A beta-matrix takes the form

$$\boldsymbol{\beta}_{\mathbf{n}} = \mathbf{S}_2(\mathbf{S}_2 + d\cdot\mathbf{S}_1)^{-1} = \left(\mathbf{I} + d\cdot\mathbf{S}_1\mathbf{S}_2^{-1}\right)^{-1}, \tag{4.1}$$

where d is a positive constant. This is a matrix-valued functional of a F-matrix. Let

$$\widehat{G}_{\mathbf{n}}(x) = p\left(F_{0,\mathbf{n}}(x) - F_{0,\mathbf{y_n}}(x)\right),$$

where $F_{0,\mathbf{n}}(x)$ and $F_{0,\mathbf{y_n}}(x)$ are the ESD and LSD of $\boldsymbol{\beta}_{\mathbf{n}}$, respectively.

Theorem 4.1. *Under the same assumptions as in Theorem 3.2, the vector of LSS*

$$\left(\int f_1(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x),\ldots,\int f_k(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x)\right)$$

of the beta matrix

$$\boldsymbol{\beta}_{\mathbf{n}}=\mathbf{S}_2(\mathbf{S}_2+d\cdot\mathbf{S}_1)^{-1}$$

with a positive constant d, verifies

$$\int f_i(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x)=\int f_i\left(\frac{1}{dx+1}\right)\mathrm{d}\widetilde{G}_{\mathbf{n}}(x),$$

where $\int f_i(\frac{1}{dx+1})\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)$ is as in (3.2). Moreover, as $\mathbf{n}\to\infty$, this vector converges weakly to a Gaussian vector $(X_{f_1},\ldots,X_{f_k})$ whose means and covariances are the same as those in Theorem 3.2 except that $f_i(x)$ and $f_j(x)$ are replaced by $f_i(\frac{1}{d\cdot x+1})$ and $f_j(\frac{1}{d\cdot x+1})$, respectively.

4.2. *Popular LSS of F-matrices*

This section aims to illustrate how one can use Theorems 3.1 and 3.2 for some LSS popular in multivariate statistical analysis. In particular, the goal is to compute explicitly the asymptotic means and covariance functions of the limiting Gaussian distribution. We only restrict our attention to the real variables. Three examples will be given, but we only detail the computations for the first one, because those of the other examples are very similar.

Example 4.1. *If $f=\log(a+bx)$, $f'=\log(a'+b'x)$, $a,a',b,b'>0$, then*

$$EX_f=\frac{1}{2}\log\left(\frac{(c^2-d^2)h^2}{(ch-y_2d)^2}\right)-\frac{\beta_x y_1(1-y_2)^2d^2}{2(ch-dy_2)^2}+\frac{\beta_y(1-y_2)}{2}\left[\frac{2dy_2}{ch-dy_2}+\frac{d^2(y_2^2-y_2)}{(ch-dy_2)^2}\right]$$

and

$$\operatorname{cov}(X_f,X_{f'})=2\log\left(\frac{cc'}{cc'-dd'}\right)+\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2dd'}{(ch-dy_2)(c'h-d'y_2)},$$

where $c>d>0$, $c'>d'>0$ satisfying $c^2+d^2=\frac{a(1-y_2)^2+b(1+h^2)}{(1-y_2)^2}$, $(c')^2+(d')^2=\frac{a'(1-y_2)^2+b'(1+h^2)}{(1-y_2)^2}$, $cd=\frac{bh}{(1-y_2)^2}$ and $c'd'=\frac{b'h}{(1-y_2)^2}$.

Proof. We have $f(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2})=\log(|c+d\xi|^2)$ and $f'(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2})=\log(|c'+d'\xi|^2)$.

In the formula

$$\begin{aligned}E(X_f)&=\lim_{r\downarrow 1}\frac{1}{4\pi\mathrm{i}}\oint_{|\xi|=1}f\left(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\right)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right)\mathrm{d}\xi\\&=\lim_{r\downarrow 1}\frac{1}{4\pi\mathrm{i}}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right)\mathrm{d}\xi\\&\quad+\frac{\beta_x y_1(1-y_2)^2}{2\pi\mathrm{i}h^2}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\frac{1}{(\xi+y_2/h)^3}\,\mathrm{d}\xi\\&\quad+\frac{\beta_y(1-y_2)}{4\pi\mathrm{i}}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\frac{\xi^2-y_2/h^2}{(\xi+y_2/h)^2}\left[\frac{1}{\xi-\sqrt{y_2}/h}+\frac{1}{\xi+\sqrt{y_2}/h}-\frac{2}{\xi+y_2/h}\right]\mathrm{d}\xi\end{aligned}$$

we have

$$\begin{aligned}
&\lim_{r\downarrow 1}\frac{1}{4\pi i}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right)d\xi\\
&=\lim_{r\downarrow 1}\frac{1}{4\pi i}\oint_{|\xi|=1}\bigl(\log(c+d\xi)+\log\bigl((c+d\xi^{-1})\bigr)\bigr)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right)d\xi\\
&=\lim_{r\downarrow 1}\frac{1}{4\pi i}\oint_{|\xi|=1}\log(c+d\xi)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right)d\xi\\
&\quad+\lim_{r\downarrow 1}\frac{1}{4\pi i}\oint_{|\xi|=1}\log(c+d\xi)\left(\frac{1}{r\xi^{-1}+1}+\frac{1}{r\xi^{-1}-1}-\frac{2}{\xi^{-1}+h^{-1}y_2}\right)\xi^{-2}\,d\xi\quad (\text{making } \xi^{-1}\to\xi)\\
&=\lim_{r\downarrow 1}\frac{1}{4\pi i}\int_{|\xi|=1}\log(c+d\xi)\left(\frac{1}{r\xi+1}+\frac{1}{r\xi-1}-\frac{2}{\xi+h^{-1}y_2}\right.\\
&\quad\left.+\frac{1}{\xi(r+\xi)}+\frac{1}{\xi(r-\xi)}-\frac{2}{\xi(1+h^{-1}y_2\xi)}\right)d\xi\\
&=\frac{1}{4}\bigl(\log\bigl[\bigl(c^2-d^2\bigr)^2\bigr]-2\log\bigl[\bigl(c-y_2dh^{-1}\bigr)^2\bigr]\bigr)=\frac{1}{2}\log\left(\frac{(c^2-d^2)h^2}{(ch-y_2d)^2}\right),
\end{aligned}$$

$$\begin{aligned}
&\frac{\beta_x y_1(1-y_2)^2}{2\pi i h^2}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\frac{1}{(\xi+y_2/h)^3}\,d\xi\\
&=\frac{\beta_x y_1(1-y_2)^2}{2\pi i h^2}\oint_{|\xi|=1}\bigl(\log(c+d\xi)+\log\bigl(c+d\xi^{-1}\bigr)\bigr)\frac{1}{(\xi+y_2/h)^3}\,d\xi\\
&=\frac{\beta_x y_1(1-y_2)^2}{2\pi i h^2}\oint_{|\xi|=1}\log(c+d\xi)\left[\frac{1}{(\xi+y_2/h)^3}+\frac{\xi}{(1+y_2\xi/h)^3}\right]d\xi\\
&=\frac{\beta_x y_1(1-y_2)^2}{4\pi i h^2}\frac{2\pi i}{2}\frac{-2d^2}{(c-dy_2/h)^2}=\frac{-\beta_x y_1(1-y_2)^2d^2}{2(ch-dy_2)^2}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\beta_y(1-y_2)}{4\pi i}\oint_{|\xi|=1}\log\bigl(|c+d\xi|^2\bigr)\frac{\xi^2-y_2/h^2}{(\xi+y_2/h)^2}\left[\frac{1}{\xi-\sqrt{y_2}/h}+\frac{1}{\xi+\sqrt{y_2}/h}-\frac{2}{\xi+y_2/h}\right]d\xi\\
&=\frac{\beta_y(1-y_2)}{4\pi i}\oint_{|\xi|=1}\bigl(\log(c+d\xi)+\log\bigl(c+d\xi^{-1}\bigr)\bigr)\frac{\xi^2-y_2/h^2}{(\xi+y_2/h)^2}\\
&\quad\times\left[\frac{1}{\xi-\sqrt{y_2}/h}+\frac{1}{\xi+\sqrt{y_2}/h}-\frac{2}{\xi+y_2/h}\right]d\xi\\
&=\frac{\beta_y(1-y_2)}{4\pi i}\oint_{|\xi|=1}\log(c+d\xi)\left\{\frac{\xi^2-y_2/h^2}{(\xi+y_2/h)^2}\left[\frac{1}{\xi-\sqrt{y_2}/h}+\frac{1}{\xi+\sqrt{y_2}/h}-\frac{2}{\xi+y_2/h}\right]d\xi\right.\\
&\quad\left.+\frac{1-y_2\xi^2/h^2}{(1+y_2\xi/h)^2}\left[\frac{1}{\xi(1-\sqrt{y_2}\xi/h)}+\frac{1}{\xi(1+\sqrt{y_2}\xi/h)}-\frac{2}{\xi(1+y_2\xi/h)}\right]d\xi\right\}\\
&=\frac{\beta_y(1-y_2)}{4}\left[\frac{2d((\sqrt{y_2}-y_2)/h)}{c-dy_2/h}+\frac{2d((-\sqrt{y_2}-y_2)/h)}{c-dy_2/h}+\frac{8dy_2/h}{c-dy_2/h}+\frac{2d^2((y_2^2-y_2)/h^2)}{(c-dy_2/h)^2}\right]\\
&=\frac{\beta_y(1-y_2)}{2}\left[\frac{2dy_2}{ch-dy_2}+\frac{d^2(y_2^2-y_2)}{(ch-dy_2)^2}\right].
\end{aligned}$$

So we obtain

$$EX_f = \frac{1}{2}\log\left(\frac{(c^2-d^2)h^2}{(ch-y_2d)^2}\right) - \frac{\beta_x y_1(1-y_2)^2 d^2}{2(ch-dy_2)^2} + \frac{\beta_y(1-y_2)}{2}\left[\frac{2dy_2}{ch-dy_2} + \frac{d^2(y_2^2-y_2)}{(ch-dy_2)^2}\right].$$

Furthermore, in the formula

$$\begin{aligned}\operatorname{cov}(X_f, X_{f'}) = &-\frac{1}{2\pi^2}\lim_{r\downarrow 1}\oint\oint_{|\xi_1|=|\xi_2|=1}\log\big(|c+d\xi_1|^2\big)\log\big(\left|c'+d'\xi_2\right|^2\big)\frac{\mathrm{d}\xi_1\,\mathrm{d}\xi_2}{(\xi_1-r\xi_2)^2}\\ &-\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2}{4\pi^2h^2}\oint_{|\xi_1|=1}\frac{\log(|c+d\xi_1|^2)}{(\xi_1+y_2/h)^2}\,\mathrm{d}\xi_1\oint_{|\xi_2|=1}\frac{\log(|c'+d'\xi_2|^2)}{(\xi_2+y_2/h)^2}\,\mathrm{d}\xi_2\end{aligned}$$

we have

$$\begin{aligned}&\lim_{r\downarrow 1}-\frac{1}{2\pi^2}\oint\oint_{|\xi_1|=|\xi_2|=1}\log\big(|c+d\xi_1|^2\big)\log\big(\left|c'+d'\xi_2\right|^2\big)\frac{\mathrm{d}\xi_1\,\mathrm{d}\xi_2}{(\xi_1-r\xi_2)^2}\\ &\quad=\lim_{r\downarrow 1}-\frac{1}{2\pi^2}\oint_{|\xi_1|=1}\log\big(|c+d\xi_1|^2\big)\left[\oint_{|\xi_2|=1}\big(\log\big(c'+d'\xi_2\big)+\log\big(c'+d'\xi_2^{-1}\big)\big)\frac{\mathrm{d}\xi_2}{(\xi_1-r\xi_2)^2}\right]\mathrm{d}\xi_1\\ &\quad=\lim_{r\downarrow 1}-\frac{1}{2\pi^2}\oint_{|\xi_1|=1}\log\big(|c+d\xi_1|^2\big)\left[\oint_{|\xi_2|=1}\log\big(c'+d'\xi_2\big)\left(\frac{\mathrm{d}\xi_2}{(\xi_1-r\xi_2)^2}+\frac{\mathrm{d}\xi_2}{(\xi_1\xi_2-r)^2}\right)\right]\mathrm{d}\xi_1\\ &\quad=-\lim_{r\downarrow 1}\frac{i}{\pi}\oint_{|\xi_1|=1}\log\big(|c+d\xi_1|^2\big)\frac{d'}{r^2c'+d'r\xi_1}\,\mathrm{d}\xi_1\quad\text{(the second term is analytic)}\\ &\quad=-\frac{\mathrm{i}}{\pi}\oint_{|\xi_1|=1}\big(\log(c+d\xi_1)+\log\big(c+d\xi_1^{-1}\big)\big)\frac{d'}{c'+d'\xi_1}\,\mathrm{d}\xi_1\\ &\quad=-\frac{\mathrm{i}}{\pi}\oint_{|\xi_1|=1}\log(c+d\xi_1)\left(\frac{d'}{c'+d'\xi_1}+\frac{d'}{\xi_1(c'\xi_1+d')}\right)\mathrm{d}\xi_1\\ &\quad=2\log\left(\frac{cc'}{cc'-dd'}\right)\end{aligned}$$

and

$$\oint_{|\xi|=1}\frac{\log(|c+d\xi|^2)}{(\xi+y_2/h)^2}\,\mathrm{d}\xi=\oint_{|\xi|=1}\left(\frac{\log(c+d\xi)}{(\xi+y_2/h)^2}+\frac{\log(c+d\xi)}{(\xi^{-1}+y_2/h)^2}\xi^{-2}\right)\mathrm{d}\xi=\frac{2\pi hd}{ch-dy_2}\mathrm{i}.$$

Therefore,

$$\begin{aligned}&-\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2}{4\pi^2h^2}\oint_{|\xi_1|=1}\frac{\log(|c+d\xi_1|^2)}{(\xi_1+y_2/h)^2}\,\mathrm{d}\xi_1\oint_{|\xi_2|=1}\frac{\log(|c'+d'\xi_2|^2)}{(\xi_2+y_2/(hr_2))^2}\,\mathrm{d}\xi_2\\ &\quad=\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2}{4\pi^2h^2}\frac{4\pi^2h^2dd'}{(ch-dy_2)(c'h-d'y_2)}\\ &\quad=\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2dd'}{(ch-dy_2)(c'h-d'y_2)}.\end{aligned}$$

So we obtain

$$\operatorname{cov}(X_f, X_{f'}) = 2\log\left(\frac{cc'}{cc'-dd'}\right)+\frac{(\beta_x y_1+\beta_y y_2)(1-y_2)^2dd'}{(ch-dy_2)(c'h-d'y_2)}.$$

□

Example 4.2. *For any positive integers $k \geq l \geq 1$ and $f_k(x) = x^k$ and $f_l(x) = x^l$, we have*

$$
\begin{aligned}
E(X_{f_k}) &= \lim_{r\downarrow 1} \frac{1}{4\pi \mathrm{i}} \oint_{|\xi|=1} \frac{(1+h\xi)^k(\xi+h)^k}{(1-y_2)^{2k}\cdot \xi^k}\left(\frac{1}{\xi+r^{-1}} + \frac{1}{\xi - r^{-1}} - \frac{2}{\xi + y_2/h}\right) \mathrm{d}\xi \\
&\quad + \frac{\beta_x \cdot y_1(1-y_2)^2}{2\pi \mathrm{i}\cdot h^2} \oint_{|\xi|=1} \frac{(1+h\xi)^k(\xi+h)^k}{(1-y_2)^{2k}\cdot \xi^k} \frac{1}{(\xi+y_2/h)^3} \mathrm{d}\xi \\
&\quad + \frac{\beta_y \cdot (1-y_2)}{4\pi \mathrm{i}} \oint_{|\xi|=1} \frac{(1+h\xi)^k(\xi+h)^k}{(1-y_2)^{2k}\cdot \xi^k} \frac{\xi^2 - y_2/h^2}{(\xi+y_2/h)^2} \\
&\quad \times \left[\frac{1}{\xi - \sqrt{y_2}/h} + \frac{1}{\xi + \sqrt{y_2}/h} - \frac{2}{\xi + y_2/h}\right] \mathrm{d}\xi \\
&= \frac{1}{2(1-y_2)^{2k}} \Bigg[(1-h)^{2k} + (1+h)^{2k} - 2(1-y_2)^k \left(1 - \frac{h^2}{y_2}\right)^k \\
&\quad + \sum_{i_1+i_2+i_3=k-1} \frac{k\cdot k! i_3!((-1)^{i_3} + (-1)^{2i_3+1})}{(k-i_1)!(k-i_2)!} h^{k+i_1-i_2} \\
&\quad + \sum_{i_1+i_2+i_3=k-1} \frac{2k\cdot k! i_3! \cdot h^{k+i_1-i_2}}{(k-i_1)!(k-i_2)!} \left(-\frac{h}{y_2}\right)^{i_3+1}\Bigg] \\
&\quad + \frac{\beta_x \cdot y_1}{h^2(1-y_2)^{2(k-1)}} \Bigg[\sum_{i_1+i_2+i_3=2} \frac{k\cdot k!(k+i_3-1)!(-1)^{i_3}}{2(k-i_1)!(k-i_2)!} h^{i_1}(1-y_2)^{k-i_1} \left(\frac{h^2-y_2}{h}\right)^{k-i_2} \left(-\frac{h}{y_2}\right)^{k+i_3} \\
&\quad + \sum_{i_1+i_2+i_3=k-1} \frac{k\cdot k!(2+i_3)!(-1)^{i_3}}{(k-i_1)!(k-i_2)!2!} \cdot h^{k+i_1-i_2} \left(\frac{h}{y_2}\right)^{3+i_3}\Bigg] \\
&\quad + \frac{\beta_y}{2(1-y_2)^{2k-1}} \Bigg[\sum_{i_1+i_2+i_3+i_4=1} \frac{k\cdot k!(k+i_4-1)!(-1)^{i_4}}{(k-i_1)!(k-i_2)!} h^{i_1}(1-y_2)^{k-i_1} \\
&\quad \times \left(\frac{h^2-y_2}{h}\right)^{k-i_2} \left(\frac{\sqrt{y_2}-y_2}{h}\right)^{1-i_3} \left(-\frac{h}{y_2}\right)^{k+i_4} \\
&\quad + \sum_{i_1+i_2+i_3+i_4=1} \frac{k\cdot k!(k+i_4-1)!(-1)^{i_4}}{(k-i_1)!(k-i_2)!} h^{i_1}(1-y_2)^{k-i_1} \\
&\quad \times \left(\frac{h^2-y_2}{h}\right)^{k-i_2} \left(\frac{-\sqrt{y_2}-y_2}{h}\right)^{1-i_3} \left(-\frac{h}{y_2}\right)^{k+i_4} \\
&\quad + \sum_{i_1+i_2+i_3+i_4=k-1} \frac{k\cdot k!(i_4+1)!(-1)^{i_4}}{(k-i_1)!(k-i_2)!} \cdot h^{k+i_1-i_2} \left(\left(\frac{\sqrt{y_2}}{h}\right)^{1-i_3} + \left(-\frac{\sqrt{y_2}}{h}\right)^{1-i_3}\right) \left(\frac{h}{y_2}\right)^{2+i_4} \\
&\quad - \sum_{i_1+i_2+i_3+i_4+i_5=2} \frac{k\cdot k!(k+i_5-1)! h^{i_1}(1-y_2)^{k-i_1}}{(k-i_1)!(k-i_2)!(-1)^{i_5}} \\
&\quad \times \left(\frac{h^2-y_2}{h}\right)^{k-i_2} \left(\frac{\sqrt{y_2}-y_2}{h}\right)^{1-i_3} \left(\frac{-\sqrt{y_2}-y_2}{h}\right)^{1-i_4} \left(-\frac{h}{y_2}\right)^{k+i_5} \\
&\quad - \sum_{i_1+i_2+i_3+i_4+i_5=k-1} \frac{k\cdot k!(i_5+2)!(-1)^{i_5}}{(k-i_1)!(k-i_2)!} \cdot h^{k+i_1-i_2} \left(\frac{\sqrt{y_2}}{h}\right)^{1-i_3} \left(-\frac{\sqrt{y_2}}{h}\right)^{1-i_4} \left(\frac{h}{y_2}\right)^{3+i_5}\Bigg],
\end{aligned}
$$

$$\operatorname{cov}(X_{f_k}, X_{f_l}) = -\frac{1}{2\pi^2} \lim_{r\downarrow 1} \oint\oint_{|\xi_1|=|\xi_2|=1} \frac{(1+h\xi_1)^l(\xi_1+h)^l \cdot (1+h\xi_2)^k(\xi_2+h)^k}{(1-y_2)^{2l+2k} \cdot (\xi_1 - r\xi_2)^2 \xi_1^l \xi_2^k} \, d\xi_1 \, d\xi_2$$
$$- \frac{(\beta_x y_1 + \beta_y y_2)(1-y_2)^2}{4\pi^2 h^2} \oint_{|\xi_1|=1} \frac{(1+h\xi_1)^l(\xi_1+h)^l}{(\xi_1 + y_2/h)^2 \xi_1^l} \, d\xi_1 \oint_{|\xi_2|=1} \frac{(1+h\xi_2)^k(\xi_2+h)^k}{(\xi_2 + y_2/(hr_2))^2 \xi_2^k} \, d\xi_2,$$

where

$$\begin{aligned}
&-\frac{1}{2\pi^2} \lim_{r\downarrow 1} \oint\oint_{|\xi_1|=|\xi_2|=1} \frac{(1+h\xi_1)^l(\xi_1+h)^l \cdot (1+h\xi_2)^k(\xi_2+h)^k}{(1-y_2)^{2l+2k} \cdot (\xi_1 - r\xi_2)^2 \xi_1^l \xi_2^k} \, d\xi_1 \, d\xi_2 \\
&\quad = -\frac{i}{\pi(1-y_2)^{2l+2k}} \oint_{|\xi_2|=1} \frac{(1+h\xi_2)^k(\xi_2+h)^k}{\xi_2^k} \left(\oint_{|\xi_1|=1} \frac{(1+h\xi_1)^l(\xi_1+h)^l}{\xi_1^l(\xi_1 - r\xi_2)^2} \, d\xi_1 \right) d\xi_2 \\
&\quad = -\frac{\mathrm{i}}{\pi(1-y_2)^{2l+2k}} \sum_{i_1+i_2+i_3=l-1} \frac{l \cdot l!(i_3+1)! h^{l+i_1-i_2}}{(l-i_1)!(l-i_2)!} \oint_{|\xi_2|=1} \frac{(1+h\xi_2)^k(\xi_2+h)^k}{\xi_2^{k+i_3+2}} \, d\xi_2 \\
&\quad = \frac{2}{(1-y_2)^{2l+2k}} \sum_{i_1+i_2+i_3=l-1} \sum_{j_1+j_2=k+i_3+1} \frac{l \cdot l!(i_3+1)!k!k!}{(l-i_1)!(l-i_2)!(k+i_3+1)!(k-j_1)!(k-j_2)!} h^{l+k+i_1-i_2+j_1-j_2}
\end{aligned}$$

and

$$\begin{aligned}
&\oint_{|\xi_1|=1} \frac{(1+h\xi_1)^l(\xi_1+h)^l}{(\xi_1 + y_2/h)^2 \xi_1^l} \, d\xi_1 \\
&\quad = 2\pi\mathrm{i} \Bigg[\sum_{i_1+i_2+i_3=1} \frac{l \cdot l!(l+i_3-1)!(-1)^{i_3}}{(l-i_1)!(l-i_2)!} h^{i_1} (1-y_2)^{l-i_1} \left(\frac{h^2-y_2}{h} \right)^{l-i_2} \left(-\frac{h}{y_2} \right)^{l+i_3} \\
&\qquad + \sum_{i_1+i_2+i_3=l-1} \frac{l \cdot l!(1+i_3)!(-1)^{i_3}}{(l-i_1)!(l-i_2)!} h^{l+i_1-i_2} \left(\frac{h}{y_2} \right)^{2+i_3} \Bigg].
\end{aligned}$$

So we obtain

$$\begin{aligned}
&\operatorname{cov}(X_{f_k}, X_{f_l}) \\
&\quad = \frac{2}{(1-y_2)^{2l+2k}} \sum_{i_1+i_2+i_3=l-1} \sum_{j_1+j_2=k+i_3+1} \frac{l \cdot l!(i_3+1)!k!k!}{(l-i_1)!(l-i_2)!(k+i_3+1)!(k-j_1)!(k-j_2)!} h^{l+k+i_1-i_2+j_1-j_2} \\
&\qquad + \frac{(\beta_x y_1 + \beta_y y_2)(1-y_2)^2}{h^2} \\
&\qquad \times \Bigg\{ \Bigg[\sum_{i_1+i_2+i_3=1} \frac{l \cdot l!(l+i_3-1)!(-1)^{i_3}}{(l-i_1)!(l-i_2)!} h^{i_1} (1-y_2)^{l-i_1} \left(\frac{h^2-y_2}{h} \right)^{l-i_2} \left(-\frac{h}{y_2} \right)^{l+i_3} \\
&\qquad + \sum_{i_1+i_2+i_3=l-1} \frac{l \cdot l!(1+i_3)!(-1)^{i_3}}{(l-i_1)!(l-i_2)!} h^{l+i_1-i_2} \left(\frac{h}{y_2} \right)^{2+i_3} \Bigg] \\
&\qquad \times \Bigg[\sum_{i_1+i_2+i_3=1} \frac{k \cdot k!(k+i_3-1)!(-1)^{i_3}}{(k-i_1)!(k-i_2)!} h^{i_1} (1-y_2)^{k-i_1} \left(\frac{h^2-y_2}{h} \right)^{k-i_2} \left(-\frac{h}{y_2} \right)^{k+i_3} \\
&\qquad + \sum_{i_1+i_2+i_3=k-1} \frac{k \cdot k!(1+i_3)!(-1)^{i_3}}{(k-i_1)!(k-i_2)!} h^{k+i_1-i_2} \left(\frac{h}{y_2} \right)^{2+i_3} \Bigg] \Bigg\}.
\end{aligned}$$

Example 4.3. *If $f = \mathrm{e}^x$, then by Taylor expansion, we have*

$$EX_f = \sum_{l=0}^{+\infty} \frac{1}{l!} EX_{f_l} \quad \text{and} \quad \operatorname{var}(X_f) = \sum_{k,l=0}^{+\infty} \operatorname{cov}(X_{f_k}, X_{f_l}),$$

where $f_l(x) = x^l$, EX_{f_l} and $\operatorname{cov}(X_{f_l}, X_{f_k})$ can be obtained in Example 4.2.

5. Comments and conclusions

In this paper, we have established the CLTs for LSS of an F matrix with explicit expressions of the asymptotic means and covariance functions. As a consequence, we have established the CLT for the LSS of a beta matrix, which is a matrix function of an F matrix.

It is well-known that high-dimensional problems are very important in practice. In the past, dimension reduction and feature extraction played pivotal roles in high-dimensional statistical problems. But a large portion of information contained in the original data would inevitably get lost if the dimension is very large. For example, in variable selection of multivariate linear regression models, one will lose all information contained in the unselected variables; in principal component analysis, all information contained in the "non-principal" components would be gone.

More work remains to be done. In the future, we hope to establish the CLTs for more general large-dimensional problems in practice.

6. Proofs

6.1. *Two lemmas*

Lemma 6.1. *Let $m_0(z) = \underline{m}_{y_2}(-\underline{m}(z))$ where $\underline{m}(z)$ is the solution of (2.10), then we have the following identities*

$$z = -\frac{m_0(z)(m_0(z) + 1 - y_1)}{(m_0(z) + 1/(1-y_2))(1-y_2)} \quad \text{and} \quad \underline{m}(z) = \frac{(1-y_2)(m_0(z) + 1/(1-y_2))}{m_0(z)(m_0(z)+1)}, \tag{6.1}$$

$$m_0'(z) = -\frac{(1-y_2)^2(m_0(z) + 1/(1-y_2))^2}{(1-y_2)m_0^2(z) + 2m_0(z) + 1 - y_1}, \tag{6.2}$$

$$\int \frac{\mathrm{d}F_{y_2}(x)}{x + \underline{m}(z)} = \frac{m_0(z)}{(1-y_2)\cdot(m_0(z) + 1/(1-y_2))}, \tag{6.3}$$

$$1 - y_1 \int \frac{\underline{m}^2(z)\,\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2} = \frac{(1-y_2)m_0^2(z) + 2m_0(z) + 1 - y_1}{(1-y_2)m_0^2(z) + 2m_0(z) + 1}, \tag{6.4}$$

$$\int \frac{x\cdot\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2} = \frac{m_0^2(z)}{(1-y_2)m_0^2(z) + 2m_0(z) + 1}, \tag{6.5}$$

$$\underline{m}'(z) = -\frac{(1-y_2)m_0^2 + 2m_0 + 1}{m_0^2(m_0+1)^2}\cdot m_0'(z) = -\bigl(1 - y_2 m_0^2(1+m_0)^{-2}\bigr)m_0^{-2}m_0'(z), \tag{6.6}$$

$$\left(\log\frac{(1-y_2)m_0^2(z) + 2m_0(z) + 1 - y_1}{(1-y_2)m_0^2(z) + 2m_0(z) + 1}\right)' = \frac{-2y_1\int \underline{m}^3(z)(z)x(x+\underline{m}(z))^{-3}\,\mathrm{d}F_{y_2}(x)}{[1 - y_1\int \underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)]^2}, \tag{6.7}$$

$$\bigl[\log\bigl(1 - y_2 m_0^2(1+m_0)^{-2}\bigr)\bigr]' = \frac{2y_2\underline{m}'(z)m_0^3(1+m_0)^{-3}}{(1 - y_2 m_0^2(1+m_0)^{-2})^2}, \tag{6.8}$$

where $m_0'(z) = \frac{\mathrm{d}}{\mathrm{d}z}m_0(z)$ and $\underline{m}'(z) = \frac{\mathrm{d}}{\mathrm{d}z}\underline{m}(z)$.

Proof. Because

$$\underline{m}_{y_2}(z) = -\frac{1-y_2}{z} + y_2 \cdot m_{y_2}(z) \tag{6.9}$$

then we obtain $m'_{y_2}(z) = -\frac{1-y_2}{y_2} \cdot \frac{1}{z^2} + \frac{1}{y_2} \cdot \underline{m}'_{y_2}(z)$, that is,

$$m'_{y_2}\big(-\underline{m}(z)\big) = -\frac{1-y_2}{y_2}\frac{1}{(\underline{m}(z))^2} + \frac{1}{y_2} \cdot \underline{m}'_{y_2}\big(-\underline{m}(z)\big).$$

So we have

$$\int \frac{\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2} = m'_{y_2}\big(-\underline{m}(z)\big) = -\frac{1-y_2}{y_2} \cdot \frac{1}{(\underline{m}(z))^2} + \frac{1}{y_2} \cdot \underline{m}'_{y_2}\big(-\underline{m}(z)\big). \tag{6.10}$$

By (2.10), we have

$$\begin{aligned} z &= -\frac{1}{\underline{m}(z)} + \int \frac{y_1\,\mathrm{d}F_{y_2}(x)}{x+\underline{m}(z)} \\ &= -\frac{1}{\underline{m}(z)} - \frac{y_1(1-y_2)}{y_2\underline{m}(z)} + \frac{y_1}{y_2}\underline{m}_{y_2}\big(-\underline{m}(z)\big) \\ &= \frac{y_1+y_2-y_1y_2}{y_2} \cdot \frac{1}{-\underline{m}(z)} + \frac{y_1}{y_2}\underline{m}_{y_2}\big(-\underline{m}(z)\big). \end{aligned} \tag{6.11}$$

Using the notation $h^2 = y_1 + y_2 - y_1y_2$ and differentiating both sides of the above identity, we obtain

$$1 = \frac{h^2}{y_2(\underline{m}(z))^2}\underline{m}'(z) - \frac{y_1}{y_2}\underline{m}'_{y_2}\big(-\underline{m}(z)\big)\underline{m}'(z).$$

This implies $\underline{m}'(z) = \frac{y_2(\underline{m}(z))^2}{h^2 - y_1(\underline{m}(z))^2\underline{m}'_{y_2}(-\underline{m}(z))}$ or

$$y_1\big(\underline{m}(z)\big)^2\underline{m}'_{y_2}\big(-\underline{m}(z)\big) = h^2 - \frac{y_2(\underline{m}(z))^2}{\underline{m}'(z)}, \tag{6.12}$$

where $\underline{m}'_{y_2}(-\underline{m}(z)) = \frac{\mathrm{d}}{\mathrm{d}\xi}\underline{m}_{y_2}(\xi)|_{\xi=-\underline{m}(z)}$ instead of $\frac{\mathrm{d}}{\mathrm{d}z}\underline{m}_{y_2}(-\underline{m}(z))$. So by (6.10) and (6.12), we have

$$1 - y_1\int \frac{(\underline{m}(z))^2\,\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2} = \frac{h^2}{y_2} - \frac{y_1(\underline{m}(z))^2\underline{m}'_2(-\underline{m}(z))}{y_2} = \frac{(\underline{m}(z))^2}{\underline{m}'(z)}. \tag{6.13}$$

That is,

$$\underline{m}'(z) = \frac{\underline{m}^2(z)}{1 - y_1\int(\underline{m}(z))^2\,\mathrm{d}F_{y_2}(x)/(x+\underline{m}(z))^2}. \tag{6.14}$$

The Stieltjes transform $\underline{m}_{y_2}(z)$ satisfies $z = -\frac{1}{\underline{m}_{y_2}(z)} + \frac{y_2}{1+\underline{m}_{y_2}(z)}$. Differentiating both sides, we obtain

$$1 = \left(\frac{1}{(\underline{m}_{y_2}(z))^2} - \frac{y_2}{(1+\underline{m}_{y_2}(z))^2}\right)\underline{m}'_{y_2}(z).$$

Therefore, $\underline{m}'_{y_2}(z) = \frac{(\underline{m}_{y_2}(z))^2}{1-y_2(\underline{m}_{y_2}(z))^2(1+\underline{m}_{y_2}(z))^{-2}}$ and thus

$$\underline{m}'_{y_2}\big(-\underline{m}(z)\big) = \frac{[\underline{m}_{y_2}(-\underline{m}(z))]^2}{1 - y_2 \cdot [\underline{m}_{y_2}(-\underline{m}(z))]^2 \cdot [1+\underline{m}_{y_2}(-\underline{m}(z))]^{-2}}. \tag{6.15}$$

Because $z=-\frac{1}{\underline{m}_{y_2}(z)}+\frac{y_2}{1+\underline{m}_{y_2}(z)}$, then we have

$$-\underline{m}(z)=-\frac{1}{\underline{m}_{y_2}(-\underline{m}(z))}+\frac{y_2}{1+\underline{m}_{y_2}(-\underline{m}(z))}=\frac{(y_2-1)\cdot(\underline{m}_{y_2}(-\underline{m}(z))-1/(y_2-1))}{\underline{m}_{y_2}(-\underline{m}(z))\cdot(\underline{m}_{y_2}(-\underline{m}(z))+1)}.$$

Let $m_0=\underline{m}_{y_2}(-\underline{m}(z))$ by (6.11), then we prove the conclusion (6.1) of the lemma

$$\underline{m}(z)=\frac{(1-y_2)\cdot(m_0+1/(1-y_2))}{m_0\cdot(m_0+1)}\quad\text{and}\quad z=-\frac{m_0(m_0+1-y_1)}{(m_0+1/(1-y_2))\cdot(1-y_2)}.\tag{6.16}$$

Differentiating the second identity in (6.16), we obtain

$$1=-\frac{[(2m_0+1-y_1)(m_0+1/(1-y_2))-m_0(m_0+1-y_1)]\underline{m}_0'}{(m_0+1/(1-y_2))^2(1-y_2)}.$$

Solving $\underline{m}_0'$, we obtain the second assertion (6.2) of the lemma

$$\underline{m}_0'=-\frac{(1-y_2)^2(m_0+1/(1-y_2))^2}{(1-y_2)m_0^2+2m_0+1-y_1}.$$

By the identity $\underline{m}_{y_2}(-\underline{m}(z))=\frac{1-y_2}{\underline{m}(z)}+y_2\cdot m_{y_2}(-\underline{m}(z))$, we obtain

$$\begin{aligned}\int\frac{\mathrm{d}F_{y_2}(x)}{x+\underline{m}(z)}=m_{y_2}\bigl(-\underline{m}(z)\bigr)&=\frac{\underline{m}_{y_2}(-\underline{m}(z))}{y_2}-\frac{1-y_2}{y_2}\cdot\frac{1}{\underline{m}(z)}\\&=\frac{\underline{m}_{y_2}(-\underline{m}(z))}{y_2}-\frac{1-y_2}{y_2}\cdot\frac{\underline{m}_{y_2}(-\underline{m}(z))(\underline{m}_{y_2}(-\underline{m}(z))+1)}{(1-y_2)(\underline{m}_{y_2}(-\underline{m}(z))+1/(1-y_2))}\\&=\frac{\underline{m}_{y_2}(-\underline{m}(z))}{(1-y_2)\cdot(\underline{m}_{y_2}(-\underline{m}(z))+1/(1-y_2))}\\&=\frac{m_0}{(1-y_2)(m_0+1/(1-y_2))}.\end{aligned}$$

This is the third conclusion (6.3) of the lemma. By (6.13), (6.15) and (6.16), we obtain the 4th conclusion (6.4) of the lemma

$$\begin{aligned}1-y_1\underline{m}^2\int\frac{\mathrm{d}F_{y_2}(x)}{(x+\underline{m})^2}&=\frac{h^2}{y_2}-\frac{y_1}{y_2}\frac{(y_2-1)^2(m_0-1/(y_2-1))^2\cdot m_0^2(1+m_0)^2}{m_0^2(m_0+1)^2\cdot[(1+m_0)^2-y_2\cdot m_0^2]}\\&=\frac{(1-y_2)m_0^2+2m_0+1-y_1}{(1-y_2)m_0^2+2m_0+1},\end{aligned}\tag{6.17}$$

where $m_0^2+\frac{2}{1-y_2}m_0+\frac{1-y_1}{1-y_2}=(m_0+\frac{1+h}{1-y_2})\cdot(m_0+\frac{1-h}{1-y_2})$. We obtain

$$\begin{aligned}\int\frac{x\cdot\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2}&=\int\frac{\mathrm{d}F_{y_2}(x)}{x+\underline{m}(z)}-\underline{m}(z)\int\frac{\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2}\\&=\frac{m_0/(1-y_2)}{(m_0+1/(1-y_2))}-\frac{m_0(m_0+1)/(1-y_2)}{[(1-y_2)m_0^2+2m_0+1](m_0+1/(1-y_2))}\\&=\frac{m_0^2}{(1-y_2)m_0^2+2m_0+1}.\end{aligned}$$

This is the 5th conclusion (6.5) of the lemma. By (6.16), we obtain

$$\begin{aligned}\underline{m}'(z) &= \frac{(1-y_2)(m_0(m_0+1)-(m_0+1/(1-y_2))(2m_0+1))m_0'}{m_0^2(m_0+1)^2}\\ &= \frac{-(1-y_2)(m_0+1/(1-y_2))^2+y_2/(1-y_2)}{m_0^2\cdot(m_0+1)^2}\underline{m}_0'\\ &= -\frac{(1-y_2)m_0^2+2m_0+1}{m_0^2(m_0+1)^2}\cdot\underline{m}_0'.\end{aligned}$$

This is the 6th conclusion (6.6) of the lemma. Furthermore, we obtain

$$\begin{aligned}\left(\log\frac{(1-y_2)m_0^2(z)+2m_0(z)+1-y_1}{(1-y_2)m_0^2(z)+2m_0(z)+1}\right)' &= \left[\log\left(1-y_1\int\underline{m}^2(z)\big(x+\underline{m}(z)\big)^{-2}\,\mathrm{d}F_{y_2}(x)\right)\right]'\\ &= \frac{(1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x))'}{1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)}\\ &= \frac{-2y_1\int\underline{m}(z)\underline{m}'(z)x(x+\underline{m}(z))^{-3}\,\mathrm{d}F_{y_2}(x)}{1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)}\\ &= \frac{-2y_1\int\underline{m}^3(z)(z)x(x+\underline{m}(z))^{-3}\,\mathrm{d}F_{y_2}(x)}{[1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)]^2},\end{aligned}$$

where the last equation holds by (6.14). This is the 7th conclusion (6.7) of the lemma.

$$\begin{aligned}\left[\log\left(1-y_2m_0^2(1+m_0)^{-2}\right)\right]' &= \frac{-2y_2m_0'(z)m_0(1+m_0)^{-3}}{1-y_2m_0^2(1+m_0)^{-2}}\\ &= \frac{-2y_2m_0'(z)[1-y_2m_0^2(1+m_0)^{-2}]m_0^{-2}m_0^3(1+m_0)^{-3}}{(1-y_2m_0^2(1+m_0)^{-2})^2}\\ &= \frac{2y_2\underline{m}'(z)m_0^3(1+m_0)^{-3}}{(1-y_2m_0^2(1+m_0)^{-2})^2}.\end{aligned}$$

This is the 8th conclusion (6.8) of the lemma.

So the proof of Lemma 6.1 is completed. □

In the computation of the mean function and covariance function of Bai and Silverstein [5], without their conditions on the 4th moment of X, by their (1.15), their (4.10) and (2.7) each should contain an additional term

$$-\beta_x\frac{pb_p^2}{n_1}\cdot E\left(e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_1\cdot e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\left(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\right)^{-1}\mathbf{S}_2^{-1/2}e_1\right)$$

and

$$\frac{\beta_x b_p(z_1)b_p(z_2)}{n_1^2}\sum_{j=1}^{n_1}\sum_{i=1}^{p}e_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z_1)\mathbf{S}_2^{-1/2}e_ie_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z_2)\mathbf{S}_2^{-1/2}e_i$$

respectively, where $b_p(z)=\frac{1}{1+n_1^{-1}E\operatorname{tr}S_2^{-1}D_1^{-1}(z)}$ and the definitions of D_1, e_i, E_j are in Lemma 6.2. The following lemma proves the convergence $E(e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_1\cdot e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I})^{-1}\mathbf{S}_2^{-1/2}e_1)$ and the uniform convergence of the diagonal elements $e_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z_1)\mathbf{S}_2^{-1/2}e_i$ for all i and j.

Lemma 6.2. *Suppose that the assumptions of Theorem 3.2 hold. Then for any z, we have*

$$\max_{i,j}\left|e_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z)\mathbf{S}_2^{-1/2}e_i+\frac{1}{z}\int\frac{\mathrm{d}F_{y_2}(x)}{x+\underline{m}(z)}\right|\to 0 \quad \textit{in } p, \tag{6.18}$$

$$\begin{aligned}&E\big(e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_1\cdot e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_1\big)\\&\quad\to\frac{1}{z^2}\int\frac{\mathrm{d}F_{y_2}(x)}{x+\underline{m}}\int\frac{x\,\mathrm{d}F_{y_2}(x)}{(x+\underline{m})^2},\end{aligned} \tag{6.19}$$

where

$$e_i=(\underbrace{0,\dots,0}_{i-1},1,\underbrace{0,\dots,0}_{p-i})',\qquad \mathbf{D}(z)=\mathbf{S}_2^{-1/2}\mathbf{S}_1\mathbf{S}_2^{-1/2}-z\mathbf{I},$$

$$\mathbf{D}_j(z)=\mathbf{D}(z)-r_jr_j^*,\qquad r_j=\frac{1}{\sqrt{n_1}}\mathbf{S}_2^{-1/2}\mathbf{X}_{\cdot j}$$

for $j=1,\dots,n_1$ and E_j denotes the conditional expectation given $r_1,\dots,r_j$.

In the proof of Lemma 6.2, we need Lemma 9.1 of Bai and Silverstein [6], Kolmogorov inequality and Burkholder inequality which are quoted below for easy reference.

Lemma 6.3. *Suppose that x_i, $i=1,\dots,n$, are independent, with $Ex_i=0$, $E|x_i|^2=1$, $\sup E|x_i|^4=\nu<\infty$ and $|x_i|\le\eta\sqrt{n}$ with $\eta>0$. Assume that $\mathbf{A}$ is a complex matrix. Then for any given $2\le p\le b\log(n\nu^{-1}\eta^4)$ and $b>1$, we have*

$$E\big|\boldsymbol{\alpha}^*\mathbf{A}\boldsymbol{\alpha}-\operatorname{tr}(\mathbf{A})\big|^p\le\nu n^p\big(n\eta^4\big)^{-1}\big(40b^2\|\mathbf{A}\|\eta^2\big)^p,$$

where $\|\mathbf{A}\|$ is the operator norm and $\boldsymbol{\alpha}=(x_1,\dots,x_n)^T$.

Lemma 6.4 (A generalization of Kolmogorov's inequality). *If $\{S_i,\mathcal{F}_i,1\le i\le n\}$ is a sub martingale, then for each λ,*

$$\lambda P\Big(\max_{i\le n}S_i\ge\lambda\Big)\le E(S_nI_{(\max_{i\le n}S_i>\lambda)}) \tag{6.20}$$

(*see Theorem* 2.1 *of Hall and Heyde* [19]).

Lemma 6.5 (Burkholder inequality). *Let $\{X_k\}$ be a complex martingale difference sequence with respect to the increasing σ-field $\mathcal{F}_k$, and let E_k denote conditional expectation w.r.t. $\mathcal{F}_k$. Then, for $p\ge 2$,*

$$E\left|\sum_{k=1}^n X_k\right|^p\le K_p\left(E\left(\sum_{k=1}^n E_{k-1}|X_k|^2\right)^{p/2}+E\sum_{k=1}^n|X_k|^p\right) \tag{6.21}$$

(*see Lemma* 2.13 *of Bai and Silverstein* [6]).

Proof of Lemma 6.2. First, we claim that for any random matrices $\mathbf{M}$ with non-random bound $\|\mathbf{M}\|\le K$, fixed $t>0$, $i\le p$ and z with $|\Im z|=v>0$, we have

$$P\Big(\sup_{j\le n_1}\big|E_je_i'\mathbf{S}_2^{-1/2}\mathbf{D}_j^{-1}(z)\mathbf{S}_2^{-1/2}\mathbf{M}e_i-E_je_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}\mathbf{M}e_i\big|\ge\varepsilon\Big)=\mathrm{o}\big(n_1^{-t}\big). \tag{6.22}$$

In fact,

$$\begin{aligned}
&\left|E_j e_i' \mathbf{S}_2^{-1/2}\mathbf{D}_j^{-1}(z)\mathbf{S}_2^{-1/2}\mathbf{M}e_i - E_j e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}\mathbf{M}e_i\right| \\
&\quad = \left|E_j e_i'\mathbf{S}_2^{-1/2}\big(\mathbf{D}_j^{-1}(z) - \mathbf{D}^{-1}(z)\big)\mathbf{S}_2^{-1/2}\mathbf{M}e_i\right| \\
&\quad = \left|E_j \frac{e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_j^{-1}(z) r_j r_j^* \mathbf{D}_j^{-1}\mathbf{S}_2^{-1/2}\mathbf{M}e_i}{1 + r_j^*\mathbf{D}_j^{-1} r_j}\right| \\
&\quad \le K \cdot E_j \left|e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_j^{-1}(z) r_j\right|^2.
\end{aligned}$$

By noticing $\frac{1}{n_1}|e_i'(\mathbf{S}_2^{-1/2}\mathbf{D}_j^{-1}(z)\mathbf{S}_2^{-1}\mathbf{D}_j^{-1}(z)\mathbf{S}_2^{-1/2})e_i| \le K/n_1$ and applying Lemma 6.3 by choosing $\eta = [\log n_1]$, one can easily prove (6.22).

To show the convergence of the first conclusion (6.18) in Lemma 6.2, we consider

$$e_i'\mathbf{S}_2^{-1/2} E_j \mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2} e_i = E_j e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2} e_i,$$

where $\mathbf{D}(z) = \mathbf{S}_2^{-1/2}\mathbf{S}_1\mathbf{S}_2^{-1/2} - z\cdot\mathbf{I}$, E_j denotes the conditional expectation with respect to the σ-field generated by $r_1, \ldots, r_j$ and E_0 denotes conditional expectation given $\mathbf{S}_2$. Now, we further extend the definition of E_j for negative j, that is, E_j denotes the conditional expectation when $\mathbf{Y}_{\cdot -j}, \mathbf{Y}_{\cdot -j+1}, \ldots, \mathbf{Y}_{\cdot n_2}$ are given. Note that E_{-n_2-1} denotes the unconditional expectation, and

$$\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2} = (\mathbf{S}_1 - z\cdot S_2)^{-1} := \widetilde{\mathbf{D}}^{-1}(z).$$

That is, the limits of the diagonal elements of

$$E_j\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2} = E_j\widetilde{\mathbf{D}}^{-1}(z)$$

are identical. To this end, employing Kolmogorov inequality (6.20) for martingales, we have

$$\begin{aligned}
I_i &:= P\Big(\sup_{-n_2\le j\le n_1}\left|E_j e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - E e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\right| \ge \varepsilon\Big) \\
&\le \varepsilon^{-4} E\left|E_{n_1} e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - E_{-n_2-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\right|^4 \\
&\qquad \text{(using Kolmogorov inequality (6.20))} \\
&= \varepsilon^{-4} E\left|\sum_{k=-n_2}^{n_1}(E_k - E_{k-1})e_i'\widetilde{\mathbf{D}}^{-1}(z)e_i\right|^4 \\
&= \varepsilon^{-4} E\left|\sum_{k=-n_2}^{n_1}(E_k - E_{k-1})e_i'\big(\widetilde{\mathbf{D}}^{-1}(z) - \widetilde{\mathbf{D}}_k^{-1}(z)\big)e_i\right|^4 \\
&\le \varepsilon^{-4} K_1 E\left|\sum_{k=1}^{n_1}(E_k - E_{k-1})e_i'\big(\widetilde{\mathbf{D}}^{-1}(z) - \widetilde{\mathbf{D}}_k^{-1}(z)\big)e_i\right|^4 \\
&\qquad + \varepsilon^{-4} K_1 E\left|\sum_{k=-n_2}^{0}(E_k - E_{k-1})e_i'\big(\widetilde{\mathbf{D}}^{-1}(z) - \widetilde{\mathbf{D}}_k^{-1}(z)\big)e_i\right|^4 \\
&\qquad \text{(using } E(|X_1| + |X_2|)^r \le 2^{r-1}(E|X_1|^r + E|X_2|^r),\ r \ge 1)
\end{aligned}$$

$$\leq \frac{K_1}{\varepsilon^4 n_1^4} E\left|\sum_{k=1}^{n_1}(E_k - E_{k-1})\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4 + \frac{K_1}{\varepsilon^4 n_2^4} E\left|\sum_{k=-n_2}^{0}(E_k - E_{k-1})\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^4,$$

where the last inequality uses the common-used formula of the inverse matrix

$$\left(\mathbf{A}+\mathbf{B}\mathbf{D}'\mathbf{B}'\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}+\mathbf{D}^{-1}\right)^{-1}\mathbf{B}'\mathbf{A}^{-1}, \tag{6.23}$$

K_1 is a positive constant and

$$\widetilde{\mathbf{D}}_k = \begin{cases} \widetilde{\mathbf{D}} - \frac{1}{n_1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*, & \text{if } k > 0, \\ \widetilde{\mathbf{D}}, & \text{if } k = 0, \\ \widetilde{\mathbf{D}} + \frac{z}{n_2}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*, & \text{if } k < 0. \end{cases}$$

Thus, by Burkholder inequality for martingale difference sequence, we have

$$\begin{aligned}
I_i &= \frac{K_1}{\varepsilon^4 n_1^4} E\left|\sum_{k=1}^{n_1}(E_k - E_{k-1})\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4 \\
&\quad + \frac{K_1}{\varepsilon^4 n_2^4} E\left|\sum_{k=-n_2}^{0}(E_k - E_{k-1})\frac{z\cdot e_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^4 \\
&\leq \frac{K}{\varepsilon^4 n_1^4}\cdot\left[E\left(\sum_{k=1}^{n_1}E_{k-1}\left|(E_k - E_{k-1})\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^2\right)^2\right. \\
&\quad \left. + \sum_{k=1}^{n_1}E\left|(E_k - E_{k-1})\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4\right] \quad \text{(using Burkholder inequality (6.21))} \\
&\quad + \frac{K}{\varepsilon^4 n_2^4}\left[E\left(\sum_{k=-n_2}^{0}E_{k-1}\left|(E_k - E_{k-1})\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^2\right)^2\right. \\
&\quad \left. + \sum_{k=-n_2}^{0}E\left|(E_k - E_{k-1})\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^4\right] \quad \text{(using Burkholder inequality (6.21))} \\
&\leq \frac{2K}{\varepsilon^4 n_1^4}\cdot\left[E\left(\sum_{k=1}^{n_1}E_{k-1}\left|E_k\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^2 + \sum_{k=1}^{n_1}\left|E_{k-1}\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^2\right)^2\right. \\
&\quad \left. + \sum_{k=1}^{n_1}E\left|E_k\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4 + \sum_{k=1}^{n_1}E\left|E_{k-1}\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4\right] \\
&\quad \text{(using } E(|X_1|+|X_2|)^r \leq 2^{r-1}(E|X_1|^r + E|X_2|^r), r \geq 1) \\
&\quad + \frac{2K}{\varepsilon^4 n_2^4}\left[E\left(\sum_{k=-n_2}^{0}E_{k-1}\left|E_k\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^2 + \left|E_{k-1}\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot -k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot -k}/n_2}\right|^2\right)^2\right.
\end{aligned}$$

$$+\sum_{k=-n_2}^{0} E\left|E_k \frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^4 + \sum_{k=-n_2}^{0} E\left|E_{k-1} \frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^4\Bigg]$$

(using $E(|X_1|+|X_2|)^r \le 2^{r-1}(E|X_1|^r + E|X_2|^r), r \ge 1$)

$$\le \frac{4K}{\varepsilon^4 n_1^4}\cdot\left[E\left(\sum_{k=1}^{n_1}\left|\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^2\right)^2 + \sum_{k=1}^{n_1} E\left|\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4\right]$$

$$+\frac{4K}{\varepsilon^4 n_2^4}\left[E\left(\sum_{k=-n_2}^{0}\left|\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^2\right)^2 + \sum_{k=-n_2}^{0} E\left|\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^4\right]$$

(using $E(E_k|X|)^2 \le E|X|^2$, $(E_k|X|)^2 \le E_k|X|^2$ and $E(E_k|X|)^4 \le E|X|^4$, for r.v. X). (6.24)

We quote the inequalities (3.4) of Bai and Silverstein [3] and (3.4) of Bai and Silverstein [5] as follows.

$$\left|\frac{1}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right| \le \frac{|z|}{v}, \qquad E\left|\frac{\mathbf{X}_{\cdot j}^*\widetilde{\mathbf{D}}_j^{-1}\mathbf{X}_{\cdot j} - \operatorname{tr}\widetilde{\mathbf{D}}_j^{-1}}{n_1}\right|^p \le \frac{K_p\delta^{2p-4}}{n_1} \quad (p \text{ even}) \quad \text{and}$$
$$\left|\frac{1}{1+n_1^{-1}\operatorname{tr}\widetilde{\mathbf{D}}_k^{-1}}\right| \le \frac{|z|}{v}. \tag{6.25}$$

Through some computations similar to (4.22) of Bai and Silverstein [5], we have

$$E\left|e_i'\mathbf{M}\mathbf{X}_{\cdot k}\right|^2 = \mathrm{O}(1) \quad \text{and} \quad E\left|e_i'\mathbf{M}\mathbf{X}_{\cdot k}\right|^4 = \mathrm{O}(1), \tag{6.26}$$

where $\mathbf{M}$ is a non-random matrix with finite operator norm.

Therefore, by (6.25) and (6.26), we have

$$\frac{K}{\varepsilon^4 n_1^4}\cdot E\left(\sum_{k=1}^{n_1}\left|\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^2\right)^2 \le \frac{K|z^4|}{\varepsilon^4 n_1^4 v^4}\cdot E\left(\sum_{k=1}^{n_1}\left|e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot k}^*\widetilde{\mathbf{D}}_k^{-1}e_i\right|^2\right)^2 = \mathrm{O}\left(n_1^{-2}\right)$$

and

$$\frac{K}{n_1^4}\sum_{k=1}^{n_1} E\left|\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot 1}^*\widetilde{\mathbf{D}}_k^{-1}e_i}{1+\mathbf{X}_{\cdot k}\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}/n_1}\right|^4 \le \frac{K|z|^4}{v^4 n_1^4}\sum_{k=1}^{n_1} E\left|e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{X}_{\cdot k}\mathbf{X}_{\cdot 1}^*\widetilde{\mathbf{D}}_k^{-1}e_i\right|^4 = \mathrm{O}\left(n_1^{-3}\right).$$

Furthermore, by noticing that

$$\left|\frac{1}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right| = \left|\frac{\bar z}{\bar z - |z|^2\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_k^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right| \le \frac{|z|}{v},$$

we can similarly prove that

$$\frac{K}{\varepsilon^4 n_2^4} E\left(\sum_{k=-n_2}^{0} E_{k-1}\left|\frac{ze_i'\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^2\right)^2 = \mathrm{O}\left(n_2^{-2}\right)$$

and

$$\frac{K}{\varepsilon^4 n_2^4}\cdot\sum_{k=-n_2}^{0} E\left|\frac{e_i'\widetilde{\mathbf{D}}_k^{-1}\mathbf{Y}_{\cdot-k}\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}e_i}{1-z\mathbf{Y}_{\cdot-k}^*\widetilde{\mathbf{D}}_{-k}^{-1}\mathbf{Y}_{\cdot-k}/n_2}\right|^4 = \mathrm{o}\left(n_2^{-3}\right).$$

That is,

$$I_i = \mathrm{O}\big(n_1^{-2}\big) + \mathrm{O}\big(n_2^{-2}\big).$$

Then we obtain

$$\begin{aligned}
&P\Big(\max_{i,j}\big|E_j e_i' \mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\big| \geq \varepsilon\Big)\\
&\quad = P\Bigg[\bigcup_{i=1}^{p}\Big(\max_{j}\big|E_j e_i' \mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\big| \geq \varepsilon\Big)\Bigg]\\
&\quad \leq \sum_{i=1}^{p} P\Big(\max_{j}\big|E_j e_i' \mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\big| \geq \varepsilon\Big)\\
&\quad = \sum_{i=1}^{p} I_i = p\cdot\big(\mathrm{O}\big(n_1^{-2}\big)+\mathrm{O}\big(n_2^{-2}\big)\big) = \mathrm{o}(1)
\end{aligned}$$

because the dimension p tends to infinity proportionally to the sample sizes n_1 and n_2. So we have

$$\max_{i,j}\big|E_j e_i' \mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i - Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i\big| \to 0 \quad \text{in } p.$$

If the $\mathbf{X}$ variables and $\mathbf{Y}$ variables are respectively identically distributed, then

$$Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i = \frac{1}{p}\operatorname{tr} E\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}.$$

Similar to (2.17) of Bai and Silverstein [5], we have

$$\begin{aligned}
Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}^{-1}(z)\mathbf{S}_2^{-1/2}e_i &= \frac{1}{p}E\big[\operatorname{tr}\big(\mathbf{S}_2^{-1}\mathbf{D}^{-1}(z)\big)\big] = \frac{n_1}{p}\cdot\bigg(\frac{1}{b_p(z)}-1\bigg)\\
&\to -\frac{1+z\underline{m}(z)}{zy_1\underline{m}(z)} = -\int\frac{\mathrm{d}F_{y_2}(x)}{z(x+\underline{m}(z))},
\end{aligned}\tag{6.27}$$

where $b_p(z) = \frac{1}{1+n_1^{-1}E\operatorname{tr}\mathbf{S}_2^{-1}\mathbf{D}_1^{-1}(z)}$. Thus, the first conclusion (6.18) in Lemma 6.2 follows.

Next, we shall show that the above limit holds true under the assumptions of Theorem 3.2. Let $\widetilde{\mathbf{D}}_{j,w} = \widetilde{\mathbf{D}} - \frac{1}{n_1}\mathbf{X}_{\cdot j}\mathbf{X}_{\cdot j}^* + \frac{1}{n_1}\mathbf{W}_j\mathbf{W}_j^*$, where $\mathbf{W}_j$ consists of i.i.d. entries distributed as X_{11}, that is, we change the jth term $\frac{1}{n_1}\mathbf{X}_{\cdot j}\mathbf{X}_{\cdot j}^*$ with an analogue $\frac{1}{n_1}\mathbf{W}_j\mathbf{W}_j^*$ with i.i.d. entries. We have

$$\begin{aligned}
Ee_i'\widetilde{\mathbf{D}}^{-1}e_i - Ee_i'\widetilde{\mathbf{D}}_{j,w}^{-1}e_i &= Ee_i'\big(\widetilde{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}_j^{-1}\big)e_i - Ee_i'\big(\widetilde{\mathbf{D}}_{j,w}^{-1}-\widetilde{\mathbf{D}}_j^{-1}\big)e_i\\
&= n_1^{-1}Ee_i'\widetilde{\mathbf{D}}_j^{-1}\big[\mathbf{X}_{\cdot j}\mathbf{X}_{\cdot j}^*\beta_j - \mathbf{W}_j\mathbf{W}_j^*\beta_{j,w}\big]\widetilde{\mathbf{D}}_j^{-1}e_i,
\end{aligned}$$

where $\beta_j = (1+n_1^{-1}\mathbf{X}_j^*\widetilde{\mathbf{D}}_j^{-1}\mathbf{X}_j)^{-1}$ and $\beta_{j,w} = (1+n_1^{-1}\mathbf{W}_j^*\widetilde{\mathbf{D}}_j^{-1}\mathbf{W}_j)^{-1}$ by (6.23). Let $\hat{\beta}_j = (1+n_1^{-1}\operatorname{tr}\widetilde{\mathbf{D}}_j^{-1})^{-1}$, $\hat{\gamma}_j = n_1^{-1}[\mathbf{X}_{\cdot j}^*\widetilde{\mathbf{D}}_j^{-1}\mathbf{X}_{\cdot j} - \operatorname{tr}\widetilde{\mathbf{D}}_j^{-1}]$ and $\hat{\gamma}_{j,w} = n_1^{-1}[\mathbf{W}_j^*\widetilde{\mathbf{D}}_j^{-1}\mathbf{W}_j - \operatorname{tr}\widetilde{\mathbf{D}}_j^{-1}]$. Noticing that $\beta_{j,w} = \hat{\beta}_j - \hat{\beta}_j\beta_{j,w}\hat{\gamma}_{j,w}$ and similar decomposition for β_j, we have by (6.25)

$$\begin{aligned}
&\big|Ee_i'\widetilde{\mathbf{D}}^{-1}e_i - Ee_i'\widetilde{\mathbf{D}}_{j,w}^{-1}e_i\big|\\
&\quad = n_1^{-1}\big|Ee_i'\widetilde{\mathbf{D}}_j^{-1}\big[\mathbf{X}_{\cdot j}\mathbf{X}_{\cdot j}^*\hat{\beta}_j\beta_j\hat{\gamma}_j - \mathbf{W}_j\mathbf{W}_j^*\beta_{j,w}\hat{\beta}_j\hat{\gamma}_{j,w}\big]\widetilde{\mathbf{D}}_j^{-1}e_i\big|\\
&\quad \leq \frac{K}{n_1}\big[\big(E\big|e_i'\widetilde{\mathbf{D}}_j^{-1}\mathbf{X}_{\cdot j}\big|^4E|\hat{\gamma}_j|^2\big)^{1/2} + \big(E\big|e_i'\widetilde{\mathbf{D}}_j^{-1}\mathbf{W}_j\big|^4E|\hat{\gamma}_{j,w}|^2\big)^{1/2}\big] = \mathrm{O}\big(n_1^{-3/2}\big).
\end{aligned}$$

Using the same approach, we can replace all terms $\frac{1}{n_1}\mathbf{X}_{\cdot j}\mathbf{X}_{\cdot j}^*$ in $\mathbf{S}_1$ by $\frac{1}{n_1}\mathbf{W}_j\mathbf{W}_j^*$, the total error will be controlled by $\mathrm{O}(n_1^{-1/2})$, and replace all terms in $\mathbf{S}_2$ by i.i.d. entries with a total error controlled by $\mathrm{O}(n_2^{-1/2})$. And then, we can show that the first conclusion (6.18) in Lemma 6.2 holds under the the assumptions of Theorem 3.2.

Now we consider the convergence of $E(e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_1\cdot e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I})^{-1}\mathbf{S}_2^{-1/2}e_1)$. First, we have

$$
\begin{aligned}
&E\big(e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big)\\
&\quad=E\big\{\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i+E_{-n_2-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\big)\\
&\qquad\times\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i+E_{-n_2-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big)\big\}\\
&\quad=E\big\{\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i+Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\big)\\
&\qquad\times\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i+Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big)\big\}\\
&\quad=Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\\
&\qquad+E\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big),
\end{aligned}
$$

where

$$
\begin{aligned}
&E_{n_1-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i=e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i,\\
&E_{-n_2-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i=Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i,\\
&E_{n_1-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i=e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i
\end{aligned}
$$

and

$$
E_{-n_2-1}e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i=Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i.
$$

Furthermore,

$$
\begin{aligned}
&\big|E\big((E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big)\big|^2\\
&\quad\le E\big|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\big|^2\cdot E\big|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big|^2\\
&\quad\le\big(E\big|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\big|^4\\
&\qquad\cdot E\big|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big|^4\big)^{1/2}.
\end{aligned}
$$

Similar to in the proof of I_i of Lemma 6.2, we have $E|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i|^4=\mathrm{O}(n_1^{-2})+\mathrm{O}(n_2^{-2})$ and $E|(E_{n_1-1}-E_{-n_2-1})e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I})^{-1}\mathbf{S}_2^{-1/2}e_i|^4=\mathrm{O}(n_1^{-2})+\mathrm{O}(n_2^{-2})$. Then we have

$$
\begin{aligned}
&E\big(e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i\big)\\
&\quad=Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot Ee_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_i+\mathrm{o}(1).
\end{aligned}
$$

Similar to (6.27), we have $\frac{1}{p}E[\operatorname{tr}(\mathbf{S}_2^{-1}\mathbf{D}_{n_1}^{-1}(z))]\to-\int\frac{\mathrm{d}F_{y_2}(x)}{z(x+\underline{m})}$. Using the decomposition (4.13) of Bai and Silverstein [5] and similar arguments (4.14)–(4.20), we have $\frac{1}{p}E\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I})^{-1}\mathbf{S}_2^{-1/2}e_i\to-\frac{1}{z}\int\frac{x\,\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2}$. Then we obtain

$$
E\big(e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_1\cdot e_1'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\big(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I}\big)^{-1}\mathbf{S}_2^{-1/2}e_1\big)\to\frac{1}{z^2}\int\frac{\mathrm{d}F_{y_2}(x)}{x+\underline{m}}\int\frac{x\,\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z))^2}.
$$

Similarly, we can obtain that the above limit holds true under the assumptions of Theorem 3.2.

The proof of Lemma 6.2 is completed. □

6.2. *Proofs of Theorem* 3.1 *and Corollary* 3.1

Proof of Theorem 3.1. Following the same techniques of truncation and normalization given in Bai and Silverstein [5] (see lines −9 to −6 from the bottom of p. 559), we may assume the following additional assumptions:

- $EX_{jk}=0$, $EY_{jk}=0$ and $E|X_{jk}|^2=1$, $E|Y_{jk}|^2=1$;
- $E|X_{jk}|^4=1+\kappa+\mathrm{o}(1)$ and $E|Y_{jk}|^4=1+\kappa+\mathrm{o}(1)$;
- for the complex case, $EX_{jk}^2=\mathrm{o}(n_1^{-1})$ and $EY_{jk}^2=\mathrm{o}(n_2^{-1})$.

In fact, through truncation and normalization, the means and variances of X_{jk} and Y_{jk} can equal to zero and one, respectively. But the second moments of complex r.v.'s X_{jk} and Y_{jk} have $\mathrm{o}(n_1^{-1})$ and $\mathrm{o}(n_2^{-1})$ respectively, and their fourth moments have $\mathrm{o}(1)$ except for $1+\kappa$.

Split our proofs into two steps. Write

$$n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr]=n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]+n_1\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr],$$

where $\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)$ and $\underline{m}_{\mathbf{y_n}}(z)$ are unique roots whose imaginary parts having the same sign as that of z to the following equations by (2.10)

$$z=-\frac{1}{\underline{m}^{\{y_{n_1},H_{n_2}\}}}+y_{n_1}\cdot\int\frac{\mathrm{d}F_{n_2}(t)}{t+\underline{m}^{\{y_{n_1},H_{n_2}\}}}\quad\text{and}\quad z=-\frac{1}{\underline{m}_{\mathbf{y_n}}}+y_{n_1}\cdot\int\frac{\mathrm{d}F_{y_{n_2}}(t)}{t+\underline{m}_{\mathbf{y_n}}}.$$

In the following Steps 1 and 2, we unify the real and complex cases, by the notation κ, β_x and β_y.

Step 1. Consider conditional distribution of

$$n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]. \tag{6.28}$$

given $\mathbf{S}_2=\{\text{all }\mathbf{S}_2\}$. Going along the lines of the proof of Lemma 1.1 of Bai and Silverstein [5], we can similarly prove that the conditional distribution of

$$n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]=p\bigl[m_{\mathbf{n}}(z)-m^{\{y_{n_1},H_{n_2}\}}(z)\bigr]$$

given $\mathbf{S}_2$ converges to a Gaussian process $M_1(z)$ on the contour $\mathcal{C}$ with mean function

$$E\bigl(M_1(z)|\mathbf{S}_2\bigr)=(\kappa-1)\cdot\frac{y_1\int\underline{m}(z)^3x[x+\underline{m}(z)]^{-3}\,\mathrm{d}F_{y_2}(x)}{[1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)]^2} \tag{6.29}$$

for $z\in\mathcal{C}$ and covariance function

$$\operatorname{cov}\bigl(M_1(z_1),M_1(z_2)|\mathbf{S}_2\bigr)=\kappa\cdot\left(\frac{\underline{m}'(z_1)\cdot\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2}-\frac{1}{(z_1-z_2)^2}\right) \tag{6.30}$$

for $z_1,z_2\in\mathcal{C}$. Note that the mean and covariance of the limiting distribution are independent of the conditioning $\mathbf{S}_2$, which shows that the limiting distribution of this part is independent of the limit of the next part because the asymptotic mean and covariances are non-random.

Step 2. Now, we consider the limiting process of

$$n_1\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr]=p\bigl[m^{\{y_{n_1},H_{n_2}\}}(z)-m_{\mathbf{y_n}}(z)\bigr]. \tag{6.31}$$

By (2.10), we have

$$z=-\frac{1}{\underline{m}_{\mathbf{y_n}}}+y_{n_1}\cdot\int\frac{t}{1+t\cdot\underline{m}_{\mathbf{y_n}}}\,\mathrm{d}H_{y_{n_2}}(t)=-\frac{1}{\underline{m}_{\mathbf{y_n}}}+y_{n_1}\cdot\int\frac{\mathrm{d}F_{y_{n_2}}(t)}{t+\underline{m}_{\mathbf{y_n}}}.$$

On the other hand, $\underline{m}^{\{y_{n_1},H_{n_2}\}}$ is the solution to the equation

$$z=-\frac{1}{\underline{m}^{\{y_{n_1},H_{n_2}\}}}+y_{n_1}\cdot\int\frac{t\cdot\mathrm{d}H_{n_2}(t)}{1+t\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}}=-\frac{1}{\underline{m}^{\{y_{n_1},H_{n_2}\}}}+y_{n_1}\cdot\int\frac{\mathrm{d}F_{n_2}(t)}{t+\underline{m}^{\{y_{n_1},H_{n_2}\}}}.$$

By the definition of Stieltjes transform, then the above two equations become

$$z=-\frac{1}{\underline{m}_{\mathbf{y_n}}}+y_{n_1}\cdot m_{y_{n_2}}(-\underline{m}_{\mathbf{y_n}})\quad\text{and}\quad z=-\frac{1}{\underline{m}^{\{y_{n_1},H_{n_2}\}}}+y_{n_1}\cdot m_{n_2}\bigl(-\underline{m}^{\{y_{n_1},H_{n_2}\}}\bigr).$$

The difference of the above two identities yields

$$\begin{aligned}0&=\frac{\underline{m}^{\{y_{n_1},H_{n_2}\}}-\underline{m}_{\mathbf{y_n}}}{\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}}+y_{n_1}\bigl[m_{n_2}\bigl(-\underline{m}^{\{y_1,H_{n_2}\}}\bigr)-m_{n_2}(-\underline{m}_{\mathbf{y_n}})+m_{n_2}(-\underline{m}_{\mathbf{y_n}})-m_{y_{n_2}}(-\underline{m}_{\mathbf{y_n}})\bigr]\\&=\frac{\underline{m}^{\{y_{n_1},H_{n_2}\}}-\underline{m}_{\mathbf{y_n}}}{\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}}-y_{n_1}\int\frac{(\underline{m}^{\{y_{n_1},H_{n_2}\}}-\underline{m}_{\mathbf{y_n}})\,\mathrm{d}F_{n_2}(t)}{(t+\underline{m}^{\{y_{n_1},H_{n_2}\}})(t+\underline{m}_{\mathbf{y_n}})}+y_{n_1}\bigl[m_{n_2}(-\underline{m}_{\mathbf{y_n}})-m_{y_{n_2}}(-\underline{m}_{\mathbf{y_n}})\bigr].\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}&n_1\cdot\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr]\\&\quad=-y_{n_1}\cdot\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}\cdot\frac{n_1[m_{n_2}(-\underline{m}_{\mathbf{y_n}})-m_{y_{n_2}}(-\underline{m}_{\mathbf{y_n}})]}{1-y_{n_1}\cdot\int\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}\,\mathrm{d}F_{n_2}(t)/((t+\underline{m}_{\mathbf{y_n}})\cdot(t+\underline{m}^{\{y_{n_1},H_{n_2}\}}))}\\&\quad=-\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{y_{n_1},H_{n_2}}\cdot\frac{n_2[\underline{m}_{n_2}(-\underline{m}_{\mathbf{y_n}})-\underline{m}_{y_{n_2}}(-\underline{m}_{\mathbf{y_n}})]}{1-y_{n_1}\cdot\int\underline{m}_{\mathbf{y_n}}\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}\,\mathrm{d}F_{n_2}(t)/((t+\underline{m}_{\mathbf{y_n}})(t+\underline{m}^{\{y_{n_1},H_{n_2}\}}))}.\end{aligned}\tag{6.32}$$

We then consider the limiting process of

$$n_2\cdot\bigl[\underline{m}_{n_2}\bigl(-\underline{m}_{\mathbf{y_n}}(z)\bigr)-\underline{m}_{y_{n_2}}\bigl(-\underline{m}_{\mathbf{y_n}}(z)\bigr)\bigr].$$

Noticing that for any $z\in\mathbb{C}^+\cup\overline{\mathbb{C}^+}$, $\underline{m}_{\mathbf{y_n}}(z)\to\underline{m}(z)$, the limiting distribution of

$$n_2\cdot\bigl[\underline{m}_{n_2}\bigl(-\underline{m}_{\mathbf{y_n}}(z)\bigr)-\underline{m}_{y_{n_2}}\bigl(-\underline{m}_{\mathbf{y_n}}(z)\bigr)\bigr]$$

is the same as that of

$$n_2\cdot\bigl[\underline{m}_{n_2}\bigl(-\underline{m}(z)\bigr)-\underline{m}_{y_{n_2}}\bigl(-\underline{m}(z)\bigr)\bigr].$$

It can be shown that when z runs along $\mathcal{C}$ clockwise, $-\underline{m}(z)$ would enclose the support of F_{y_2} clockwise without intersect the support. Again, using Lemma 1.1 in Bai and Silverstein [5] (with minor modification), we conclude that $n_2[\underline{m}_{n_2}(-\underline{m}(z))-\underline{m}_{y_{n_2}}(-\underline{m}(z))]$ converges weakly to a Gaussian process $M_2(\cdot)$ on $z\in\mathcal{C}$ with mean function and covariance function

$$E\bigl(M_2(z)\bigr)=(\kappa-1)\cdot\frac{y_2\cdot[\underline{m}_{y_2}(-\underline{m}(z))]^3\cdot[1+\underline{m}_{y_2}(-\underline{m}(z))]^{-3}}{[1-y_2\cdot(\underline{m}_{y_2}(-\underline{m}(z))/(1+\underline{m}_{y_2}(-\underline{m}(z))))^2]^2}\tag{6.33}$$

and

$$\operatorname{cov}\bigl(M_2(z_1),M_2(z_2)\bigr)=\kappa\left(\frac{\underline{m}'_{y_2}(-\underline{m}(z_1))\cdot\underline{m}'_{y_2}(-\underline{m}(z_2))}{[\underline{m}_{y_2}(-\underline{m}(z_1))-\underline{m}_{y_2}(-\underline{m}(z_2))]^2}-\frac{1}{(\underline{m}(z_1)-\underline{m}(z_2))^2}\right)\tag{6.34}$$

for $z_1,z_2\in\mathcal{C}$.

Because

$$\frac{-\underline{m}_{\mathbf{y_n}}(z)\cdot \underline{m}^{\{y_{n_1},H_{n_2}\}}(z)}{1-y_{n_1}\cdot\int \underline{m}_{\mathbf{y_n}}(z)\cdot \underline{m}^{\{y_{n_1},H_{n_2}\}}\,\mathrm{d}F_{n_2}(t)/((t+\underline{m}_{\mathbf{y_n}}(z))(t+\underline{m}^{\{y_{n_1},H_{n_2}\}}))}$$

converges to

$$-\underline{m}'(z)=\frac{-\underline{m}^2(z)}{1-y_1\cdot \underline{m}^2(z)\cdot\int \mathrm{d}F_{y_2}(t)/[t+\underline{m}(z)]^2}$$

by (6.14), then we conclude that by (6.32)

$$n_1\cdot\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr]$$

converges weakly to a Gaussian process $M_3(\cdot)$ satisfying

$$M_3(z)=-\underline{m}'(z)\cdot M_2(z)$$

with the means $E(M_3(z))=-\underline{m}'(z)\cdot EM_2(z)$ and covariance functions $\operatorname{cov}(M_3(z_1),M_3(z_2))=\underline{m}'(z_1)\underline{m}'(z_2)\cdot\operatorname{cov}(M_2(z_1),M_2(z_2))$. Because the limit of

$$n_1\cdot\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]$$

conditioning on $\mathcal{S}_2$ is independent of the ESD of S_{n_2}, we know that the limits of

$$n_1\cdot\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]\quad\text{and}\quad n_1\cdot\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr]$$

are asymptotically independent. Finally, we conclude that $n_1\cdot[\underline{m}_{\mathbf{n}}(z)-\underline{m}_{\mathbf{y_n}}(z)]$ converges weakly to a two-dimensional Gaussian process $M_1(z)+M_3(z)$ with mean function

$$E\bigl(M_1(z)+M_3(z)\bigr)=(\kappa-1)\cdot\frac{y_1\int \underline{m}^3(z)x[x+\underline{m}(z)]^{-3}\,\mathrm{d}F_{y_2}(x)}{[1-y_1\int \underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)]^2} \tag{6.35}$$

$$-(\kappa-1)\cdot\frac{y_2\cdot \underline{m}'(z)[\underline{m}_{y_2}(-\underline{m}(z))]^3\cdot[1+\underline{m}_{y_2}(-\underline{m}(z))]^{-3}}{[1-y_2\cdot(\underline{m}_{y_2}(-\underline{m}(z))/(1+\underline{m}_{y_2}(-\underline{m}(z))))^2]^2} \tag{6.36}$$

and covariance function

$$\begin{aligned}&\operatorname{cov}\bigl(M_1(z_1)+M_3(z_1),M_1(z_2)+M_3(z_2)\bigr)\\&\quad=\kappa\cdot\frac{\underline{m}'(z_1)\underline{m}'(z_2)\underline{m}'_{y_2}(-\underline{m}(z_1))\cdot\underline{m}'_{y_2}(-\underline{m}(z_2))}{[\underline{m}_{y_2}(-\underline{m}(z_1))-\underline{m}_{y_2}(-\underline{m}(z_2))]^2}-\frac{\kappa}{(z_1-z_2)^2}.\end{aligned} \tag{6.37}$$

Because $\int f(x)\,\mathrm{d}G(x)=-\frac{1}{2\pi\mathrm{i}}\int f(z)m_G(z)\,\mathrm{d}z$, then we obtain that the LSS of F matrix

$$\biggl(\int f_1(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x),\ldots,\int f_k(x)\,\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)\biggr)$$

converges weakly to a Gaussian vector $(X_{f_1},\ldots,X_{f_k})$ where

$$EX_{f_i}=-\frac{1}{2\pi\mathrm{i}}\oint f_i(z)E\bigl(M_1(z)+M_3(z)\bigr)\,\mathrm{d}z$$

and

$$\operatorname{cov}(X_{f_i},X_{f_j})=-\frac{1}{4\pi^2}\oint\oint f_i(z)f_j(z)\operatorname{cov}\bigl(M_1(z_1)+M_3(z_1),M_1(z_2)+M_3(z_2)\bigr)\,\mathrm{d}z_1\,\mathrm{d}z_2.$$

Recall $m_0(z) = \underline{m}_{y_2}(-\underline{m}(z))$. By Lemma 6.1, we have

$$-\frac{\oint f_i(z)\cdot(6.35)\,\mathrm{d}z}{2\pi\mathrm{i}} = \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\left(\frac{(1-y_2)m_0^2(z)+2m_0(z)+1-y_1}{(1-y_2)m_0^2(z)+2m_0(z)+1}\right),$$

$$-\frac{\oint f_i(z)\cdot(6.36)\,\mathrm{d}z}{2\pi\mathrm{i}} = \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)$$

and

$$-\frac{\oint\oint f_i(z_1)f_j(z_2)\cdot(6.37)\,\mathrm{d}z}{4\pi^2} = -\frac{\kappa}{4\pi^2}\oint\oint\frac{f_i(z_1)f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)-m_0(z_2))^2}.$$

Finally we obtain

$$EX_{f_i} = \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\left(\frac{(1-y_2)m_0^2(z)+2m_0(z)+1-y_1}{(1-y_2)m_0^2(z)+2m_0(z)+1}\right) + \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)$$

and

$$\operatorname{cov}(X_{f_i},X_{f_j}) = -\frac{\kappa}{4\pi^2}\oint\oint\frac{f_i(z_1)f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)-m_0(z_2))^2}.$$

Proof of Theorem 3.1 is completed. □

Proof of Corollary 3.1. In Lemma 6.1 it is proved that $m_0(z)$ satisfies the equation

$$z = -\frac{m_0(z)(m_0(z)+1-y_1)}{(1-y_2)(m_0(z)+1/(1-y_2))}. \tag{6.38}$$

It is also known that the support of the LSD $F_{y_1,y_2}(x)$ of F matrix is

$$\left[a = \frac{(1-h)^2}{(1-y_2)^2}, b = \frac{(1+h)^2}{(1-y_2)^2}\right],$$

when $y_1 \le 1$ or the above interval with a singleton $\{0\}$ when $y_1 > 1$. Therefore, $m_0(a) = \underline{m}_{y_2}(-\underline{m}(a))$ and $m_0(b) = \underline{m}_{y_2}(-\underline{m}(b))$ are real numbers which are the real roots of equations

$$a = \frac{m_0(a)\cdot[m_0(a)+1-y_1]}{[m_0(a)-1/(y_2-1)]\cdot(y_2-1)} \quad\text{and}\quad b = \frac{m_0(b)\cdot[m_0(b)+1-y_1]}{[m_0(b)-1/(y_2-1)]\cdot(y_2-1)}.$$

So we obtain $m_0(b) = -\frac{1+h}{1-y_2}$ and $m_0(a) = -\frac{1-h}{1-y_2}$. Clearly, when z runs in the positive direction around the support interval $[a,b]$ of $F_{\mathbf{y}}(x)$, $m_0(z) = \underline{m}_{y_2}(-\underline{m}(z))$ runs in the positive direction around the interval

$$I = \left(-\frac{1+h}{1-y_2}, -\frac{1-h}{1-y_2}\right).$$

Let $m_0(z) = -\frac{1+hr\xi}{1-y_2}$, where $r > 1$ but very close to 1, $|\xi| = 1$. By (6.38), we have

$$z = \frac{1+h^2+hr^{-1}\bar{\xi}+hr\xi}{(1-y_2)^2}.$$

This shows that when ξ anticlockwise runs along the unit circle, z anticlockwise runs a contour which closely incloses the interval $[a,b]$ when r is close to 1, where $a,b=\frac{(1\mp h)^2}{(1-y_2)^2}$. Therefore, as $r\downarrow 1$, we have

$$\begin{aligned}
EX_{f_i} &= \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\biggl(\frac{(1-y_2)m_0^2(z)+2m_0(z)+1-y_1}{(1-y_2)m_0^2(z)+2m_0(z)+1}\biggr) \\
&\quad + \frac{\kappa-1}{4\pi\mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\bigl(1-y_2m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr) \\
&= \lim_{r\downarrow 1}\frac{\kappa-1}{4\pi\mathrm{i}}\oint_{|\xi|=1} f_i\biggl(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\biggr)\biggl[\frac{1}{\xi-r^{-1}}+\frac{1}{\xi+r^{-1}}-\frac{1}{\xi-\sqrt{y_2}/h}-\frac{1}{\xi+\sqrt{y_2}/h}\biggr]\mathrm{d}\xi \\
&\quad + \frac{\kappa-1}{4\pi\mathrm{i}}\oint_{|\xi|=1} f_i\biggl(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\biggr)\biggl[\frac{1}{\xi-\sqrt{y_2}/h}+\frac{1}{\xi+\sqrt{y_2}/h}-\frac{2}{\xi+y_2/h}\biggr]\mathrm{d}\xi \\
&= \lim_{r\downarrow 1}\frac{\kappa-1}{4\pi\mathrm{i}}\oint_{|\xi|=1} f_i\biggl(\frac{1+h^2+2h\Re(\xi)}{(1-y_2)^2}\biggr)\biggl[\frac{1}{\xi-r^{-1}}+\frac{1}{\xi+r^{-1}}-\frac{2}{\xi+y_2/h}\biggr]\mathrm{d}\xi.
\end{aligned}$$

By Lemma 6.1, we have

$$\underline{m}'(z) = -\frac{(1-y_2)m_0^2(z)+2m_0+1}{m_0^2(z)\cdot(m_0(z)+1)^2}\cdot m_0'(z) = \frac{(1-y_2)^2}{hr}\cdot\frac{(\xi+\sqrt{y_2}/hr)(\xi-\sqrt{y_2}/hr)\xi'}{(\xi+y_2/hr)^2(\xi+1/hr)^2}$$

and

$$\underline{m}(z) = -\frac{(1-y_2)^2}{hr}\frac{\xi}{(\xi+1/hr)(\xi+y_2/hr)}.$$

Making variable change $m_0(z_j)=-\frac{1+hr_j\xi_j}{1-y_2}$ for $j=1,2$ where $r_2>r_1>1$. Similarly, one can prove that as $1<r_1<r_2\to 1^+$, we have

$$\begin{aligned}
&\operatorname{cov}(X_{f_i},X_{f_j}) \\
&\quad = -\frac{\kappa}{4\pi^2}\oint\oint\frac{f_i(z_1)f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)-m_0(z_2))^2} \\
&\quad = -\lim_{\substack{1<r_1<r_2\\ r_2\to 1^+}}\frac{\kappa}{4\pi^2}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1}\frac{f_i((1+h^2+2h\Re(\xi_1))/(1-y_2)^2)f_j((1+h^2+2h\Re(\xi_2))/(1-y_2)^2)}{(r_1\xi_1-r_2\xi_2)^2}\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2 \\
&\quad = -\lim_{r\downarrow 1}\frac{\kappa}{4\pi^2}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1}\frac{f_i((1+h^2+2h\Re(\xi_1))/(1-y_2)^2)f_j((1+h^2+2h\Re(\xi_2))/(1-y_2)^2)}{(\xi_1-r\xi_2)^2}\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2.
\end{aligned}$$

Proof of Corollary 3.1 is completed. □

6.3. *Proofs of Theorem 3.2 and Corollary 3.2*

Proof of Theorem 3.2. Similar to the proof of Theorem 3.1, we use the same truncation and centralization technique of Bai and Silverstein [5]. That is, we may assume the same additional assumptions as described in the proof of Theorem 3.1 except the fourth moments of X_{jk} and Y_{jk}. Here we assume that $E|X_{jk}|^4=\beta_x+\kappa+\mathrm{o}(1)$ and $E|Y_{jk}|^4=\beta_y+\kappa+\mathrm{o}(1)$.

Similar to the proof of Theorem 3.1, we split

$$n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr] = n_1\bigl[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\bigr]+n_1\bigl[\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)-\underline{m}_{\mathbf{y_n}}(z)\bigr].$$

In Step 1, checking the proof of Lemma 1.1 of Bai and Silverstein [5], one finds that conditional distribution of

$$n_1\big[\underline{m}_{\mathbf{n}}(z)-\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)\big] \tag{6.39}$$

given $\mathbf{S}_2$ still converges to a Gaussian process $M_1(z)$ on the contour $\mathcal{C}$. When computing the asymptotic mean function where their formula (1.15) is used, without their condition on the 4th moment of $\mathbf{X}$-variables, the mean function should include an additional term by Lemma 6.2

$$\beta_x\cdot\frac{y_1\cdot\underline{m}^3(z)\cdot\int\mathrm{d}F_{y_2}(x)/(x+\underline{m}(z))\int x\cdot\mathrm{d}F_{y_2}(x)/(x+\underline{m}(z))^2}{1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(x)} \tag{6.40}$$

which is the limit of

$$\frac{\underline{m}(z)\cdot\beta_x\cdot(p/n_1)\cdot b_p^2E(e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}\mathbf{S}_2^{-1/2}e_i\cdot e_i'\mathbf{S}_2^{-1/2}\mathbf{D}_{n_1}^{-1}(\underline{m}\mathbf{S}_2^{-1}+\mathbf{I})^{-1}\mathbf{S}_2^{-1/2}e_i)}{1-y_1\int\underline{m}^2(z)(x+\underline{m}(z))^{-2}\,\mathrm{d}F_{y_2}(t)}$$

ever dropped in (4.10) and (4.12) of Bai and Silverstein [5]. Similarly, when computing the asymptotic covariance function, one finds that there should be an additional term by Lemma 6.2

$$\beta_x\cdot y_1\cdot\int\frac{\underline{m}'(z_1)\cdot x\cdot\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z_1))^2}\int\frac{\underline{m}'(z_2)\cdot x\cdot\mathrm{d}F_{y_2}(x)}{(x+\underline{m}(z_2))^2} \tag{6.41}$$

which is the limit of

$$\frac{\partial^2}{\partial z_1\,\partial z_2}\left(\frac{\beta_x\cdot b_p(z_1)b_p(z_2)}{n_1^2}\sum_{j=1}^{n_1}\sum_{i=1}^{p}e_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z_1)\mathbf{S}_2^{-1/2}e_i\cdot e_i'\mathbf{S}_2^{-1/2}E_j\mathbf{D}_j^{-1}(z_2)\mathbf{S}_2^{-1/2}e_i\right)$$

ever dropped in (2.7) of Bai and Silverstein [5]. Note that the mean and covariance of the limiting distribution are independent of the conditioning $\mathbf{S}_2$, which shows that the limiting distribution of this part is independent of the limit of the next part because the asymptotic mean and covariances are non-random.

As in Step 2 of the proof of Theorem 3.1, we have the same formula (6.32). The process

$$n_2\cdot\big[\underline{m}_{n_2}\big(-\underline{m}_{\mathbf{y_n}}(z)\big)-\underline{m}_{y_{n_2}}\big(-\underline{m}_{\mathbf{y_n}}(z)\big)\big]$$

also tends to Gaussian process $M_2(z)$ on the contour $\mathcal{C}$.

Checking the computation of the asymptotic mean function and covariance function, we find that they each have an additional terms respectively

$$\beta_y\cdot\frac{y_2\cdot m_0^3(z)\cdot(1+m_0(z))^{-3}}{1-y_2\cdot m_0^2(z)\cdot(1+m_0(z))^{-2}}$$

and

$$\beta_y\cdot y_2\cdot\frac{m_0'(z_1)}{(1+m_0(z_1))^2}\cdot\frac{m_0'(z_2)}{(1+m_0(z_2))^2},$$

which are the special cases of (6.40) and (6.41) respectively and $m_0(z)=\underline{m}_{y_2}(-\underline{m}(z))$.

Because

$$\begin{aligned}&\frac{-\underline{m}_{\mathbf{y_n}}(z)\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}(z)}{1-y_{n_1}\cdot\int\underline{m}_{\mathbf{y_n}}(z)\cdot\underline{m}^{\{y_{n_1},H_{n_2}\}}\,\mathrm{d}F_{n_2}(t)/((t+\underline{m}_{\mathbf{y_n}}(z))(t+\underline{m}^{\{y_{n_1},H_{n_2}\}}))}\\&\quad\to-\underline{m}'(z)=\frac{-\underline{m}^2(z)}{1-y_1\cdot\underline{m}^2(z)\cdot\int\mathrm{d}F_{y_2}(t)/[t+\underline{m}(z)]^2},\end{aligned}$$

we conclude that $n_1 \cdot [\underline{m}^{\{y_{n_1}, H_{n_2}\}}(z) - \underline{m}_{\mathbf{y_n}}(z)]$ converges weakly to a Gaussian process $M_3(\cdot)$ satisfying

$$M_3(z) = -\underline{m}'(z) \cdot M_2(z)$$

with the means $E(M_3(z)) = -\underline{m}'(z) \cdot EM_2(z)$ and covariance functions $\operatorname{cov}(M_3(z_1), M_3(z_2)) = \underline{m}'(z_1)\underline{m}'(z_2) \cdot \operatorname{cov}(M_2(z_1), M_2(z_2))$. Because the limit of $n_1 \cdot [\underline{m}_{\mathbf{n}}(z) - \underline{m}^{\{y_{n_1}, H_{n_2}\}}(z)]$ conditioning on $\{\text{all } \mathbf{S}_2\}$ is independent of the ESD of $\mathbf{S}_2$, we know that the limits of

$$n_1 \cdot \big[\underline{m}_{\mathbf{n}}(z) - \underline{m}^{\{y_{n_1}, H_{n_2}\}}(z)\big] \quad \text{and} \quad n_1 \cdot \big[\underline{m}^{\{y_{n_1}, H_{n_2}\}}(z) - \underline{m}_{\mathbf{y_n}}(z)\big]$$

are asymptotically independent. Then we obtain that $n_1 \cdot [\underline{m}_{\mathbf{n}}(z) - \underline{m}_{\mathbf{y_n}}(z)]$ converges weakly to a two-dimensional Gaussian process $M_1(z) + M_3(z)$ with means

$$\begin{aligned}
E\big(M_1(z) + M_3(z)\big) = {} & (\kappa - 1) \cdot \frac{y_1 \int \underline{m}^3(z) x [x + \underline{m}(z)]^{-3} \, \mathrm{d}F_{y_2}(x)}{[1 - y_1 \int \underline{m}^2(z)(x + \underline{m}(z))^{-2} \, \mathrm{d}F_{y_2}(x)]^2} \\
& + \beta_x \cdot \frac{y_1 \cdot \underline{m}^3(z) \cdot \int \mathrm{d}F_{y_2}(x)/(x + \underline{m}(z)) \int x \cdot \mathrm{d}F_{y_2}(x)/(x + \underline{m}(z))^2}{1 - y_1 \int \underline{m}^2(z)(x + \underline{m}(z))^{-2} \, \mathrm{d}F_{y_2}(x)} \\
& - (\kappa - 1) \cdot \underline{m}'(z) \frac{y_2 \cdot [m_0(z)]^3 \cdot [1 + m_0(z)]^{-3}}{[1 - y_2 \cdot (m_0(z)/(1 + m_0(z)))^2]^2} \\
& - \beta_y \cdot \underline{m}'(z) \frac{y_2 \cdot m_0^3(z) \cdot (1 + m_0(z))^{-3}}{1 - y_2 \cdot m_0^2(z) \cdot (1 + m_0(z))^{-2}}
\end{aligned}$$

and

$$\begin{aligned}
& \operatorname{cov}\big(M_1(z_1) + M_3(z_1), M_1(z_2) + M_3(z_2)\big) \\
& \quad = \beta_x \cdot y_1 \cdot \int \frac{\underline{m}'(z_1) \cdot x \cdot \mathrm{d}F_{y_2}(x)}{(x + \underline{m}(z_1))^2} \int \frac{\underline{m}'(z_2) \cdot x \cdot \mathrm{d}F_{y_2}(x)}{(x + \underline{m}(z_2))^2} - \frac{\kappa}{(z_1 - z_2)^2} \\
& \qquad + \kappa \cdot \frac{\underline{m}'(z_1)\underline{m}'(z_2) m_0'(z_1) \cdot m_0'(z_2)}{[m_0(z_1) - m_0(z_2)]^2} + \beta_y \cdot y_2 \cdot \frac{\underline{m}'(z_1) m_0'(z_1)}{(1 + m_0(z_1))^2} \cdot \frac{\underline{m}'(z_2) m_0'(z_2)}{(1 + m_0(z_2))^2}.
\end{aligned}$$

Because $\int f(x) \, \mathrm{d}G(x) = -\frac{1}{2\pi \mathrm{i}} \int f(z) m_G(z) \, \mathrm{d}z$, then we obtain that the Linear spectral statistics of F matrix

$$\left(\int f_1(x) \, \mathrm{d}\widetilde{G}_{\mathbf{n}}(x), \ldots, \int f_k(x) \, \mathrm{d}\widetilde{G}_{\mathbf{n}}(x) \right)$$

converge weakly to a Gaussian vector $(X_{f_1}, \ldots, X_{f_k})$ where

$$EX_{f_i} = -\frac{1}{2\pi \mathrm{i}} \oint f_i(z) E\big(M_1(z) + M_3(z)\big) \, \mathrm{d}z$$

and

$$\operatorname{cov}(X_{f_i}, X_{f_j}) = -\frac{1}{4\pi^2} \oint \oint f_i(z) f_j(z) \operatorname{cov}\big(M_1(z_1) + M_3(z_1), M_1(z_2) + M_3(z_2)\big) \, \mathrm{d}z_1 \, \mathrm{d}z_2.$$

Then by Lemma 6.1, we have

$$\begin{aligned}
EX_{f_i} = {} & \frac{\kappa - 1}{4\pi \mathrm{i}} \oint f_i(z) \, \mathrm{d} \log\left(\frac{(1 - y_2) m_0^2(z) + 2m_0(z) + 1 - y_1}{(1 - y_2) m_0^2(z) + 2m_0(z) + 1} \right) \\
& + \frac{\beta_x \cdot y_1}{2\pi \mathrm{i}} \oint f_i(z) \big(m_0(z) + 1\big)^{-3} \, \mathrm{d}m_0(z)
\end{aligned}$$

$$+\frac{\kappa-1}{4\pi \mathrm{i}}\oint f_i(z)\,\mathrm{d}\log\bigl(1-y_2 m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)$$
$$+\frac{\beta_y}{4\pi \mathrm{i}}\oint f_i(z)\bigl(1-y_2 m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)\,\mathrm{d}\log\bigl(1-y_2 m_0^2(z)\bigl(1+m_0(z)\bigr)^{-2}\bigr)$$

and

$$\operatorname{cov}(X_{f_i},X_{f_j}) = -\frac{\kappa}{4\pi^2}\oint\oint\frac{f_i(z_1)f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)-m_0(z_2))^2} - \frac{(\beta_x y_1+\beta_y y_2)}{4\pi^2}\oint\oint\frac{f_i(z_1)f_j(z_2)\,\mathrm{d}m_0(z_1)\,\mathrm{d}m_0(z_2)}{(m_0(z_1)+1)^2(m_0(z_2)+1)^2}. \qquad \square$$

Proof of Corollary 3.2. By the same variable change $m_0(z)=-\frac{1+hr\xi}{1-y_2}$ as given in proof of Corollary 3.1, the proof of the corollary can be done by only technical calculus and hence the details are omitted. $\square$

6.4. *Proof of Theorem* 4.1

Proof of Theorem 4.1. As a consequence of the CLTs for F-matrix, we establish the CLT for LSS of beta matrix of the form: $\mathbf{S}_2(\mathbf{S}_2+d\cdot\mathbf{S}_1)^{-1}=(\mathbf{I}+d\cdot\mathbf{S}_1\mathbf{S}_2^{-1})^{-1}$, a matrix functional of F-matrix, where d is a positive constant. Because the ith eigenvalues λ_i' and λ_i of beta matrix and F matrix have the relation

$$\lambda_i'=\frac{1}{1+d\cdot\lambda_i},$$

then the ESD of the beta matrix satisfies

$$F_{0,\mathbf{n}}(x)=\frac{1}{p}\sum_{i=1}^{p}I_{\{\lambda_i'\le x\}}=\frac{1}{p}\sum_{i=1}^{p}I_{\{1/(1+d\lambda_i)\le x\}}=\frac{1}{p}\sum_{i=1}^{p}I_{\{\lambda_i\ge 1/d(1/x-1)\}}=1-\frac{1}{p}\sum_{i=1}^{p}I_{\{\lambda_i<1/d(1/x-1)\}}$$

that is,

$$F_{0,\mathbf{n}}(x)=1-\frac{1}{p}\sum_{i=1}^{p}I_{\{\lambda_i\le 1/d(1/x-1)\}}+\frac{I_{\{\lambda_i=1/d(1/x-1)\}}}{p}=1-f_{\mathbf{n}}\left(\frac{1}{d}\left(\frac{1}{x}-1\right)\right)+\frac{I_{\{\lambda_i=1/d(1/x-1)\}}}{p}.$$

Then we have

$$F_{0,\mathbf{y_n}}(x)=1-F_{\mathbf{y_n}}\left(\frac{1}{d}\left(\frac{1}{x}-1\right)\right)$$

except some discontinuous points. So we obtain

$$\begin{aligned}\int_0^1 f_i(x)\,\mathrm{d}F_{0,\mathbf{y_n}}(x)&=\int_0^1 f_i(x)\,\mathrm{d}\left(1-F_{\mathbf{y_n}}\left(\frac{1}{d}\left(\frac{1}{x}-1\right)\right)\right)\\&=\int_{+\infty}^0 f_i\left(\frac{1}{1+dt}\right)\mathrm{d}\bigl(1-F_{\mathbf{y_n}}(t)\bigr)\\&=\int_0^{+\infty} f_i\left(\frac{1}{1+dx}\right)\mathrm{d}F_{\mathbf{y_n}}(x)\end{aligned}$$

and

$$\int f_k(x)\,\mathrm{d}F_{0,\mathbf{n}}=\frac{1}{p}\sum_{i=1}^{p}f\bigl(\lambda_i'\bigr)=\frac{1}{p}\sum_{i=1}^{p}f\left(\frac{1}{1+d\cdot\lambda_i}\right)=\int f_k\left(\frac{1}{1+dx}\right)\mathrm{d}f_{\mathbf{n}}.$$

Then the linear spectral statistics of the beta matrix and F matrix satisfy

$$\left(\int f_1(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x),\ldots,\int f_k(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x)\right)$$
$$=\left(\int f_1\left(\frac{1}{dx+1}\right)\mathrm{d}\widetilde{G}_{\mathbf{n}}(x),\ldots,\int f_k\left(\frac{1}{dx+1}\right)\mathrm{d}\widetilde{G}_{\mathbf{n}}(x)\right).$$

Then under Assumptions [A]–[B], the linear spectral statistics

$$\left(\int f_1(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x),\ldots,\int f_k(x)\,\mathrm{d}\widehat{G}_{\mathbf{n}}(x)\right)$$

of the beta matrix $\mathbf{S}_2(\mathbf{S}_2+d\cdot\mathbf{S}_1)^{-1}$, converge weakly to a Gaussian vector $(X_{f_1},\ldots,X_{f_k})$ whose means and covariances are the same as Theorem 3.2 except that $f_i(x)$ and $f_j(x)$ are replaced by $f_i(\frac{1}{d\cdot x+1})$ and $f_j(\frac{1}{d\cdot x+1})$, respectively. □

Acknowledgements

The author wishes to express her thanks to Professor Z. D. Bai for raising this problem and giving many valuable suggestions, which are of great help to the completion of the paper. The author thanks Professors J. W. Silverstein, J. F. Yao and X. M. He for their suggestions and their helps in polishing my English. This work is supported by NSFC 10701021, 11071035 and NENU-STC07001.

References

[1] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields* **134** (2006) 283–338. MR2222385
[2] Z. D. Bai. A note on asymptotic joint distribution of the eigenvalues of a noncentral multivariate F-matrix. Technical report, Central for Multivariate Analysis, Univ. Pittsburgh, 1984.
[3] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large dimensional random matrices. *Ann. Probab.* **26** (1998) 316–345. MR1617051
[4] Z. D. Bai and J. W. Silverstein. Exact separation of eigenvalues of large dimensional sample covariance matrices. *Ann. Probab.* **27** (1999) 1536–1555. MR1733159
[5] Z. D. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** (2004) 553–605. MR2040792
[6] Z. D. Bai and J. W. Silverstein. *Spectral Analysis of Large-Dimensional Random Matrices*, 2nd edition. Springer, New York, 2010. MR2567175
[7] Z. D. Bai and J. F. Yao. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli* **11** (2005) 1059–1092. MR2189081
[8] Z. D. Bai, Y. Q. Yin and P. R. Krishnaiah. On LSD of product of two random matrices when the underlying distribution is isotropic. *J. Multivariate Anal.* **19** (1986) 189–200. MR0847583
[9] A. Boutet De Monvel, L. Pastur and M. Shcherbina. On the statistical mechanics approach in the random matrix theory, integrated density of states. *J. Stat. Phys.* **79** (1995) 585–611. MR1327898
[10] T. Cabanal-Duvillard. Fluctuations de la loi empirique de grande matrices aléatoires. *Ann. Inst. H. Poincaré Probab. Statist.* **37** (2001) 373–402. MR1831988
[11] S. Chatterjee. Fluctuations of eigenvalues and second order poincaré inequalities. *Probab. Theory Related Fields* **143** (2009) 1–40. MR2449121
[12] O. Costin and J. L. Lebowitz. Gaussian fluctuation in random matrices. *Phys. Rev. Lett.* **75** (1995) 69–72.
[13] P. Diaconis and S. N. Evans. Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** (2001) 2615–2633. MR1828463
[14] I. Dumitriu and A. Edelman. Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. *J. Math. Phys.* **47** (2006) 063302. MR2239975
[15] V. L. Girko. *Theory of Random Determinants.* Kluwer Academic Publishers, London, 1990. MR1080966
[16] A. Guionnet. Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. *Ann. Inst. H. Poincaré Probab. Statist.* **38** (2002) 341–384. MR1899457
[17] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim and L. Pastur. A new approach for capacity analysis of large dimensional multi-antenna channels. *IEEE Trans. Inform. Theory* **54** (2008) 3987–4004. MR2450782
[18] W. Hachem, P. Loubaton and J. Najim. A CLT for information-theoretical statistics of Gram random matrices with a given variance profile. *Ann. Appl. Probab.* **18** (2008) 2071–2130. MR2473651

[19] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application*. Academic Press, New York, 1980. MR0624435
[20] C. P. Hughes, J. P. Keating and N. O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.* **220** (2001) 429–451. MR1844632
[21] S. Israelson. Asymptotic fluctuations of a particle system with singular interaction. *Stochastic Process. Appl.* **93** (2001) 25–56. MR1819483
[22] T. Jiang. Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles. *Probab. Theory Related Fields* **144** (2009) 221–246. MR2480790
[23] K. Johansson. On random matrices from the classical compact groups. *Ann. Math.* **145** (1997) 519–545. MR1454702
[24] K. Johansson. On the fluctuation of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** (1998) 151–204. MR1487983
[25] I. M. Johnstone. High dimensional statistical inference and random matrices. In *International Congress of Mathematicians*, Vol. I 307–333. Eur. Math. Soc. Zürich, Switzerland, 2007. MR2334195
[26] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** (1982) 1–38. MR0650926
[27] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2+it)$. *Comm. Math. Phys.* **214** (2000) 57–89. MR1794265
[28] A. M. Khorunzhy, B. A. Knoruzhenko and L. A. Pastur. Asymptotic properties of large random matrices with independent entrices. *J. Math. Phys.* **37** (1996) 5033–5060. MR1411619
[29] J. A. Mingo and R. Speicher. Second order freeness and fluctuations of random matrices I, Gaussian and Wishart matrices and cyclic Fock spaces. *J. Funct. Anal.* **235** (2006) 226–270. MR2216446
[30] K. C. S. Pillai. Percentage points of the largest root of the multivariate beta matrix. *Biometrika* **54** (1967) 189–194. MR0215433
[31] K. C. S. Pillai and B. N. Flury. Percentage points of the largest characteristic root of the multivariate beta matrix. *Comm. Statist.* **13** (1984) 2199–2237. MR0754832
[32] B. Ridelury and J. W. Silverstein. Gaussian fluctuations for non-Hermitian random matrix ensembles. *Ann. Probab.* **34** (2006) 2118–2143. MR2294978
[33] B. Rider and B. Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not.* **2007** (2007) Article ID rnm006. MR2361453
[34] J. W. Silverstein. The limiting eigenvalue distribution of a multivariate F-matrix. *SIAM J. Math. Anal.* **16** (1985) 641–646. MR0783987
[35] J. W. Silverstein and S. I. Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Multivariate Anal.* **54** (1995) 295–309. MR1345541
[36] Y. A. Sinaĭ and A. Soshnikov. Central limit theorems for traces of large random matrices with independent entries. *Bol. Soc. Brasil. Mat.* **29** (1998) 1–24. MR1620151
[37] Y. A. Sinaĭ and A. Soshnikov. A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices. *Funct. Anal. Appl.* **32** (1998) 114–131. MR1647832
[38] A. Soshnikov. Gaussian limits for determinantal random point fields. *Ann. Probab.* **28** (2002) 171–181. MR1894104
[39] K. Wieand. Eigenvalue distributions of random unitary matrices. *Probab. Theory Related Fields* **123** (2002) 202–224. MR1900322
[40] Y. Q. Yin, Z. D. Bai and P. R. Krishnaiah. Limiting behavior of the eigenvalues of a multivariate F-matrix. *J. Multivariate Anal.* **13** (1983) 508–516. MR0727036