Bernoulli **21**(3), 2015, 1538–1574
DOI: 10.3150/14-BEJ613

Convergence of the empirical spectral distribution function of Beta matrices

ZHIDONG BAI[1,*], JIANG HU[1,**], GUANGMING PAN[2] and WANG ZHOU[3]

[1]*KLASMOE and School of Mathematics & Statistics, Northeast Normal University, Changchun, P.R.C., 130024. E-mail:* [*]*baizd@nenu.edu.cn;* [**]*huj156@nenu.edu.cn*
[2]*Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371. E-mail: gmpan@ntu.edu.sg*
[3]*Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546. E-mail: stazw@nus.edu.sg*

Let $\mathbf{B}_n = \mathbf{S}_n(\mathbf{S}_n + \alpha_n \mathbf{T}_N)^{-1}$, where $\mathbf{S}_n$ and $\mathbf{T}_N$ are two independent sample covariance matrices with dimension p and sample sizes n and N, respectively. This is the so-called Beta matrix. In this paper, we focus on the limiting spectral distribution function and the central limit theorem of linear spectral statistics of $\mathbf{B}_n$. Especially, we do not require $\mathbf{S}_n$ or $\mathbf{T}_N$ to be invertible. Namely, we can deal with the case where $p > \max\{n, N\}$ and $p < n + N$. Therefore, our results cover many important applications which cannot be simply deduced from the corresponding results for multivariate F matrices.

Keywords: Beta matrices; CLT; LSD; multivariate statistical analysis

1. Introduction

In the last two decades, more and more large dimensional data sets appear in scientific research. When the dimension of data or number of parameters becomes large, the classical methods could reduce statistical efficiency significantly. In order to analyze those large data sets, many new statistical techniques, such as large dimensional multivariate statistical analysis (MSA) based on the random matrix theory (RMT), have been developed. In this paper, we will investigate a widely used type of random matrices in MSA which are called Beta matrices.

Firstly we introduce some definitions and terminology associated with Beta matrices. Let $\mathbf{X}_n = (x_{ij})_{p\times n}$, where $\{x_{ij}\}$ are independent and identically distributed (i.i.d.) random variables with mean zero and variance one, and similarly let $\mathbf{T}_N = N^{-1}\mathbb{X}_N\mathbb{X}_N^*$ be another sample covariance matrix independent of $\mathbf{S}_n$, where $\mathbb{X}_N = (\mathbb{x}_{ij})_{p\times N}$ and $\{\mathbb{x}_{ij}\}$ are i.i.d. random variables with mean zero and variance one. The Beta matrix is defined as

$$\mathbf{B}_n = \mathbf{S}_n(\mathbf{S}_n + \alpha_n \mathbf{T}_N)^{-1}, \tag{1.1}$$

where α_n is a positive constant. For any $n \times n$ matrix $\mathbf{A}$ with only real eigenvalues, we denote $F^{\mathbf{A}}$ as the empirical spectral distribution function (ESDF) of $\mathbf{A}$, that is $F^{\mathbf{A}}(x) = \frac{1}{n}\sum_{i=1}^n I(\lambda_i^{\mathbf{A}} \le x)$, where $\lambda_i^{\mathbf{A}}$ denotes the ith smallest eigenvalue of $\mathbf{A}$ and $I(\cdot)$ is the indicator function. In addition, we shall call $\int f(x)\,\mathrm{d}F^{\mathbf{A}}(x) = \frac{1}{n}\sum_{k=1}^n f(\lambda_k^{\mathbf{A}})$ a linear spectral statistics (LSS) of matrix $\mathbf{A}$. In this paper, we focus on the limiting ESDF and the central limit theorem (CLT) of LSS of $\mathbf{B}_n$.

1350-7265 © 2015 ISI/BS

One motivation to study Beta matrices is that their ESDFs are very useful in MSA, such as in the test of equality of k ($k \geq 2$) covariance matrices, multivariate analysis of variance, the independence test of sets of variables, canonical correlation analysis and so on. There is a huge literature regarding this kind of matrices. One may refer to [1,9,11] for more details. For pedagogical reasons, we provide one statistical application of Beta matrices as follows.

Let $\{\mathbf{z}_1^{(1)}, \dots, \mathbf{z}_n^{(1)}\}$ be an i.i.d. sample drawn from a p-dimensional distribution and $\{\mathbf{z}_1^{(2)}, \dots, \mathbf{z}_N^{(2)}\}$ be an i.i.d. sample drawn from another p-dimensional distribution. Suppose $\boldsymbol{\mu}_i = \mathbb{E}\mathbf{z}_1^{(i)} = \mathbf{0}$ and $\boldsymbol{\Sigma}_i = \mathbb{V}\mathrm{ar}\, \mathbf{z}_1^{(i)}$, $i = 1, 2$. Write $\mathbf{z}_j^{(1)} = \boldsymbol{\Sigma}_1^{1/2}\mathbf{X}_{(\cdot,j)}$ and $\mathbf{z}_j^{(2)} = \boldsymbol{\Sigma}_1^{1/2}\mathbb{X}_{(\cdot,j)}$ where $\mathbf{X}_{(\cdot,j)}$ ($\mathbb{X}_{(\cdot,j)}$) is the jth column of $\mathbf{X}_n$ ($\mathbb{X}_N$) and $\boldsymbol{\Sigma}_i^{1/2}$ is any square root of $\boldsymbol{\Sigma}_i$. We wish to test

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \quad \text{v.s.} \quad H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2.$$

This is one of the most elementary problems in MSA, for which there are lots of test statistics. If we write $\mathbf{Z}_n^{(1)} = n^{-1}\sum_{i=1}^n \mathbf{z}_i^{(1)}(\mathbf{z}_i^{(1)})^*$ and $\mathbf{Z}_N^{(2)} = N^{-1}\sum_{j=1}^n \mathbf{z}_j^{(2)}(\mathbf{z}_j^{(2)})^*$, then all the following L_j, $j = 1, 2, \dots, 5$ are the most frequently used test statistics for H_0 (see Chapter 8 in [11]).

$$\begin{aligned}
L_1 &= \log \frac{|\mathbf{Z}_n^{(1)}|^n \cdot |\mathbf{Z}_N^{(2)}|^N}{|c_n\mathbf{Z}_n^{(1)} + c_N\mathbf{Z}_N^{(2)}|^{n+N}} = \int \bigl(n\log(x/c_n) - N\log\bigl((1-x)/c_N\bigr)\bigr)\,\mathrm{d}F^{\mathbf{B}_n}(x), \\
L_2 &= \mathrm{tr}\bigl(\mathbf{Z}_N^{(2)}\bigl(\mathbf{Z}_n^{(1)}\bigr)^{-1}\bigr) = p\int \frac{1-x}{\alpha_n x}\,\mathrm{d}F^{\mathbf{B}_n}(x), \\
L_3 &= \log\bigl|\mathbf{Z}_n^{(1)}\bigl(\mathbf{Z}_n^{(1)} + \alpha_n\mathbf{Z}_N^{(2)}\bigr)^{-1}\bigr| = p\int \log x\,\mathrm{d}F^{\mathbf{B}_n}(x), \\
L_4 &= \mathrm{tr}\bigl(\mathbf{Z}_n^{(1)}\bigl(\mathbf{Z}_n^{(1)} + \alpha_n\mathbf{Z}_N^{(2)}\bigr)^{-1}\bigr) = p\int x\,\mathrm{d}F^{\mathbf{B}_n}(x), \\
L_5 &= c_n\,\mathrm{tr}\bigl(\mathbf{Z}_n^{(1)}\bigl(c_n\mathbf{Z}_n^{(1)} + c_N\mathbf{Z}_N^{(2)}\bigr)^{-1} - \mathbf{I}\bigr)^2 + c_N\,\mathrm{tr}\bigl(\mathbf{Z}_N^{(2)}\bigl(c_n\mathbf{Z}_n^{(1)} + c_N\mathbf{Z}_N^{(2)}\bigr)^{-1} - \mathbf{I}\bigr)^2 \\
&= c_n p\int \bigl(c_n^{-1}x - 1\bigr)^2\,\mathrm{d}F^{\mathbf{B}_n}(x) + c_N p\int \bigl(c_N^{-1}(1-x) - 1\bigr)^2\,\mathrm{d}F^{\mathbf{B}_n}(x),
\end{aligned} \tag{1.2}$$

where $c_n = n/(n+N)$, $c_N = N/(n+N)$ and $\alpha_n = N/n$. Apparently all the above test statistics are linear functionals of the ESDF of Beta matrices $\mathbf{B}_n$, which are all the LSS of $\mathbf{B}_n$. It is already well known that the classical limit theorems for those LSS are not valid when the dimension is large. So it is crucial to investigate the sequence $\{F^{\mathbf{B}_n}\}$ in the large dimensional case. The following result tells us the limiting behavior of $\{F^{\mathbf{B}_n}\}$ as $p, n, N \to \infty$.

Theorem 1.1 (Limiting spectral distribution function (LSDF)). *Assume on a common probability space*:

(i) *For each* i, j, n, $x_{ij} = x_{nij}$ *are i.i.d. with* $\mathbb{E}x_{11} = 0$, $\mathbb{E}|x_{11}|^2 = 1$.
(ii) $\alpha_n \to \alpha > 0$ *and* $y_n = p/n \to y > 0$.
(iii) *For each* k, l, N, $\mathbb{x}_{kl} = \mathbb{x}_{Nkl}$ *are i.i.d. with* $\mathbb{E}\mathbb{x}_{11} = 0$, $\mathbb{E}|\mathbb{x}_{11}|^2 = 1$.
(iv) $Y_N = p/N \to Y > 0$ *and* $\frac{p}{n+N} \to \frac{yY}{y+Y} \in (0, 1)$.

(v) $\sup_n \mathbb{E}|x_{11}|^4 < \infty$ *and* $\sup_N \mathbb{E}|\mathbb{x}_{11}|^4 < \infty$.

Then with probability 1, $F^{\mathbf{B}_n} \to F$ *weakly, where* F *is a non-random distribution function whose density function is*

$$\begin{cases} \dfrac{\sqrt{((\alpha(1-Y)-1+y)^2+4\alpha)(t_r-t)(t-t_l)}}{2\pi t(1-t)(y(1-t)+\alpha tY)}, & \textit{when } t_l < t < t_r; \\ 0, & \textit{otherwise}, \end{cases}$$

where $t_l, t_r = (\frac{2\alpha-(1-y)[\alpha(1-Y)-1+y]\mp 2\alpha\sqrt{y-yY+Y}}{(\alpha(1-Y)-1+y)^2+4\alpha})$. *In addition, when* $y > 1$, $F(t)$ *has a point mass* $1-1/y$ *at* $t=0$; *when* $Y > 1$, $F(t)$ *has a point mass* $1-1/Y$ *at* $t=1$.

Remark 1.2. Condition $yY/(y+Y) < 1$ is to guarantee that the random matrix $\mathbf{S}_n + \alpha_n \mathbf{T}_N$ is invertible almost surely because $yY/(y+Y) > 1$ ensures that the dimension p could be eventually larger than the number of observations $n+N$. This would imply that $\mathbf{S}_n + \alpha_n \mathbf{T}_N$ is singular. Condition (v) gives us the a.s. bounds of the limit of the smallest and largest eigenvalues, $\lambda_1^{\mathbf{S}_n+\alpha_n\mathbf{T}_N}$ and $\lambda_p^{\mathbf{S}_n+\alpha_n\mathbf{T}_N}$ respectively, of the random matrix $\mathbf{S}_n + \alpha_n \mathbf{T}_N$ since by the definition of $\mathbf{B}_n$ we can rewrite

$$\begin{aligned} &\mathbf{S}_n + \alpha_n \mathbf{T}_N \\ &= \frac{1}{n}\left(\mathbf{X}_i\mathbf{X}_i^* + \frac{\alpha_n n}{N}\mathbb{X}_N\mathbb{X}_N^*\right) \\ &= \frac{1}{n+N}\begin{pmatrix} x_{11} & \cdots & x_{1n} & \mathbb{x}_{11} & \cdots & \mathbb{x}_{1N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{p1} & \cdots & x_{pn} & \mathbb{x}_{p1} & \cdots & \mathbb{x}_{pN} \end{pmatrix} \mathbf{\Gamma} \begin{pmatrix} x_{11} & \cdots & x_{1n} & \mathbb{x}_{11} & \cdots & \mathbb{x}_{1N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{p1} & \cdots & x_{pn} & \mathbb{x}_{p1} & \cdots & \mathbb{x}_{pN} \end{pmatrix}^*. \end{aligned}$$

Here

$$\mathbf{\Gamma} = \begin{pmatrix} \begin{pmatrix} \frac{n+N}{n} & & \\ & \ddots & \\ & & \frac{n+N}{n} \end{pmatrix}_{n\times n} & \\ & \begin{pmatrix} \frac{(n+N)\alpha_n}{N} & & \\ & \ddots & \\ & & \frac{(n+N)\alpha_n}{N} \end{pmatrix}_{N\times N} \end{pmatrix}_{(n+N)\times(n+N)}$$

is a diagonal matrix. Thus under (v), for any $\varepsilon > 0$ and any $l > 0$, there exist two positive constants $\nu_1 = \min\{1, \alpha Y/y\} \cdot (1+y/Y)(1-\sqrt{\frac{yY}{y+Y}})^2$ and $\nu_2 = \max\{1, \alpha Y/y\} \cdot (1+y/Y)(1+$

$\sqrt{\frac{yY}{y+Y}})^2$ such that almost surely

$$\lim_{p,n,N\to\infty} \lambda_1^{\mathbf{S}_n+\alpha_n\mathbf{T}_N} \geq \nu_1, \qquad \lim_{p,n,N\to\infty} \lambda_p^{\mathbf{S}_n+\alpha_n\mathbf{T}_N} \leq \nu_2 \tag{1.3}$$

and

$$\mathbb{P}\bigl(\lambda_1^{\mathbf{S}_n+\alpha_n\mathbf{T}_N} < \nu_1 - \varepsilon\bigr) = \mathrm{o}\bigl(n^{-l}\bigr), \qquad \mathbb{P}\bigl(\lambda_p^{\mathbf{S}_n+\alpha_n\mathbf{T}_N} > \nu_2 + \varepsilon\bigr) = \mathrm{o}\bigl(n^{-l}\bigr). \tag{1.4}$$

One may refer to [2] for the proof of (1.3) and (1.4).

Remark 1.3. Under the assumptions (i) and (ii) in Theorem 1.1, it is proved that the ESDF of the sequence $\{\mathbf{S}_n\}$ has a non-random limit which is known as the Marchenko–Pastur (M–P) distribution [2,10]. Yin [15] and Silverstein [12] investigated the LSDF of the sequence $\{\mathbf{S}_n\mathbf{T}_N\}$ assuming (i)–(iii) of Theorem 1.1. If $\mathbf{T}_N$ is invertible, Bai *et al.* [5] gave the LSDF of the sequence $\{\mathbf{S}_n\mathbf{T}_N^{-1}\}$.

Remark 1.4. If $\max\{y, Y\} < 1$, by (v) we know that at least one of the matrices $\mathbf{S}_n$ and $\mathbf{T}_N$ is invertible a.s. Without loss of generality, we assume $Y < 1$. So $\mathbf{T}_N$ is invertible a.s. Then we have

$$\mathbf{B}_n = \mathbf{S}_n\mathbf{T}_N^{-1}\bigl(\mathbf{S}_n\mathbf{T}_N^{-1} + \alpha_n\mathbf{I}\bigr)^{-1}, \tag{1.5}$$

which is a function of $\mathbf{S}_n\mathbf{T}_N^{-1}$. Via $\tilde{t} = \alpha_n t/(1-t)$ we can recover Theorem 5.3 in [5] from our Theorem 1.1 directly. Thus our Theorem 1.1 includes Theorem 5.3 in [5] as a special case.

Remark 1.5. From the density function in Theorem 1.1, we can find that the condition $\frac{p}{n+N} \to \frac{yY}{y+Y} \in (0,1)$ is necessary, which is to make sense of $\sqrt{y+Y-yY}$.

For the purpose of multivariate inference, it is of interest to know the limiting distribution of these LSS (1.2). Thus, we will give the central limit theorems (CLT) of LSS of Beta matrices. In order to present this result, we need more notation. Denote

$$\mathfrak{B}_n(x) = p\bigl(F^{\mathbf{B}_n}(x) - F_0(x)\bigr),$$

where F_0 is the limit distribution of $F^{\mathbf{B}_n}$ with α, y, Y replaced by α_n, y_n, Y_N, respectively. For any function of bounded variation G on the real line, its Stieltjes transform is defined by

$$s_G(z) = \int \frac{1}{\lambda - z}\,\mathrm{d}G(\lambda), \qquad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}.$$

Then we have the following theorem.

Theorem 1.6. *In addition to the conditions* (i)–(iv) *in Theorem* 1.1, *we further assume that*:

1. $\mathbb{E}x_{11}^2 = \mathbb{E}\mathbb{x}_{11}^2 = \mathfrak{t}$, $\mathbb{E}|x_{11}|^4 = \mathfrak{m}_x$, $\mathbb{E}|\mathbb{x}_{11}|^4 = \mathfrak{m}_{\mathbb{x}}$ *and* $\max_{p,n,N}\{\mathfrak{m}_x, \mathfrak{m}_{\mathbb{x}}\} < \infty$, *where* $\mathfrak{t} = 0$, *when both* $\mathbf{X}_n$ *and* $\mathbb{X}_N$ *are complex valued*, *and* $\mathfrak{t} = 1$ *if both real*.
2. *Let* $f_1, \dots, f_k$ *be functions analytic on an open region containing the interval* $[c_l, c_r]$ *where* $c_l = \nu_2^{-1}(1-\sqrt{y})^2$, $c_r = 1 - \alpha\nu_2^{-1}(1-\sqrt{Y})^2$, *and* ν_2 *is defined in Remark* 1.2.

Then, as $\min(n, N, p) \to \infty$, *the random vector*

$$\left(\int f_i \,\mathrm{d}\mathfrak{B}_n(x)\right), \qquad i = 1, \dots, k,$$

converges weakly to a Gaussian vector $(G_{f_1}, \dots, G_{f_k})$ *with mean functions*

$$\begin{aligned}
\mathbb{E}G_{f_i} &= \frac{\mathfrak{t}}{4\pi i}\oint f_i\left(\frac{z}{\alpha+z}\right)\mathrm{d}\log\left(\frac{(1-Y)\dddot{s}^2(z) + 2\dddot{s}(z) + 1 - y}{(1-Y)\dddot{s}^2 + 2\dddot{s}(z) + 1}\right) \\
&\quad + \frac{\mathfrak{t}}{4\pi i}\oint f_i\left(\frac{z}{\alpha+z}\right)\mathrm{d}\log\bigl(1 - Y\dddot{s}^2(z)\bigl(1+\dddot{s}(z)\bigr)^{-2}\bigr) \\
&\quad + \frac{\mathfrak{m}_x - \mathfrak{t} - 2}{2\pi i}\oint y f_i\left(\frac{z}{\alpha+z}\right)\bigl(\dddot{s}(z)+1\bigr)^{-3}\,\mathrm{d}\dddot{s}(z) \\
&\quad + \frac{\mathfrak{m}_{\mathbb{x}} - \mathfrak{t} - 2}{4\pi i}\oint f_i\left(\frac{z}{\alpha+z}\right)\bigl(1 - Y\dddot{s}^2(z)\bigl(1+\dddot{s}(z)\bigr)^{-2}\bigr)\,\mathrm{d}\log\bigl(1 - Y\dddot{s}^2(z)\bigl(1+\dddot{s}(z)\bigr)^{-2}\bigr)
\end{aligned}$$

and covariance functions

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}(G_{f_i}, G_{f_j}) \\
&\quad = -\frac{\mathfrak{t}+1}{4\pi^2}\oint\oint \frac{f_i(z_1/(\alpha+z_1))f_j(z_2/(\alpha+z_2))\,\mathrm{d}\dddot{s}(z_1)\,\mathrm{d}\dddot{s}(z_2)}{(\dddot{s}(z_1) - \dddot{s}(z_2))^2} \\
&\qquad - \frac{y(\mathfrak{m}_x - \mathfrak{t} - 2) + Y(\mathfrak{m}_{\mathbb{x}} - \mathfrak{t} - 2)}{4\pi^2}\oint\oint \frac{f_i(z_1/(\alpha+z_1))f_j(z_2/(\alpha+z_2))\,\mathrm{d}\dddot{s}(z_1)\,\mathrm{d}\dddot{s}(z_2)}{(\dddot{s}(z_1)+1)^2(\dddot{s}(z_2)+1)^2},
\end{aligned}$$

where

$$\begin{aligned}
s(z) &= \frac{(1+y)(1-z) - \alpha z(1-Y) + \sqrt{((1-y)(1-z) + \alpha z(1-Y))^2 - 4\alpha z(1-z)}}{2z(1-z)(y(1-z) + \alpha z Y)} - \frac{1}{z}, \\
\dot{s}(z) &= \frac{\alpha}{(\alpha+z)^2}s\left(\frac{z}{\alpha+z}\right) - \frac{1}{\alpha+z}, \qquad \ddot{s}(z) = -z^{-1}(1-y) + y\dot{s}(z), \\
s_{\mathrm{mp}}^Y(z) &= \frac{1 - Y - z + \sqrt{(z-1-Y)^2 - 4Y}}{2Yz}, \qquad \dddot{s}(z) = Y s_{\mathrm{mp}}^Y\bigl(-\ddot{s}(z)\bigr) + \bigl(\ddot{s}(z)\bigr)^{-1}(1-Y).
\end{aligned}$$

All the above contour integrals can be evaluated on any contour enclosing the interval $[\frac{\alpha c_l}{1-c_l}, \frac{\alpha c_r}{1-c_r}]$.

Remark 1.7. Actually, this result should be right under the condition that f_i is analytic (or continuously differentiable) on an open region containing the interval $[t_l, t_r]$. However its proof is more difficult at the current stage because we do not have the following results of Beta matrices: the exact separation of eigenvalues, the limit of the smallest and the largest eigenvalues and the convergence rate of the ESDF.

Remark 1.8. In this theorem, the notions $s(z)$ and $s_{\mathrm{mp}}^{Y}(z)$ are the Stieltjes transforms of the LSDFs of $\mathbf{B}_n$ and $\mathbf{T}_N$ respectively. If $Y < 1$, Zheng in [16] established the CLT of the LSS of F matrix $\mathbf{S}_n\mathbf{T}_N^{-1}$ whose proof is based on [4]. It is apparent that our Theorem 1.6 covers Zheng's result. In addition, notice that the conclusions in Theorem 1.6 and Theorem 4.1 in [16] have the same form. The reason is that, by calculation we can easily get

$$\dot{s}(z) = \frac{1}{zy} - \frac{1}{z} - \frac{y(z(1-Y)+1-y)+2zY - y\sqrt{((1-y)+z(1-Y))^2 - 4z}}{2z(yz+Y)}$$

which has the same expression of the Stieltjes transform of the LSDF of F matrices (see (2.6) in [16]). Here we want to remind the reader that, when we use the last formula to calculate the density function, that is, calculating $\pi^{-1}\lim_{z\downarrow x+i0}\Im\dot{s}(z)$, we can find that the condition $Y < 1$ is not needed but $y + Y > yY$ is necessary (see page 79 in [2] for more details).

Remark 1.9. If $\{x_{ij}\}$ and $\{\mathbb{x}_{ij}\}$ are independent standard normal random variables and $p < \max\{n, N\}$, Beta matrices can be seen as Beta–Jacobi ensemble with some parameter β. Some related results about this ensemble can be found in [8] and the references therein.

This paper is organized as follows: In Section 2, we present the proof of Theorem 1.1. Theorem 1.6 is proved in Section 3 and Section 4. Some technical lemmas are given in Section 5.

2. Proof of Theorem 1.1

In this section, we will give the proof of Theorem 1.1. The main tool we use here is the Stieltjes transform. Its function can be explained by the following two lemmas.

Lemma 2.1 (Lemma 1.1 in [6]). *For any random matrix $\mathbf{A}_n$, let $F^{\mathbf{A}_n}$ denote the ESDF of $\mathbf{A}_n$ and $s_{F^{\mathbf{A}_n}}(z)$ its Stieltjes transform. Then, if $F^{\mathbf{A}_n}$ is tight with probability one and for each $z \in \mathbb{C}^+$, $s_{F^{\mathbf{A}_n}}(z)$ converges almost surely to a non-random limit $s_F(z)$ as $n \to \infty$, then there exists a non-random probability distribution F taking $s_F(z)$ as its Stieltjes transform such that with probability one, as $n \to \infty$, $F^{\mathbf{A}_n}$ converges weakly to F.*

Lemma 2.2 (Theorem 2.1 in [14]). *Let G be a function of bounded variation and $x_0 \in \mathbb{R}$. Suppose that $\lim_{z\in\mathbb{C}^+\to x_0}\Im s_G(z)$ exists. Its limit is denoted by $\Im s_G(x_0)$. Then G is differentiable at x_0, and its derivative is $\pi^{-1}\Im s_G(x_0)$.*

Theorem 1.1 follows from the following Theorem 2.3.

Theorem 2.3. *Under the conditions* (i) *and* (ii) *in Theorem 1.1, we assume that*:

(1) $\{\mathbf{A}_p\}$ *is a sequence of* $p \times p$ *Hermitian matrices with uniformly bounded spectral norm in* n *with probability one and the ESDFs of* $\{\mathbf{A}_p\}$ *almost surely tend to a non-random limit* $F^{\mathbf{A}}$ *as* $p \to \infty$.
(2) *The smallest eigenvalue of matrices* $\{\mathbf{S}_n + \alpha_n \mathbf{A}_p\}$ *almost surely tends to a positive value as* $n \to \infty$ *and* $p \to \infty$.

Then we have $F^{\underline{\mathbf{B}}_n} \xrightarrow{a.s.} \underline{F}$, *where* $\underline{\mathbf{B}}_n = \mathbf{S}_n(\mathbf{S}_n + \alpha_n \mathbf{A}_p)^{-1}$ *and* $\underline{F}$ *is a non-random distribution function whose Stieltjes transform* $\underline{s} = \underline{s}(z) = s_{\underline{F}}(z)$ *satisfies*

$$\underline{s} = \int \frac{(1 - y(1-z)(z\underline{s}+1)) + \alpha t}{(1-z)(1-y(1-z)(z\underline{s}+1)) - \alpha z t} \, \mathrm{d}F^{\mathbf{A}}(t), \tag{2.1}$$

and in the set $\{\underline{s} : \underline{s} \in \mathbb{C}^+\}$ *the solution to* (2.1) *is unique.*

By Lemma 2.1, we know that to prove Theorem 2.3 we just need to prove three conclusions: (1) $\{F^{\underline{\mathbf{B}}_n}\}$ is tight a.s. (2) $s_{F^{\underline{\mathbf{B}}_n}} \xrightarrow{\text{a.s.}} s$ with s satisfying (2.1). (3) The solution to (2.1) is unique in $\mathbb{C}^+$. Now we prove Theorem 2.3 step by step.

2.1. Proof of Theorem 2.3

Step 1: Applying Lemma 5.2 directly, we have for any $x_1, x_2 \geq 0$

$$\begin{aligned} F^{\underline{\mathbf{B}}_n}\{(x_1 x_2, \infty)\} &\leq F^{\mathbf{S}_n}\{(x_1, \infty)\} + F^{(\mathbf{S}_n + \alpha_n \mathbf{A}_p)^{-1}}\{(x_2, \infty)\} \\ &= F^{\mathbf{S}_n}\{(x_1, \infty)\} + F^{(\mathbf{S}_n + \alpha_n \mathbf{A}_p)}\{(0, 1/x_2)\}. \end{aligned} \tag{2.2}$$

It is known that, under the assumptions of Theorem 2.3, with probability one $F^{\mathbf{S}_n}$ tends to the M–P distribution F^y_{mp}, which has a density function

$$f^y_{\text{mp}}(x) = \begin{cases} \dfrac{1}{2\pi x y}\sqrt{(b-x)(x-a)}, & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

and has a point mass $1 - 1/y$ at the origin if $y > 1$, where $a = (1-\sqrt{y})^2$ and $b = (1+\sqrt{y})^2$. Thus $\{F^{\mathbf{S}_n}\}$ is tight almost surely, that is, the first term on the right-hand side of (2.2) can be arbitrarily small by choosing x_1 large.

On the other hand, by the second assumption of Theorem 1.1, the second term on the right-hand side of (2.2) can be arbitrarily small as n is large, provided that $1/x_2$ is smaller than the smallest eigenvalue of the matrices $\{\mathbf{S}_n + \alpha_n \mathbf{A}_p\}$. Thus $\{F^{\underline{\mathbf{B}}_n}\}$ is tight almost surely.

Step 2: Recalling the definition of Stieltjes transform we have that for $z \in \mathbb{C}^+$

$$s_{F^{\underline{\mathbf{B}}_n}}(z) = \frac{1}{p}\sum_{i=1}^{p} \frac{1}{\lambda_i^{\underline{\mathbf{B}}_n} - z} = \frac{1}{p}\operatorname{tr}(\underline{\mathbf{B}}_n - z\mathbf{I})^{-1}. \tag{2.4}$$

Here we have used the fact that $\underline{\mathbf{B}}_n$ has the same eigenvalues as

$$\mathbf{S}_n^{1/2}(\mathbf{S}_n + \alpha_n \mathbf{A}_p)^{-1}\mathbf{S}_n^{1/2}.$$

Denote $\underline{\mathbf{B}}_\varepsilon = \mathbf{S}_n(\mathbf{S}_n + \alpha_n \mathbf{A}_p + \varepsilon \mathbf{I})^{-1}$ with small $\varepsilon > 0$. From Lemma 5.3, we have

$$L^3\bigl(F^{\underline{\mathbf{B}}_n}, F^{\underline{\mathbf{B}}_\varepsilon}\bigr) \le \frac{1}{n}\operatorname{tr}(\underline{\mathbf{B}}_n - \underline{\mathbf{B}}_\varepsilon)(\underline{\mathbf{B}}_n - \underline{\mathbf{B}}_\varepsilon)^*.$$

By the fact

$$\begin{aligned}\underline{\mathbf{B}}_n - \underline{\mathbf{B}}_\varepsilon &= \varepsilon \mathbf{S}_n^{1/2}(\mathbf{S}_n + \alpha_n \mathbf{A}_p)^{-1/2}(\mathbf{S}_n + \alpha_n \mathbf{A}_p + \varepsilon \mathbf{I})^{-1}(\mathbf{S}_n + \alpha_n \mathbf{A}_p)^{-1/2}\mathbf{S}_n^{1/2} \\ &\le \varepsilon(\mathbf{S}_n + \alpha_n \mathbf{A}_p + \varepsilon \mathbf{I})^{-1}\end{aligned}$$

together with condition (2) in Theorem 2.3, we obtain almost surely that $L^3(F^{\underline{\mathbf{B}}_n}, F^{\underline{\mathbf{B}}_\varepsilon}) \le C\varepsilon^2$, which implies $\lim_{\varepsilon\to 0}\lim_{n\to\infty} L(F^{\underline{\mathbf{B}}_n}, F^{\underline{\mathbf{B}}_\varepsilon}) = 0$.

Next, we consider the LSDF of $\underline{\mathbf{B}}_\varepsilon$. Noticing that the matrix $\alpha_n \mathbf{A}_p + \varepsilon \mathbf{I}$ is invertible for any $\varepsilon > 0$, we have

$$\underline{\mathbf{B}}_\varepsilon = \mathbf{I} - (\widehat{\underline{\mathbf{B}}}_\varepsilon + \mathbf{I})^{-1},$$

where $\widehat{\underline{\mathbf{B}}}_\varepsilon = \mathbf{S}_n(\alpha_n \mathbf{A}_p + \varepsilon \mathbf{I})^{-1}$. Thus, we get that $F^{\underline{\mathbf{B}}_\varepsilon}(x) = F^{\widehat{\underline{\mathbf{B}}}_\varepsilon}(\frac{1}{1-x} - 1)$ and

$$s_{F^{\underline{\mathbf{B}}_\varepsilon}}(z) = \frac{1}{1-z} + \frac{1}{(1-z)^2} s_{F^{\hat{\underline{\mathbf{B}}}_\varepsilon}}\left(\frac{z}{1-z}\right). \tag{2.5}$$

Silverstein in [12] derived that for any $z \in \mathbb{C}^+$, the Stieltjes transform of the ESDF of $\widehat{\underline{\mathbf{B}}}_\varepsilon$ has a non-random limit, denoted by $s_{\hat{\varepsilon}}(z)$, which satisfies the equation

$$s_{\hat{\varepsilon}}(z) = \int \frac{1}{t(1 - y - yzs_{\hat{\varepsilon}}(z)) - z}\,\mathrm{d}F_\varepsilon^{\mathbf{A}}(t),$$

where $F_\varepsilon^{\mathbf{A}}$ is the LSDF of $(\alpha_n \mathbf{A}_p + \varepsilon \mathbf{I})^{-1}$. Note that $\Im(z/(1-z)) = |1-z|^{-2}\Im z > 0$. Thus by (2.5) we get that almost surely $s_{F^{\underline{\mathbf{B}}_\varepsilon}}(z)$ tends to a non-random limit, denoted by $\underline{s}_\varepsilon(z)$, which satisfies

$$\begin{aligned}&(1-z)^2\underline{s}_\varepsilon(z) - (1-z) \\ &\quad= \int \frac{1}{t(1 - y - y(z/(1-z))((1-z)^2\underline{s}_\varepsilon(z) - (1-z))) - z/(1-z)}\,\mathrm{d}F_\varepsilon^{\mathbf{A}}(t).\end{aligned}$$

By definition of $F_\varepsilon^{\mathbf{A}}$ and $F^{\mathbf{A}}$, we have that

$$\mathrm{d}F_\varepsilon^{\mathbf{A}}(t) = -\mathrm{d}F^{\mathbf{A}}\left(\frac{t^{-1} - \varepsilon}{\alpha}\right).$$

Therefore letting $\varepsilon \to 0$, we have

$$\underline{s} = \int \frac{(1 - y(1 - z)(z\underline{s} + 1)) + \alpha t}{(1 - z)(1 - y(1 - z)(z\underline{s} + 1)) - \alpha z t} \, \mathrm{d}F^{\mathbf{A}}(t). \tag{2.6}$$

Step 3: From Lemma 2.1, we conclude that there exists a distribution function G with support $\Psi_G \subset [0, 1]$ satisfying for any $z \in \mathbb{C}^+$,

$$\underline{s}(z) = \int_{\Psi_G} \frac{1}{x - z} \, \mathrm{d}G(x). \tag{2.7}$$

Noticing that $\Im z(\alpha + z)^{-1} = \alpha|\alpha + z|^{-2}\Im z > 0$, we infer from (2.7) that

$$\begin{aligned} \frac{\alpha}{(\alpha + z)^2} \underline{s}\left(\frac{z}{\alpha + z}\right) - \frac{1}{\alpha + z} &= \frac{\alpha}{(\alpha + z)^2} \int_{\Psi_G} \frac{1}{x - z/(\alpha + z)} \, \mathrm{d}G(x) - \frac{1}{\alpha + z} \\ &= \int_{\Psi_G} \frac{1 - x}{\alpha x - z(1 - x)} \, \mathrm{d}G(x) = \int_0^\infty \frac{1}{x - z} \, \mathrm{d}G\left(\frac{x}{\alpha + x}\right). \end{aligned}$$

Thus

$$\dot{\underline{s}} = \dot{\underline{s}}(z) = \frac{\alpha}{(\alpha + z)^2} \underline{s}\left(\frac{z}{\alpha + z}\right) - \frac{1}{\alpha + z} \tag{2.8}$$

is a Stieltjes transform of the distribution function $G(\frac{x}{\alpha+x})$ with $x \in [0, \infty)$. Notice that even if $G(x)$ has a point mass at $x = 1$, we have $\frac{1-x}{\alpha x - z(1-x)} = 0$. Thus, (2.6) can be represented as

$$\dot{\underline{s}}(z) = \int_{\mathbb{R}^+} \frac{1}{t(1 - y - yz\dot{\underline{s}}(z)) - z} \, \mathrm{d}\left(1 - F^{\mathbf{A}}\left(\frac{1}{t}\right)\right),$$

where $\mathbb{R}^+ = \{t : t \in \mathbb{R}, t > 0\}$. It is shown that the solution of the last equation is unique in $\mathbb{C}^+$ (see [12]). Thus, we obtain that (2.6) has a unique solution in $\mathbb{C}^+$, which completes the proof of Theorem 2.3.

2.2. Proof of Theorem 1.1

Using Theorem 2.3 and Remark 1.2, we know that the Stieltjes transform of F is the unique solution in $\mathbb{C}^+$ to the equation

$$s = \int \frac{(1 - y(1 - z)(zs + 1)) + \alpha t}{(1 - z)(1 - y(1 - z)(zs + 1)) - \alpha z t} \, \mathrm{d}F^Y_{\mathrm{mp}}(t). \tag{2.9}$$

Here F^Y_{mp} is the limit of $F^{\mathbf{T}_N}$ which is also the M–P distribution. After elementary calculations, we may represent the last equation as

$$s = -\frac{1}{z} - \frac{\varpi}{\alpha z^2} \int \frac{1}{t - ((1 - z)\varpi/(\alpha z))} \, \mathrm{d}F^Y_{\mathrm{mp}}(t), \tag{2.10}$$

where $\varpi = 1 - y(1-z)(zs+1)$. Recalling (2.8), we have that

$$s(z) = \frac{1}{1-z} + \frac{\alpha}{(1-z)^2}\dot{s}\left(\frac{\alpha z}{1-z}\right)$$

and

$$\varpi = 1 - y(1-z)(zs+1) = 1 - y - \frac{\alpha y z}{(1-z)}\dot{s}\left(\frac{\alpha z}{1-z}\right),$$

which implies

$$\frac{(1-z)\varpi}{\alpha z} = \frac{(1-z)(1-y)}{\alpha z} - y\dot{s}\left(\frac{\alpha z}{1-z}\right).$$

Noticing that $\Im \frac{\alpha z}{1-z} > 0$, we have $\Im \frac{(1-z)\varpi}{\alpha z} < 0$ and

$$\int \frac{1}{t - ((1-z)\varpi/(\alpha z))}\,\mathrm{d}F_{\mathrm{mp}}^{Y}(t) = \overline{s_{\mathrm{mp}}^{Y}\left(\frac{(1-\bar{z})\varpi(\bar{z})}{\alpha \bar{z}}\right)},$$

where s_{mp}^{Y} is the Stieltjes transform of the M–P distribution F_{mp}^{Y}. Since

$$s_{\mathrm{mp}}^{Y}(z) = \frac{1 - Y - z + \sqrt{(z-1-Y)^2 - 4Y}}{2Yz},$$

the equation (2.10) implies

$$s = -\frac{1}{z} - \frac{\varpi}{\alpha z^2}\left(\frac{1 - Y - ((1-z)\varpi/(\alpha z)) + \sqrt{(((1-z)\varpi/(\alpha z)) - 1 - Y)^2 - 4Y}}{2Y((1-z)\varpi/(\alpha z))}\right),$$

where, and throughout this section, the square-root of a complex number is specified as the one with positive imaginary part. The solution to this equation is

$$s(z) = \frac{(1+y)(1-z) - \alpha z(1-Y) + \sqrt{((1-y)(1-z) + \alpha z(1-Y))^2 - 4\alpha z(1-z)}}{2z(1-z)(y(1-z) + \alpha z Y)} - \frac{1}{z}.$$

Now using Lemma 2.2 and letting $z \downarrow x + \mathrm{i}0$, $\pi^{-1}\Im s(z)$ tends to the density function of the LSDF of $\mathbf{B}_n$. Thus, the density function of the LSDF of $\mathbf{B}_n$ is

$$\begin{cases} \dfrac{\sqrt{4\alpha x(1-x) - ((1-y)(1-x) + \alpha x(1-Y))^2}}{2\pi x(1-x)(y(1-x) + \alpha x Y)}, \\ \qquad \text{if } 4\alpha x(1-x) - \big((1-y)(1-x) + \alpha x(1-Y)\big)^2 > 0; \\ 0, \qquad \text{otherwise.} \end{cases}$$

Or equivalently,

$$\begin{cases} \dfrac{\sqrt{((\alpha(1-Y)-1+y)^2+4\alpha)(x_r-x)(x-x_l)}}{2\pi x(1-x)(y(1-x)+\alpha xY)}, & \text{if } x_l < x < x_r; \\ 0, & \text{otherwise,} \end{cases}$$

where $x_l, x_r = (\frac{2\alpha-(1-y)[\alpha(1-Y)-1+y]\mp 2\alpha\sqrt{y-yY+Y}}{(\alpha(1-Y)-1+y)^2+4\alpha})$. Now we determine the possible atom at 0 and 1. When $z \to 0$ with $\Im z > 0$, we have

$$\begin{aligned} &\Im\big[\big((1-y)(1-z)+\alpha z(1-Y)\big)^2 - 4\alpha z(1-z)\big] \\ &\quad = 2\Im z\big\{\big[(1-Y)\alpha - 1 + y\big]\big[(1-y)(1-\Re z)+\alpha(1-Y)\Re z\big] - 2\alpha(1-2\Re z)\big\} < 0. \end{aligned}$$

By the fact that the real part of $\sqrt{g(z)}$ has the same sign as that of the imaginary part of $g(z)$, we obtain that $\Re\sqrt{((1-y)(1-z)+\alpha z(1-Y))^2 - 4\alpha z(1-z)} < 0$. Thus

$$\sqrt{\big((1-y)(1-z)+\alpha z(1-Y)\big)^2 - 4\alpha z(1-z)} \to -|1-y|.$$

Consequently,

$$F\{0\} = -\lim_{z\to 0} zs(z) = \frac{|1-y|-1-y}{2y} + 1 = \begin{cases} \dfrac{y-1}{y}, & \text{if } y > 1; \\ 0, & \text{otherwise.} \end{cases}$$

When $z \to 1$ with $\Im z > 0$, we have

$$\begin{aligned} &\Im\big[\big((1-y)(1-z)+\alpha z(1-Y)\big)^2 - 4\alpha z(1-z)\big] \\ &\quad = 2\Im z\big\{\big[(1-Y)\alpha - 1 + y\big]\big[(1-y)(1-\Re z)+\alpha(1-Y)\Re z\big] - 2\alpha(1-2\Re z)\big\} > 0. \end{aligned}$$

Hence, we get $\Re\sqrt{((1-y)(1-z)+\alpha z(1-Y))^2 - 4\alpha z(1-z)} > 0$. Thus,

$$\sqrt{\big((1-y)(1-z)+\alpha z(1-Y)\big)^2 - 4\alpha z(1-z)} \to \alpha|1-Y|.$$

Consequently,

$$F\{1\} = -\lim_{z\to 1}(z-1)s(z) = \frac{|1-Y|-(1-Y)}{2Y} = \begin{cases} \dfrac{Y-1}{Y}, & \text{if } Y > 1; \\ 0, & \text{otherwise.} \end{cases}$$

Then the proof of Theorem 1.1 is complete.

3. Framework of proving Theorem 1.6

In this section, we will give the proof of Theorem 1.6. Recall the definition of the Stieltjes transform of a distribution function $G(x)$. Now we extend the Stieltjes transform to the whole complex plane except the interval $[c_l, c_r]$ analytically. Since every $f_k(x)$ is analytic on an open region containing the interval $[c_l, c_r]$, we assume that the analytic region contains the contour $\mathcal{C} = \{z \in \mathbb{C} : \Re z \in [c_l - \theta, c_r + \theta], \Im z = \pm\theta\} \cup \{z \in \mathbb{C} : \Re z \in \{c_l - \theta, c_r + \theta\}, \Im z \in [-\theta, \theta]\}$. Here θ can be small enough. By Cauchy's integral formula

$$f_k(x) = \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}} \frac{f_k(z)}{z - x} \, \mathrm{d}z,$$

we have for $l \geq 1$ and complex constants $a_1, \ldots, a_l$,

$$\sum_{k=1}^{l} a_k p\biggl(\int f_k(x)\, \mathrm{d}F^{\mathbf{B}_n}(x) - \int f_k(x)\, \mathrm{d}F_0(x)\biggr) = -\sum_{k=1}^{l} \frac{a_k}{2\pi \mathrm{i}} \oint_{\mathcal{C}} f_k(z) S_n(z)\, \mathrm{d}z, \tag{3.1}$$

where $S_n(z) = p(s_n(z) - s_0(z))$ and $s_0(z)$ is the Stieltjes transform of F with constants y and Y replaced by $y_n = p/n$ and $Y_n = p/N$. We remind the readers to notice that the above equality may not be correct when some eigenvalues of $\mathbf{B}_n$ fall outside the contour. However, by Remark 1.2, Lemma 5.7 and the exact separation theorem in [3], we know for $y > 1$ (or $Y > 1$) and sufficiently large n (or N), the mass at the origin (one) of $F^{\mathbf{B}_n}$ will coincide exactly with that of F_0 and with overwhelming probability all the other eigenvalues of $\mathbf{B}_n$ fall in $[c_l - \theta, c_r + \theta]$. Thus to prove Theorem 1.6, it suffices for us to derive the limiting distribution of (3.1).

Write

$$S_n(z) = p\bigl(s_n(z) - s_{N0}(z)\bigr) + p\bigl(s_{N0}(z) - s_0(z)\bigr) := S_{n1} + S_{n2},$$

where $s_{N0}(z)$ is the unique root of the equation

$$s_{N0} = \int \frac{(1 - y_n(1-z)(z s_{N0}(z) + 1)) + \alpha_n t}{(1-z)(1 - y_n(1-z)(z s_{N0} + 1)) - \alpha_n z t} \, \mathrm{d}F^{\mathbf{T}_N}(t)$$

in the set $\{s_{N0}(z) \in \mathbb{C}^+\}$. Using the notation $\dot{s}_{N0} = \dot{s}_{N0}(z) = \frac{\alpha_n}{(\alpha_n+z)^2} s_{N0}(\frac{z}{\alpha_n+z}) - \frac{1}{\alpha_n+z}$, $\dot{s}_0 = \dot{s}_0(z) = \frac{\alpha_n}{(\alpha_n+z)^2} s_0(\frac{z}{\alpha_n+z}) - \frac{1}{\alpha_n+z}$, $\ddot{s}_{N0}(z) = -z^{-1}(1 - y_n) + y_n \dot{s}_{N0}(z)$ and $\ddot{s}_0(z) = -z^{-1}(1 - y_n) + y_n \dot{s}_0(z)$ we have

$$z = -\frac{1}{\ddot{s}_{N0}} + y_n \int \frac{\mathrm{d}F^{\mathbf{T}_N}(t)}{t + \ddot{s}_{N0}} \quad \text{and} \quad z = -\frac{1}{\ddot{s}_0} + y_n \int \frac{\mathrm{d}F_{\mathrm{mp}}^{Y_N}(t)}{t + \ddot{s}_0}.$$

Making difference of the two identities above yields that

$$\frac{\ddot{s}_0 - \ddot{s}_{N0}}{\ddot{s}_0 \ddot{s}_{N0}} = y_n \int \frac{(\ddot{s}_0 - \ddot{s}_{N0})\, \mathrm{d}F^{\mathbf{T}_N}(t)}{(t + \ddot{s}_{N0})(t + \ddot{s}_0)} + y_n \int \frac{\mathrm{d}F^{\mathbf{T}_N}(t) - \mathrm{d}F_{\mathrm{mp}}^{Y_N}(t)}{t + \ddot{s}_0}.$$

Then we get

$$\ddot{s}_0 - \ddot{s}_{N0} = y_n \ddot{s}_0 \ddot{s}_{N0} \int \frac{\mathrm{d}F^{\mathbf{T}_N}(t) - \mathrm{d}F_{\mathrm{mp}}^{Y_N}(t)}{t + \ddot{s}_0} \biggl(1 - y_n \ddot{s}_0 \ddot{s}_{N0} \int \frac{\mathrm{d}F^{\mathbf{T}_N}(t)}{(t + \ddot{s}_{N0})(t + \ddot{s}_0)}\biggr)^{-1}. \tag{3.2}$$

Let $s^{\mathbf{T}_N}$ be the Stieltjes transforms of $F^{\mathbf{T}_N}$ and then from (6.32) in [16] we have the conclusion that

$$p\int \frac{\mathrm{d}F^{\mathbf{T}_N}(t) - \mathrm{d}F^{Y_N}_{\mathrm{mp}}(t)}{t+\ddot{s}_0} = p\big(s^{\mathbf{T}_N}(-\ddot{s}_0) - s^{Y_N}_{\mathrm{mp}}(-\ddot{s}_0)\big)$$

converges weakly to a Gaussian process Φ_1 on $\mathcal{C}$ with mean function

$$\mathbb{E}\Phi_1(z) = \mathfrak{t}\frac{Y[\ddot{s}(z)]^3[1+\ddot{s}(z)]^{-3}}{\{1 - Y\ddot{s}(z)/[1+\ddot{s}(z)]^2\}^2} + (\mathfrak{m}_{\mathbb{X}} - \mathfrak{t} - 2)\frac{Y[\ddot{s}(z)]^3[1+\ddot{s}(z)]^{-3}}{1 - Y[\ddot{s}(z)]^2/[1+\ddot{s}(z)]^2} \tag{3.3}$$

and covariance function

$$\begin{aligned}\mathbb{C}\mathrm{ov}\big(\Phi_1(z_1), \Phi_1(z_2)\big) = (\mathfrak{t}+1)\bigg(&\frac{(\ddot{s}(z_1))'(\ddot{s}(z_2))'}{[\ddot{s}(z_1) - \ddot{s}(z_2)]^2} - \frac{1}{(\ddot{s}(z_1) - \ddot{s}(z_2))^2}\bigg)\\ &+ (\mathfrak{m}_{\mathbb{X}} - \mathfrak{t} - 2)\frac{Y(\ddot{s}(z_1))'(\ddot{s}(z_2))'}{[1+\ddot{s}(z_1)]^2[1+\ddot{s}(z_2)]^2},\end{aligned} \tag{3.4}$$

where $\ddot{s}(z) = \ddot{s}^Y_{\mathrm{mp}}(-\ddot{s}(z))$, $\ddot{s}^Y_{\mathrm{mp}}(z) = -z^{-1}(1-Y) + Ys^Y_{\mathrm{mp}}(z)$ and $(\ddot{s}(z_i))' = \frac{\mathrm{d}}{\mathrm{d}z}\ddot{s}^Y_{\mathrm{mp}}(z)|_{z=-\ddot{s}(z_i)}$, $i = 1, 2$. And $\{S_{n2}(\cdot)\}$ forms a tight sequence on $\mathcal{C}$ and $S_{n2}(\frac{z}{\alpha+z})$ converges weakly to a Gaussian process $-(\alpha+z)^2\ddot{s}'(z)\Phi_1(z)$ with mean function

$$\mathbb{E}\big(-(1+z)^2\ddot{s}'(z)\Phi_1(z)\big) = -(\alpha+z)^2\ddot{s}'(z)\cdot(3.3)$$

and covariance function

$$\begin{aligned}&\mathbb{C}\mathrm{ov}\big(-(\alpha+z_1)^2\ddot{s}'(z_1)\Phi_1(z_1), -(\alpha+z_2)^2\ddot{s}'(z_2)\Phi_1(z_2)\big)\\ &\quad= (\alpha+z_1)^2(\alpha+z_2)^2\ddot{s}'(z_1)\ddot{s}'(z_2)\cdot(3.4).\end{aligned}$$

Recall the notation $\varpi = 1 - y(1-z)(zs+1)$ and suppose we have the following lemma.

Lemma 3.1. *Under the conditions of Theorem 1.1 and $z \in \mathcal{C}$, we have that given $\mathfrak{T}_N = \{$all $\mathbf{T}_N\}$, $\{S_{n1}(\cdot)\}$ forms a tight sequence on $\mathcal{C}$ and $S_{n1}(z)$ converges weakly to a two-dimensional Gaussian process $\Phi_2(z)$ satisfying*

$$\begin{aligned}&\mathbb{E}\big(\Phi_2(z)|\mathfrak{T}_N\big)\\ &\quad= \mathfrak{t}\frac{\int(\alpha y(1-z)\varpi^3 t/((1-z)\varpi - z\alpha t)^3)\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}{(1 - y\int((1-z)^2\varpi^2/((1-z)\varpi - z\alpha t)^2)\,\mathrm{d}F^Y_{\mathrm{mp}}(t))^2}\\ &\qquad+ (\mathfrak{m}_x - \mathfrak{t} - 2)(1-z)y\varpi^3\\ &\qquad\quad\times\frac{\int(1/(1-z)\varpi - z\alpha t)\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\int(\alpha t/((1-z)\varpi - z\alpha t)^2)\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}{1 - y\int((1-z)^2\varpi^2/((1-z)\varpi - z\alpha t)^2)\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}\end{aligned} \tag{3.5}$$

and

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}\big(\Phi_2(z_1), \Phi_2(z_2)|\mathfrak{T}_N\big) \qquad (3.6)\\
&\quad = \frac{\partial^2}{\partial z_1 \partial z_2}\bigg((\mathfrak{t}+1)\\
&\qquad\qquad \times \int\bigg[\bigg(\int \frac{y(1-z_1)(1-z_2)\varpi(z_1)\varpi(z_2)}{((1-z_1)\varpi - z_1\alpha t)((1-z_2)\varpi - z_2\alpha t)}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\bigg)\\
&\qquad\qquad\quad \times \bigg(1 - t\int \frac{y(1-z_1)(1-z_2)\varpi(z_1)\varpi(z_2)}{((1-z_1)\varpi - z_1\alpha t)((1-z_2)\varpi - z_2\alpha t)}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\bigg)^{-1}\mathrm{d}t\bigg]\\
&\qquad\qquad + (\mathfrak{m}_x - \mathfrak{t} - 2)y\\
&\qquad\qquad\quad \times \int \frac{(1-z_1)\varpi(z_1)}{(1-z_1)\varpi - z_1\alpha t}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\int \frac{(1-z_2)\varpi(z_2)}{(1-z_2)\varpi - z_2\alpha t}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\bigg). \qquad (3.7)
\end{aligned}$$

We postpone the proof of this lemma to the next section. Now we use the notation $\dot{s} = \dot{s}(z) = \frac{\alpha}{(\alpha+z)^2}s(\frac{z}{\alpha+z}) - \frac{1}{\alpha+z}$ and $\ddot{s}(z) = -z^{-1}(1-y) + y\dot{s}(z)$ to get

$$\varpi(z) = -\frac{\alpha z}{1-z}\ddot{s}\bigg(\frac{\alpha z}{1-z}\bigg), \qquad (3.8)$$

which can be used to rewrite (3.5) and (3.6) as

$$\begin{aligned}
&\mathbb{E}\bigg(\Phi_2\bigg(\frac{z}{\alpha+z}\bigg)\bigg|\mathfrak{T}_N\bigg)\\
&\quad \to \mathfrak{t}\frac{y(\alpha+z)^2\int \alpha t(\ddot{s}(z))^3(\ddot{s}(z)+t)^{-3}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}{(1 - y\int(\ddot{s}(z))^2(\ddot{s}(z)+t)^{-2}\,\mathrm{d}F^Y_{\mathrm{mp}}(t))^2} \qquad (3.9)\\
&\qquad + (\mathfrak{m}_x - \mathfrak{t} - 2)\\
&\qquad\quad \times \frac{y(\alpha+z)^2\int(\ddot{s}(z)/(\ddot{s}(z)+t))\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\int \alpha t(\ddot{s}(z))^2/(\ddot{s}(z)+t)^2\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}{1 - y\int(\ddot{s}(z))^2(\ddot{s}(z)+t)^{-2}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)} \qquad (3.10)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}\bigg(\Phi_2\bigg(\frac{z_1}{\alpha+z_1}\bigg), \Phi_2\bigg(\frac{z_2}{\alpha+z_2}\bigg)\bigg|\mathfrak{T}_N\bigg)\\
&\quad \to (\mathfrak{t}+1)(\alpha+z_1)^2(\alpha+z_2)^2\bigg(\frac{\ddot{s}'(z_1)\ddot{s}'(z_2)}{(\ddot{s}(z_1)-\ddot{s}(z_2))^2} - \frac{1}{(z_1-z_2)^2}\bigg)\\
&\qquad + (\mathfrak{m}_x - \mathfrak{t} - 2)y(\alpha+z_1)^2(\alpha+z_2)^2 \qquad (3.11)\\
&\qquad\quad \times \int \frac{\alpha t\ddot{s}'(z_1)}{(\ddot{s}(z_1)+t)^2}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)\int \frac{\alpha t\ddot{s}'(z_2)}{(\ddot{s}(z)+t)^2}\,\mathrm{d}F^Y_{\mathrm{mp}}(t).
\end{aligned}$$

Here we used the fact that (similar to (3.2))

$$z_1 - z_2 = \frac{\ddot{s}(z_1) - \ddot{s}(z_2)}{\ddot{s}(z_1)\ddot{s}(z_2)} \left(1 - y \int \ddot{s}(z_1)\ddot{s}(z_2)\big(\ddot{s}(z_1) + t\big)^{-1}\big(\ddot{s}(z_2) + t\big)^{-1} \, \mathrm{d}F^{Y}_{\mathrm{mp}}(t)\right).$$

As the mean and covariance of the limiting distribution are independent of the conditioning $\mathfrak{T}_N$, we conclude that S_{n1} and S_{n2} are asymptotically independent. Then from the above argument and page 473 in [16] we can get that $S_n(\frac{z}{1+z})$ converges weakly to a Gaussian process $-(1+z)^2\ddot{s}'(z)\Phi_1(z) + \Phi_2(\frac{z}{1+z})$ and together with (3.1) and Lemma 5.1 implies Theorem 1.6.

4. Proof of Lemma 3.1

In this section, we give the proof of Lemma 3.1. Following the similar truncation steps in [4] we may truncate and renormalize the random variables $\{x_{ij}\}$ as follows:

$$|x_{ij}| \le \delta_n \sqrt{n}, \qquad \mathbb{E}x_{ij} = 0 \quad \text{and} \quad \mathbb{E}|x_{ij}|^2 = 1.$$

Here $\delta_n \to 0$ which can be arbitrarily slow. Based on this truncation, we can verify that:

$$\mathbb{E}|x_{ij}|^4 = \mathfrak{m}_x + \mathrm{o}(1), \tag{4.1}$$

and if $\mathbf{X}_n$ is complex valued,

$$\mathbb{E}x_{ij}^2 = \mathrm{O}\big(n^{-1}\big).$$

We will introduce some notation and provide some bounds in the first part of this section. The proof of Lemma 3.1 will be given in the next part. The main procedures of the proofs, including the Stieltjes transform, the martingale decomposition and Burkholder's inequality, are routine in RMT, hence we will outline them without detailed descriptions. Interested readers are referred to Bai and Silverstein [2]. Throughout the rest of the paper, constants appearing in inequalities are represented by C which are nonrandom and may take different values from one appearance to another.

4.1. Definitions and some basic results

In this part, we introduce some notation and some useful results. First, we assume $z = u + \mathrm{i}\theta$ with $\theta > 0$. For simplicity, write $\mathbf{S} = \mathbf{S}_n$ and $\mathbf{B} = \mathbf{B}_n$. Let $\mathbf{D} = \mathbf{D}(z) = \mathbf{B} - z\mathbf{I}$, $\mathbf{F} = \mathbf{F}(z) = (1-z)\mathbf{S} - z\alpha_n\mathbf{T}_N$ and $\mathbf{I}$ be the identity matrix. Define $\mathbf{r}_i = n^{-1/2}\mathbf{X}_{(\cdot i)}$ where $\mathbf{X}_{(\cdot i)}$ is the ith column of $\mathbf{X}_n$, $\mathbf{S}_i = \mathbf{S} - \mathbf{r}_i\mathbf{r}_i^*$, $\mathbf{B}_i = \mathbf{S}_i(\mathbf{S}_i + \alpha_n\mathbf{T}_N)^{-1}$, $\mathbf{D}_i = \mathbf{D}_i(z) = \mathbf{B}_i - z\mathbf{I}$ and $\mathbf{F}_i = \mathbf{F}_i(z) = (1-z)\mathbf{S}_i - z\alpha_n\mathbf{T}_N$. Let $\mathbb{E}_i = \mathbb{E}(\cdot|\mathfrak{T}_N, \mathbf{r}_1, \dots, \mathbf{r}_i)$ and $\mathbb{E}_0 = \mathbb{E}(\cdot|\mathfrak{T}_N)$. Moreover, introduce

$$\varpi_i = \varpi_i(z) = \frac{1}{1 + (1-z)\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)\mathbf{r}_i}, \qquad \varpi_i^{\mathrm{tr}} = \varpi_i^{\mathrm{tr}}(z) = \frac{1}{1 + n^{-1}(1-z)\operatorname{tr}\mathbf{F}_i^{-1}(z)},$$

$$\varpi_i^{\mathbb{E}} = \varpi_i^{\mathbb{E}}(z) = \frac{1}{1 + n^{-1}(1-z)\mathbb{E}_0\operatorname{tr}\mathbf{F}_i^{-1}(z)},$$

$$\gamma_i = \gamma_i(z) = \mathbf{r}_i^* \mathbf{F}_i^{-1} \mathbf{r}_i - n^{-1} \mathbb{E}_0 \operatorname{tr} \mathbf{F}_i^{-1}, \qquad \eta_i = \eta_i(z) = \mathbf{r}_i^* \mathbf{F}_i^{-1} \mathbf{r}_i - n^{-1} \operatorname{tr} \mathbf{F}_i^{-1},$$
$$\xi_i = \xi_i(z) = n^{-1} \operatorname{tr} \mathbf{F}_i^{-1} - n^{-1} \mathbb{E}_0 \operatorname{tr} \mathbf{F}_i^{-1},$$
$$s_n = s_n(z) = s_{F^{\mathbf{B}_n}}(z), \qquad s = s(z) = s_{F^{y,H}}(z), \qquad s_0 = s_0(z) = s_{F^{y_n,H_n}}(z).$$

Obviously we have,

$$\gamma_i(z) = \eta_i(z) + \xi_i(z), \tag{4.2}$$
$$\varpi_i = \varpi_i^{\mathbb{E}} - (1-z)\varpi_i^{\mathbb{E}} \varpi_i \gamma_i = \varpi_i^{\mathbb{E}} - (1-z)\bigl(\varpi_i^{\mathbb{E}}\bigr)^2 \gamma_i + (1-z)^2 \bigl(\varpi_i^{\mathbb{E}}\bigr)^2 \varpi_i \gamma_i^2 \tag{4.3}$$

and

$$\varpi_i = \varpi_i^{\mathrm{tr}} - (1-z)\varpi_i^{\mathrm{tr}} \varpi_i \eta_i = \varpi_i^{\mathrm{tr}} - (1-z)\bigl(\varpi_i^{\mathrm{tr}}\bigr)^2 \eta_i + (1-z)^2 \bigl(\varpi_i^{\mathrm{tr}}\bigr)^2 \varpi_i \eta_i^2. \tag{4.4}$$

It is easy to verify that

$$\Im (1-z)^{-1} = \theta |1-z|^{-2} \tag{4.5}$$

and

$$\Im \mathbf{r}_i^* \mathbf{F}_i^{-1}(z) \mathbf{r}_i = \theta \mathbf{r}_i^* \mathbf{F}_i^{-1}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N) \mathbf{F}_i^{-1}(\bar{z}) \mathbf{r}_i \tag{4.6}$$

have the same sign. Therefore from the definition of ϖ_i, we have

$$|\varpi_i| = \left| \frac{1}{1-z} \frac{1}{1/(1-z) + \mathbf{r}_i^* \mathbf{F}_i^{-1}(z) \mathbf{r}_i} \right| \le \frac{|1-z|}{\theta}. \tag{4.7}$$

Similarly we can obtain

$$\bigl|\varpi_i^{\mathrm{tr}}\bigr| \le \frac{|1-z|}{\theta}, \qquad \bigl|\varpi_i^{\mathbb{E}}\bigr| \le \frac{|1-z|}{\theta}. \tag{4.8}$$

By the fact that

$$\bigl\| \mathbf{F}_i^{-1}(z) \bigr\| = \bigl\| \mathbf{D}_i^{-1}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N)^{-1} \bigr\| \le C\theta^{-1} \tag{4.9}$$

and Lemma 5.4, we have for any $l \ge 2$

$$\mathbb{E}\bigl|\eta_i(z)\bigr|^l \le \frac{C\delta_n^{2l-4}}{n\theta^l}. \tag{4.10}$$

In the last inequality we used $|x_{ij}| \le \delta_n \sqrt{n}$. For any invertible matrices $\mathbf{M}$, $\mathbf{M} + \mathbf{r}_i \mathbf{r}_i^*$ and $\mathbf{N}$, using

$$\mathbf{r}_i^* \bigl(\mathbf{M} + \mathbf{r}_i \mathbf{r}_i^*\bigr)^{-1} = \frac{1}{1 + \mathbf{r}_i^* \mathbf{M} \mathbf{r}_i} \mathbf{r}_i^* \mathbf{M}^{-1}, \qquad \mathbf{M}^{-1} - \mathbf{N}^{-1} = -\mathbf{N}^{-1}(\mathbf{M} - \mathbf{N})\mathbf{M}^{-1}, \tag{4.11}$$

we obtain that

$$\mathbf{F}^{-1}(z) - \mathbf{F}_i^{-1}(z) = -(1-z)\varpi_i \mathbf{F}_i^{-1}\mathbf{r}_i\mathbf{r}_i^*\mathbf{F}_i^{-1}, \tag{4.12}$$

which together with (4.6)–(4.9) implies that for any Hermitian matrix $\mathbf{M}$ with $\|\mathbf{M}\| \le C$,

$$\left|\operatorname{tr}\mathbf{F}^{-1}(z)\mathbf{M} - \operatorname{tr}\mathbf{F}_i^{-1}(z)\mathbf{M}\right| = \left|(1-z)\varpi_i \mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i\right| \le C\theta^{-1}. \tag{4.13}$$

Lemma 4.1. *Under the conditions of Theorem 1.6, we have for any non-random Hermitian matrix $\mathbf{M}$ with $\|\mathbf{M}\| \le C$ and $l \ge 2$,*

$$\mathbb{E}\left|n^{-1}\operatorname{tr}\mathbf{F}^{-1}(z)\mathbf{M} - n^{-1}\mathbb{E}_0\operatorname{tr}\mathbf{F}^{-1}(z)\mathbf{M}\right|^l \le \frac{C_l\delta_n^{2l-4}}{n^{l/2+1}\theta^{3l}}, \qquad \textit{where } z = u + \mathrm{i}\theta.$$

Proof. The martingale decomposition (one can refer to [2] for more details) gives

$$\begin{aligned}
\operatorname{tr}\mathbf{F}^{-1}\mathbf{M} - \mathbb{E}_0\operatorname{tr}\mathbf{F}^{-1}\mathbf{M} &= \sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\operatorname{tr}\bigl(\mathbf{F}^{-1}\mathbf{M} - \mathbf{F}_i^{-1}\mathbf{M}\bigr) \\
&= -(1-z)\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i \mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i \\
&= (z-1)\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i^{\mathrm{tr}} \mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i \\
&\quad + (1-z)^2\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i^{\mathrm{tr}}\varpi_i\eta_i \mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i.
\end{aligned}$$

Here we used (4.12) and (4.4). From (4.9) and Lemma 5.4, we obtain that

$$\mathbb{E}\left|\mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i - n^{-1}\operatorname{tr}\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\right|^l \le \frac{C\delta_n^{2l-4}}{n\theta^{2l}}.$$

Thus it follows from (4.8) and Lemma 5.6 that

$$\begin{aligned}
&\mathbb{E}\left|\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i^{\mathrm{tr}}\mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i\right|^l \\
&\quad = \mathbb{E}\left|\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i^{\mathrm{tr}}\bigl(\mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i - n^{-1}\operatorname{tr}\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\bigr)\right|^l \le \frac{Cn^{l/2}\delta_n^{2l-4}}{n\theta^{3l}}.
\end{aligned}$$

On the other hand, from (4.8), (4.13), (4.10) and Lemma 5.6 we also have

$$\mathbb{E}\left|\sum_{i=1}^n (\mathbb{E}_i - \mathbb{E}_{i-1})\varpi_i^{\mathrm{tr}}\varpi_i\eta_i\mathbf{r}_i^*\mathbf{F}_i^{-1}\mathbf{M}\mathbf{F}_i^{-1}\mathbf{r}_i\right|^l \le \frac{Cn^{l/2}\delta_n^{2l-4}}{n\theta^{3l}},$$

which completes the proof. □

Remark 4.2. From the last lemma and (4.13), one can easily verify that for any $l \geq 2$,

$$\mathbb{E}\big|\operatorname{tr}\mathbf{F}_i^{-1}(z)\mathbf{M} - \mathbb{E}\operatorname{tr}\mathbf{F}_i^{-1}(z)\mathbf{M}\big|^l \leq \frac{C_l n^{l/2}\delta_n^{2l-4}}{n\theta^{3l}}. \tag{4.14}$$

Furthermore, by combining (4.2), (4.10) and (4.14) with $\mathbf{M} = \mathbf{I}$, we have for any $l \geq 2$,

$$\mathbb{E}|\gamma_i|^l \leq \frac{C_l\delta_n^{2l-4}}{n}. \tag{4.15}$$

Denote $\mathbf{S}_{ij} = \mathbf{S} - \mathbf{r}_i\mathbf{r}_i^* - \mathbf{r}_j\mathbf{r}_j^*$ for $i \neq j$. Correspondingly, let $\mathbf{B}_{ij} = \mathbf{S}_{ij}(\mathbf{S}_{ij} + \alpha_n\mathbf{T}_N)^{-1}$, $\mathbf{D}_{ij} = \mathbf{D}_{ij}(z) = \mathbf{B}_{ij} - z\mathbf{I}$, $\mathbf{F}_{ij} = \mathbf{F}_{ij}(z) = (1-z)\mathbf{S}_{ij} - z\alpha\mathbf{T}_N$ and assume $\|(\mathbf{S}_{ij} + \alpha_n\mathbf{T}_N)^{-1}\| < \infty$. Moreover, we have

$$\varpi_{ij} = \varpi_{ij}(z) = \frac{1}{1 + (1-z)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\mathbf{r}_j}, \qquad \varpi_{ij}^{\mathrm{tr}} = \varpi_{ij}^{\mathrm{tr}}(z) = \frac{1}{1 + n^{-1}(1-z)\operatorname{tr}\mathbf{F}_{ij}^{-1}(z)},$$

$$\varpi_{ij}^{\mathbb{E}} = \varpi_{ij}^{\mathbb{E}}(z) = \frac{1}{1 + n^{-1}(1-z)\mathbb{E}_0\operatorname{tr}\mathbf{F}_{ij}^{-1}(z)},$$

$$\gamma_{ij} = \gamma_{ij}(z) = \mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\mathbf{r}_j - n^{-1}\mathbb{E}_0\operatorname{tr}\mathbf{F}_{ij}^{-1}(z),$$

$$\eta_{ij} = \eta_{ij}(z) = \mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z),$$

$$\xi_{ij} = \xi_{ij}(z) = n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z) - n^{-1}\mathbb{E}_0\operatorname{tr}\mathbf{F}_{ij}^{-1}(z).$$

We can get the same bound as we did in (4.2)–(4.13) by changing the subscript i to ij. Thus from now on when we consider these bounds we will ignore the subscripts. Let $\mathbf{H}_{12} = \mathbf{H}_{12}(z) = (1-z)\frac{n-1}{n}\varpi_{12}^{\mathbb{E}}\mathbf{I} - z\alpha_n\mathbf{T}_N$. We have the following lemma.

Lemma 4.3. *Under the conditions of Theorem 1.6 and $z = u + \mathrm{i}\theta$, we have for any $1 \leq k \leq p$, $1 \leq i \leq n$ and non-random matrix $\mathbf{M}$ with $\|\mathbf{M}\| \leq C$*

$$\mathbb{E}_0\mathbf{e}_k^*\mathbf{F}_i^{-1}(z)\mathbf{M}\mathbf{e}_k = \mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{M}\mathbf{e}_k + \mathrm{O}\big(n^{-1/2}\big), \tag{4.16}$$

where $\mathbf{e}_k$ is the p-dimensional vector with the kth coordinate being 1 and the remaining being zero.

Proof. Using (4.11), we can check that

$$\begin{aligned}
\mathbf{F}_i^{-1}(z) &= \mathbf{H}_{12}^{-1}(z) + \frac{\varpi_{12}^{\mathbb{E}}(1-z)}{n}\sum_{j\neq i}\mathbf{H}_{12}^{-1}(z)\big(\mathbf{F}_i^{-1}(z) - \mathbf{F}_{ij}^{-1}(z)\big) \\
&\quad + \frac{\varpi_{12}^{\mathbb{E}}(1-z)}{n}\sum_{j\neq i}\mathbf{H}_{12}^{-1}(z)\mathbf{F}_{ij}^{-1}(z) - (1-z)\sum_{j\neq i}\varpi_{ij}\mathbf{H}_{12}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z) \qquad (4.17) \\
&= \mathbf{H}_{12}^{-1}(z) + H_{(1)} - H_{(2)} - H_{(3)},
\end{aligned}$$

where

$$H_{(1)} = \frac{\varpi_{12}^{\mathbb{E}}(1-z)}{n} \sum_{j \neq i} \mathbf{H}_{12}^{-1}(z)\big(\mathbf{F}_i^{-1}(z) - \mathbf{F}_{ij}^{-1}(z)\big),$$

$$H_{(2)} = (1-z)\varpi_{12}^{\mathbb{E}} \sum_{j \neq i} \big(\mathbf{H}_{12}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z) - n^{-1}\mathbf{H}_{12}^{-1}(z)\mathbf{F}_{ij}^{-1}(z)\big),$$

$$H_{(3)} = (1-z) \sum_{j \neq i} \big(\varpi_{ij} - \varpi_{12}^{\mathbb{E}}\big)\mathbf{H}_{12}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z).$$

Note that, similar to (4.5), either the real parts or the imaginary parts of $(1-z)\varpi_{12}^{\mathbb{E}}$ and $-z$ have the same sign. Thus, we have for any $t \geq 0$

$$\left|(1-z)\frac{n-1}{n}\varpi_{12}^{\mathbb{E}} - z\alpha_n t\right|^{-1} \leq \frac{C}{\theta^3}, \tag{4.18}$$

which implies

$$\left\|\mathbf{H}_{12}^{-1}(z)\right\| \leq \frac{C}{\theta^3}. \tag{4.19}$$

Then it follows from (4.9), (4.19) and Lemma 5.4 that

$$\begin{aligned}
&\mathbb{E}_0\left|\mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\mathbf{M}\mathbf{e}_k\right|^2 \\
&\quad \leq Cn^{-2}\mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{H}_{12}^{-1}(\bar{z})\mathbf{e}_k\mathbb{E}_0\mathbf{e}_k^*\mathbf{M}\mathbf{F}_{ij}^{-1}(z)\mathbf{F}_{ij}^{-1}(\bar{z})\mathbf{M}^*\mathbf{e}_k + n^{-2}\mathbb{E}_0\left|\mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{F}_{ij}^{-1}(z)\mathbf{M}\mathbf{e}_k\right|^2.
\end{aligned}$$

From (4.19), we have

$$\left|\mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{e}_k\right| \quad \text{and} \quad \mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{H}_{12}^{-1}(\bar{z})\mathbf{e}_k \tag{4.20}$$

are both bounded from above. In addition, by (4.9) we get that

$$\mathbf{e}_k^*\mathbf{M}\mathbf{F}_{ij}^{-1}(z)\mathbf{F}_{ij}^{-1}(\bar{z})\mathbf{M}\mathbf{e}_k \leq C\left\|\mathbf{F}_{ij}^{-1}(z)\right\|^2 \leq C\theta^{-2} \tag{4.21}$$

and

$$\left|\mathbf{e}_k^*\mathbf{H}_{12}^{-1}(z)\mathbf{F}_{ij}^{-1}(z)\mathbf{M}\mathbf{e}_k\right| \leq C\theta^{-4}. \tag{4.22}$$

Thus combining (4.3), (4.8), (4.15), (4.12), (4.21), (4.22) and Hölder's inequalitywe obtain

$$\mathbb{E}_0|H_{(1)}| = \mathrm{O}\big(n^{-1}\big) \quad \text{and} \quad \mathbb{E}_0|H_{(3)}| = \mathrm{O}\big(n^{-1/2}\big).$$

Apparently we have $\mathbb{E}_0 H_{(2)} = 0$. Thus the proof of the lemma is complete. □

Lemma 4.4. *Under the conditions of Theorem 1.6 and $z = u + \mathrm{i}\theta$, we have for any $1 \le k \le p$, $1 \le j \le n$ and non-random matrix $\mathbf{M}$ with $\|\mathbf{M}\| \le C$*

$$\mathbb{E}\big|\mathbf{e}_k^*\mathbf{F}^{-1}(z)\mathbf{M}\mathbf{e}_k - \mathbb{E}_0\mathbf{e}_k^*\mathbf{F}^{-1}(z)\mathbf{M}\mathbf{e}_k\big|^2 = \mathrm{O}\big(n^{-1}\big)$$

and

$$\mathbb{E}\big|\mathbf{e}_k^*\mathbf{F}_j^{-1}(z)\mathbf{M}\mathbf{e}_k - \mathbb{E}_0\mathbf{e}_k^*\mathbf{F}_j^{-1}(z)\mathbf{M}\mathbf{e}_k\big|^2 = \mathrm{O}\big(n^{-1}\big).$$

Proof. Similarly to the proof of Lemma 4.1 and Lemma 4.3, we can easily get this lemma and we omit details. □

Lemma 4.5. *For any non-random matrix $\mathbf{M}$ with $\|\mathbf{M}\| \le C$ and $z_1 = u_1 + \mathrm{i}\theta_1, z_2 = u_2 + \mathrm{i}\theta_2$ with $\min\{\theta_1, \theta_2\} > 0$, we have*

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big) - \mathbb{E}_0\left(\frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big)\right)\right|^2 = \mathrm{O}\big(n^{-2}\big).$$

Remark 4.6. Checking the proof of Lemma 4.5, we see that Lemma 4.5 holds as well when we replace $\mathbb{E}_i(\mathbf{F}_i^{-1}(z_2))$ by $\mathbf{F}_i^{-1}(z_2)$. The main difference in the arguments is that we do not distinguish between the cases $j < i$ and $j > i$ when dealing with the latter.

Proof of Lemma 4.5. Using the martingale decomposition, we have

$$\begin{aligned}
&\frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big) - \mathbb{E}_0\left(\frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big)\right)\\
&\quad = \frac{1}{n}\sum_{j\ne i}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\big[\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big) + \operatorname{tr}\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_{ij}^{-1}(z_2)\big)\big]\\
&\quad = \frac{1}{n}\sum_{j\ne i}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})(\mathcal{K}_1 + \mathcal{K}_2 + \mathcal{K}_3),
\end{aligned}$$

where (via (4.12))

$$\begin{aligned}
\mathcal{K}_1 &= \varpi_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\varpi_{ij}(z_2)\mathbf{F}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_2)\big)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j,\\
\mathcal{K}_2 &= -\varpi_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_{ij}^{-1}(z_2)\big)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j,\\
\mathcal{K}_3 &= -\operatorname{tr}\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\varpi_{ij}(z_2)\mathbf{F}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_2)\big).
\end{aligned}$$

Note that by (4.13)

$$|\varpi_{ij}|\big\|\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\big\|^2 = \big|\varpi_{ij}\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\mathbf{F}_{ij}^{-1}(\bar{z})\mathbf{r}_j\big| \le C, \tag{4.23}$$

which implies that $\mathcal{K}_1$ is bounded.

When $j > i$, applying (4.3) to get

$$(\mathbb{E}_j - \mathbb{E}_{j-1})\mathcal{K}_1 = (\mathbb{E}_j - \mathbb{E}_{j-1})\varpi_{12}^{\mathbb{E}}(z_1)(\mathcal{K}_{11} - \mathcal{K}_{12}),$$

where $\mathcal{K}_{12} = \gamma_{kj}(z_1)\mathcal{K}_1$ and

$$\begin{aligned}\mathcal{K}_{11} = {} & \mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\varpi_{ij}(z_2)\mathbf{G}_k(z_2)\big)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j \\ & - n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\big(\varpi_{ij}(z_2)\mathbf{G}_{ij}(z_2)\big)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\end{aligned}$$

with $\mathbf{G}_{ij}(z_2) = \mathbf{F}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_2)$. We conclude from (4.23), (4.8), (4.15), Lemmas 5.4, 5.6 and $\|\mathbf{M}\| \le C$ that

$$\mathbb{E}\left|\frac{1}{n}\sum_{j>i}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})(\mathcal{K}_{11} - \mathcal{K}_{12})\right|^2 \le \frac{M}{n^2}\sum_{j>i}^{n}\big(\mathbb{E}|\mathcal{K}_{11}|^2 + \mathbb{E}|\mathcal{K}_{12}|^2\big) \le \frac{C}{n^2}.$$

On the other hand, when $j < i$, we define $\underline{\mathbf{F}}_{ij}^{-1}(z), \underline{\varpi}_{ij}(z)$ and $\underline{\gamma}_{ij}(z)$ using $\mathbf{r}_1, \dots, \mathbf{r}_{j-1}, \underline{\mathbf{r}}_{j+1}, \dots, \underline{\mathbf{r}}_{i-1}, \underline{\mathbf{r}}_{i+1}, \dots, \underline{\mathbf{r}}_n$ as $\mathbf{F}_{ij}^{-1}(z), \varpi_{ij}(z)$ and $\gamma_{kj}(z)$ are defined using $\mathbf{r}_1, \dots, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \dots, \mathbf{r}_{i-1}, \mathbf{r}_{i+1}, \dots, \mathbf{r}_n$. Here $\underline{\mathbf{r}}_1, \dots, \underline{\mathbf{r}}_n$ are i.i.d. copies of $\mathbf{r}_1$ and independent of $\{\mathbf{r}_j, j = 1, \dots, n\}$. Let

$$\mathcal{R}_{ij1}(z_1, z_2) = \mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j, \qquad \mathcal{R}_{ij2}(z_1, z_2) = \mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j.$$

Applying the equality for $\underline{\varpi}_{kj}(z_2)$ similar to (4.3) yields

$$\begin{aligned}(\mathbb{E}_j - \mathbb{E}_{j-1})\mathcal{K}_1 &= (\mathbb{E}_j - \mathbb{E}_{j-1})\big[\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\mathcal{R}_{ij1}(z_1, z_2)\mathcal{R}_{ij2}(z_1, z_2)\big] \\ &= (\mathbb{E}_j - \mathbb{E}_{j-1})(\mathcal{K}_{13} + \mathcal{K}_{14} - \mathcal{K}_{15} - \mathcal{K}_{16}),\end{aligned}$$

where

$$\begin{aligned}\mathcal{K}_{13} &= \varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\mathcal{T}_{ij1}(z_1, z_2)\mathcal{R}_{ij2}(z_1, z_2), \\ \mathcal{K}_{14} &= \varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathcal{T}_{ij2}(z_1, z_2), \\ \mathcal{K}_{15} &= \varpi_{12}^{\mathbb{E}}(z_1)\varpi_{12}^{\mathbb{E}}(z_2)\underline{\varpi}_{ij}(z_2)\underline{\gamma}_{ij}(z_2)n^{-2}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1), \\ \mathcal{K}_{16} &= \varpi_{12}^{\mathbb{E}}(z_1)\varpi_{ij}(z_1)\gamma_{ij}(z_1)\underline{\varpi}_{ij}(z_2)n^{-2}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\end{aligned}$$

with

$$\begin{aligned}\mathcal{T}_{ij1}(z_1, z_2) &= \mathcal{R}_{i1} - n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2), \\ \mathcal{T}_{ij2}(z_1, z_2) &= \mathcal{R}_{i2} - n^{-1}\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1).\end{aligned}$$

Apparently $\underline{\mathbf{F}}_{ij}^{-1}(z), \underline{\varpi}_{ij}(z)$ and $\underline{\gamma}_{ij}(z)$ have the same bound as $\mathbf{F}_{ij}^{-1}(z), \varpi_{ij}(z)$ and $\gamma_{ij}(z)$, respectively. Thus it follows from Lemma 5.4, (4.9) and (4.23) that

$$\mathbb{E}\left|\mathcal{T}_{ij1}(z_1, z_2)\right|^2 \le \frac{C}{n}, \qquad \mathbb{E}\left|\mathcal{T}_{ij2}(z_1, z_2)\right|^2 \le \frac{C}{n} \tag{4.24}$$

and

$$n^{-1}\left|\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_1)\right| < C, \qquad n^{-1}\left|\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{F}_{ij}^{-1}(z_1)\right| < C. \tag{4.25}$$

Therefore combining (4.15), (4.7), (4.8), (4.24) and (4.25), we can obtain that for $t = 3, 4, 5, 6$,

$$\mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\mathcal{K}_{1t}\right|^2 = \mathrm{O}\left(n^{-1}\right).$$

This via Lemma 5.6 implies that

$$\mathbb{E}\left|\frac{1}{n}\sum_{j<i}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\mathcal{K}_1\right|^2 = \mathrm{O}\left(n^{-2}\right).$$

The terms $\mathcal{K}_2$ and $\mathcal{K}_3$ can be similarly proved to have the same order. Then the proof of Lemma 4.5 is complete. □

Now we use (4.17) to write that

$$\begin{aligned}
&\frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\mathbf{F}_i^{-1}(z_2) \\
&\quad= \frac{1}{n}\operatorname{tr}\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbb{E}_i\mathbf{F}_i^{-1}(z_2) + \frac{1}{n}\operatorname{tr}\mathbf{M}H_{(1)}(z_1)\mathbb{E}_i\mathbf{F}_i^{-1}(z_2) \\
&\qquad - \frac{1}{n}\operatorname{tr}\mathbf{M}H_{(2)}(z_1)\mathbb{E}_i\mathbf{F}_i^{-1}(z_2) - \frac{1}{n}\operatorname{tr}\mathbf{M}H_{(3)}(z_1)\mathbb{E}_i\mathbf{F}_i^{-1}(z_2).
\end{aligned} \tag{4.26}$$

Then we have the following lemmas.

Lemma 4.7. *For any non-random matrix* $\mathbf{M}$ *with* $\|\mathbf{M}\| \le C$ *and* $z_1 = u_1 + \mathrm{i}\theta_1, z_2 = u_2 + \mathrm{i}\theta_2$ *with* $\min\{\theta_1, \theta_2\} > 0$, *we have*

$$\left|\mathbb{E}_0\operatorname{tr}\mathbf{M}H_{(1)}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\right| = \mathrm{O}_p(1) \tag{4.27}$$

and

$$\left|\mathbb{E}_0\operatorname{tr}\mathbf{M}H_{(3)}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\right| = \mathrm{O}_p(1). \tag{4.28}$$

Proof. By (4.12), we obtain that

$$\begin{aligned}
&\operatorname{tr}\mathbf{M}H_{(1)}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right) \\
&\quad= \frac{\varpi_{12}^{\mathbb{E}}(z_1)(1-z_1)}{n}\sum_{j\ne i}\varpi_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j.
\end{aligned}$$

As $\mathbf{H}_{12}^{-1}(z)$ $\mathbf{F}_i^{-1}(z)$, $\mathbf{F}_{ij}^{-1}(z)$, $\varpi_{12}^{\mathbb{E}}(z)$ and $\varpi_{12}(z)$ are all bounded when $\Im z > 0$, we can get directly that for $j > i$,

$$\left|\mathbb{E}_0 \operatorname{tr} \mathbf{M} H_{(1)}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\right| \le C.$$

When $j < i$, note that we also have

$$\left|\mathbb{E}_0 \operatorname{tr} \mathbf{M} H_{(1)}(z_1)\mathbb{E}_i\left(\mathbf{F}_{ij}^{-1}(z_2)\right)\right| \le C.$$

Then from (4.12), $\mathbb{E}|x_{ij}| < \infty$ and the definition of $\underline{\mathbf{F}}_{ij}^{-1}(z)$, $\underline{\varpi}_{ij}(z)$ and $\underline{\gamma}_{ij}(z)$ in Lemma 4.5 we have

$$\begin{aligned}
&\left|\mathbb{E}_0 \varpi_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2) - \mathbf{F}_{ij}^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j\right| \\
&\quad = \left|\mathbb{E}_0 \varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{F}_{ij}^{-1}(z_1)\mathbf{r}_j\right| = \mathrm{O}(1),
\end{aligned}$$

which completes the proof of (4.27).

Now consider (4.28). When $j < i$, using (4.3) we rewrite the left-hand side of (4.28) as

$$\begin{aligned}
&\left|(1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}_0 \sum_{j\neq i} \varpi_{ij}(z_1)\gamma_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j\right| \\
&\quad = \left|(1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}_0 \sum_{j\neq i} \varpi_{ij}(z_1)\gamma_{ij}(z_1)\mathcal{T}_{ij3}(z_1,z_2)\right. \qquad (4.29) \\
&\qquad + (1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}_0 n^{-1} \qquad (4.30) \\
&\qquad \left.\times \sum_{j\neq i} \varpi_{ij}(z_1)\gamma_{ij}(z_1) \operatorname{tr} \mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\right|,
\end{aligned}$$

where

$$\mathcal{T}_{ij3}(z_1,z_2) = \mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1).$$

From Lemma 5.4, we have $\mathbb{E}|\mathcal{T}_{ij3}(z_1,z_2)|^2 = \mathrm{O}(n^{-1})$ which together with (4.15) and Hölder's inequalityimplies

$$(4.29) = \mathrm{O}(1).$$

For (4.30), we apply (4.3) again and obtain that

$$\left|(4.30)\right| = \left|\left((1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\right)^2\mathbb{E}_0 n^{-1}\sum_{j\neq i}\varpi_{ij}(z_1)\gamma_{ij}^2(z_1)\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i\left(\mathbf{F}_i^{-1}(z_2)\right)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\right|.$$

Here we have used the fact that $|n^{-1}\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\mathbb{E}_i(\mathbf{F}_i^{-1}(z_2))\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)|$ is bounded. Thus from (4.15), we get that

$$(4.30) = \mathrm{O}(1).$$

On the other hand, when $j > i$, the above argument apparently also works if we replace $\mathbb{E}_i(\mathbf{F}_i^{-1}(z_2))$ with $\mathbb{E}_i(\mathbf{F}_{ij}^{-1}(z_2))$. And the remaining term can be expressed as

$$\begin{aligned}
&(1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}\sum_{j\neq i}\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\gamma_{ij}(z_1)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j \\
&\quad = (1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}\sum_{j\neq i}\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\gamma_{ij}(z_1)\mathcal{T}_{ij1}\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j \qquad (4.31)\\
&\qquad + (1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\mathbb{E}n^{-1}\sum_{j\neq i}\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\gamma_{ij}(z_1)\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathcal{T}_{ij2}(z_1,z_2)\\
&\qquad - (1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)\varpi_{12}^{\mathbb{E}}(z_2)\mathbb{E}n^{-2}\sum_{j\neq i}\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\gamma_{ij}(z_1)\gamma_{ij}(z_2)\\
&\qquad\qquad \cdot\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\\
&\qquad - \big((1-z_1)\varpi_{12}^{\mathbb{E}}(z_1)\big)^2\mathbb{E}n^{-2}\sum_{j\neq i}\varpi_{ij}(z_1)\underline{\varpi}_{ij}(z_2)\big(\gamma_{ij}(z_1)\big)^2 \qquad (4.32)\\
&\qquad\qquad \cdot\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1).
\end{aligned}$$

By Lemma 5.4 and a similar argument in (4.27), we can show that (4.31) and (4.32) are all bounded. Then the proof of Lemma 4.7 is complete. □

Lemma 4.8. *For any non-random matrix* $\mathbf{M}$ *with* $\|\mathbf{M}\| \le C$ *and* $z_1 = u_1 + \mathrm{i}\theta_1, z_2 = u_2 + \mathrm{i}\theta_2$ *with* $\min\{\theta_1,\theta_2\} > 0$, *we have*

$$\begin{aligned}
&\mathbb{E}_0 n^{-1}\operatorname{tr}\mathbf{M}H_{(2)}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big)\\
&\quad = -(i-1)n^{-2}(1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)\varpi_{12}^{\mathbb{E}}(z_2)\\
&\qquad \cdot\operatorname{tr}\mathbf{H}_{12}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_i^{-1}(z_1)\underline{\mathbf{F}}_i^{-1}(z_2) + \mathrm{O}_p(1).
\end{aligned}$$

Proof. It follows from (4.12), (4.8), (4.9), (4.19) and $\mathbb{E}|x_{ij}| < C$ that

$$\begin{aligned}
&\mathbb{E}_0\operatorname{tr}\mathbf{M}H_{(2)}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big)\\
&\quad = -(1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)\\
&\qquad \times\sum_{j<i}\mathbb{E}_0\underline{\varpi}_{ij}(z_2)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j \qquad (4.33)
\end{aligned}$$

$$+(1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)$$
$$\times\frac{1}{n}\sum_{j<i}\mathbb{E}_0\underline{\varpi}_{ij}(z_2)\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{F}_{ij}^{-1}(z)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j$$

and

$$(1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)\sum_{j<i}\mathbb{E}_0\underline{\varpi}_{ij}(z_2)\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{F}_{ij}^{-1}(z)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j=\mathrm{O}_p(n).$$

Applying (4.3) to rewrite the first term of (4.33) as

$$\sum_{j<i}\mathbb{E}_0\underline{\varpi}_{ij}(z_2)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j$$
$$=\varpi_{12}^{\mathbb{E}}(z_2)n^{-2}\sum_{j<i}\mathbb{E}_0\operatorname{tr}\mathbf{F}_{ij}^{-1}(z_1)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1) \tag{4.34}$$
$$+\varpi_{12}^{\mathbb{E}}(z_2)\sum_{j<i}\mathbb{E}_0\mathcal{T}_{ij1}(z_1,z_2)\mathcal{T}_{ij4}(z_1,z_2) \tag{4.35}$$
$$-\varpi_{12}^{\mathbb{E}}(z_2)\sum_{j<i}\mathbb{E}_0\underline{\varpi}_{ij}(z_2)\underline{\gamma}_{ij}(z_2)\mathbf{r}_j^*\mathbf{F}_{ij}^{-1}(z)\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{r}_j\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j, \tag{4.36}$$

where $\mathcal{T}_{ij4}(z_1,z_2)=\mathbf{r}_j^*\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbf{r}_j-n^{-1}\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)$. The arguments in (4.31) and (4.32) and (4.24) ensure that

$$(4.35)=\mathrm{O}_p(1)\quad\text{and}\quad(4.36)=\mathrm{O}_p(1).$$

In addition, from (4.12) and (4.13), we have

$$\mathbb{E}_0\operatorname{tr}\underline{\mathbf{F}}_{ij}^{-1}(z_2)\mathbf{H}_{12}^{-1}(z_1)=\mathbb{E}_0\operatorname{tr}\underline{\mathbf{F}}_i^{-1}(z_2)\mathbf{H}_{12}^{-1}(z_1)+\mathrm{O}_p(1).$$

Then using (4.17) again and repeating similar arguments as in the proof of Lemma 4.7 we obtain that

$$\mathbb{E}_0\operatorname{tr}\underline{\mathbf{F}}_i^{-1}(z_2)\mathbf{H}_{12}^{-1}(z_1)=\operatorname{tr}\mathbf{H}_{12}^{-1}(z_2)\mathbf{H}_{12}^{-1}(z_1)+\mathrm{O}_p(1).$$

Combining the above arguments, we conclude that

$$\mathbb{E}_0 n^{-1}\operatorname{tr}\mathbf{M}H_{(2)}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big)$$
$$=-(i-1)n^{-2}(1-z_1)(1-z_2)\varpi_{12}^{\mathbb{E}}(z_1)\varpi_{12}^{\mathbb{E}}(z_2)\cdot\operatorname{tr}\mathbf{H}_{12}^{-1}(z_2)\mathbf{M}\mathbf{H}_{12}^{-1}(z_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_i^{-1}(z_1)\underline{\mathbf{F}}_i^{-1}(z_2)$$
$$+\mathrm{O}_p(1),$$

which complete the proof of the lemma. □

Remark 4.9. Let $\mathbf{H}_1 = \mathbf{H}_1(z) = (1-z)\varpi_1^{\mathbb{E}}\mathbf{I} - z\alpha_n \mathbf{T}_N$. We conclude from the above arguments and the fact $|\varpi_{12}^{\mathbb{E}} - \varpi_1^{\mathbb{E}}| = \mathrm{O}(n^{-1})$ that

$$\mathbf{e}_k^* \mathbf{F}_i^{-1}(z)\mathbf{e}_k = \mathbf{e}_k^* \mathbf{H}_1^{-1}(z)\mathbf{e}_k + \mathrm{O}_p\big(n^{-1/2}\big) \tag{4.37}$$

and

$$\begin{aligned} &\frac{1}{n}\operatorname{tr}\mathbf{F}_i^{-1}(z_1)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)\big) \\ &\quad \to \frac{(1/n)\operatorname{tr}\mathbf{H}_1^{-1}(z_2)\mathbf{H}_1^{-1}(z_1)}{1-(((i-1)(1-z_1)(1-z_2)\varpi_1^{\mathbb{E}}(z_1)\varpi_1^{\mathbb{E}}(z_2))/n^2)\operatorname{tr}\mathbf{H}_1^{-1}(z_2)\mathbf{H}_1^{-1}(z_1)}. \end{aligned} \tag{4.38}$$

Here we have used the fact that the denominator of (4.38) is bounded when $\min\{\theta_1, \theta_2\} > 0$.

4.2. Proof of Lemma 3.1

Note that the contour $\mathcal{C}$ of the integration contains four segments: two horizontal lines and two vertical lines. We need to calculate the limit of $S_{n1}(z)$ at the four segments respectively. First of all, considering the top horizontal line $\mathcal{C}^t = \{z \in \mathbb{C} : \Re z \in [c_l - \theta, c_r + \theta], \Im z = \theta\}$, we know that there exists some event $\mathcal{Q}_n$ with $\mathbb{P}(\mathcal{Q}_n) \to 1$ such that,

$$\mathbb{E}\big|s_n(z) - s_n(z)I(\mathcal{Q}_n)\big| \le (\Im z)^{-1}\mathbb{P}\big(\mathcal{Q}_n^c\big) \to 0.$$

In this part, we let $\mathcal{Q} = \mathcal{Q}_n = \{\|(\mathbf{S}_n + \alpha_n \mathbf{T}_N)^{-1}\| \le C\}$ with some $C < \infty$. By (1.4) we have that for any $l > 0$, $\mathbb{P}(\mathcal{Q}^c) \le n^{-l}$. It is known that $\lambda_1^{\mathbf{S}+\alpha_n\mathbf{T}_N} \ge \lambda_1^{\mathbf{S}_i+\alpha_n\mathbf{T}_N} \ge \lambda_1^{\mathbf{S}_{ij}+\alpha_n\mathbf{T}_N}$ for any i, j, which implies

$$\mathcal{Q} \supseteq \mathcal{Q}_i \supseteq \mathcal{Q}_{ij}.$$

Here $\mathcal{Q}_i = \{\|(\mathbf{S}_i + \alpha_n \mathbf{T}_N)^{-1}\| \le C\}$ and $\mathcal{Q}_{ij} = \{\|(\mathbf{S}_{ij} + \alpha_n \mathbf{T}_N)^{-1}\| \le C\}$. Notice that we also have

$$\mathbb{P}\big(\mathcal{Q}_i^c\big) \le n^{-l} \quad \text{and} \quad \mathbb{P}\big(\mathcal{Q}_{ij}^c\big) \le n^{-l}.$$

Now we rewrite $S_{n1}(z)$ as $S_{n1} = S_{n1}^{(1)} + S_{n1}^{(2)} + \mathrm{o}_p(1)$ with

$$\begin{aligned} S_{n1}^{(1)} &= p\big(s_n(z)I(\mathcal{Q}) - \mathbb{E}_0\big[s_n(z)I(\mathcal{Q})\big]\big) \qquad \text{covariance part,} \\ S_{n1}^{(2)} &= p\big(\mathbb{E}_0 s_n(z)I(\mathcal{Q}) - s_0(z)I(\mathcal{Q})\big) \qquad \text{mean part.} \end{aligned}$$

4.2.1. *The covariance part*

The martingale decomposition used in the proof of Lemma 4.1 gives that

$$\begin{aligned}
S_{n1}^{(1)} &= \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\operatorname{tr}\big(\mathbf{D}^{-1} - \mathbf{D}_i^{-1}\big)I(\mathcal{Q}_i) + \mathrm{o}_p(1) \\
&= \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\operatorname{tr}(\mathbf{S} - \mathbf{S}_i)\mathbf{F}^{-1}(z)I(\mathcal{Q}_i) \\
&\quad + \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\operatorname{tr}(\mathbf{S}_i + \alpha_n\mathbf{T}_N)\big(\mathbf{F}^{-1} - \mathbf{F}_i^{-1}\big)I(\mathcal{Q}_i) + \mathrm{o}_p(1) \\
&= \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\eta_i(z)I(\mathcal{Q}_i) - \mathcal{D}_1 - \mathcal{D}_2 + \mathrm{o}_p(1),
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{D}_1 &= \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})(1-z)\varpi_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)\mathbf{r}_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)\mathbf{r}_i I(\mathcal{Q}_i), \\
\mathcal{D}_2 &= \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})(1-z)\varpi_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)(\mathbf{S}_i + \alpha_n\mathbf{T}_N)\mathbf{F}_i^{-1}(z)\mathbf{r}_i I(\mathcal{Q}_i).
\end{aligned}$$

Here we used (4.12) and the fact that

$$\mathbb{P}\big(I(\mathcal{Q}_i) \neq I(\mathcal{Q})\big) \le n^{-l}. \tag{4.39}$$

Check that

$$\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)\mathbf{r}_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)\mathbf{r}_i = \eta_i^2(z) + \frac{2}{n}\eta_i(z)\operatorname{tr}\mathbf{F}_i^{-1}(z) + \left(\frac{1}{n}\operatorname{tr}\mathbf{F}_i^{-1}(z)\right)^2.$$

Applying (4.4), (4.10) and Lemma 5.6 we obtain

$$\mathbb{E}\left|\mathcal{D}_1 - \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\left(\frac{2(1-z)\varpi_i^{\mathrm{tr}}\eta_i}{n}\operatorname{tr}\mathbf{F}_i^{-1} - \frac{(1-z)^2(\varpi_i^{\mathrm{tr}})^2\eta_i}{n^2}\big(\operatorname{tr}\mathbf{F}_i^{-1}\big)^2\right)I(\mathcal{Q}_i)\right|^2 = \mathrm{o}(1).$$

Similarly we have

$$\begin{aligned}
\mathbb{E}\Bigg|\mathcal{D}_2 - \sum_{i=1}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\bigg(&(1-z)\varpi_i^{\mathrm{tr}}K_i(z) \\
&- \frac{(1-z)^2(\varpi_i^{\mathrm{tr}})^2\eta_i}{n}\operatorname{tr}\mathbf{F}_i^{-2}(\mathbf{S}_i + \alpha_n\mathbf{T}_N)\bigg)I(\mathcal{Q}_i)\Bigg|^2 = \mathrm{o}(1),
\end{aligned}$$

where $K_i(z) = \mathbf{r}_i^* \mathbf{F}_i^{-1}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N)\mathbf{F}_i^{-1}(z)\mathbf{r}_i - n^{-1} \operatorname{tr} \mathbf{F}_i^{-2}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N)$. Thus we have

$$\begin{aligned}
&p\big(s_n(z) - \mathbb{E}s_n(z)\big) I(\mathcal{Q}) \\
&\quad = \sum_{i=1}^{n} \mathbb{E}_i \bigg(\big(\varpi_i^{\mathrm{tr}}\big)^2 \eta_i - (1-z)\varpi_i^{\mathrm{tr}} K_i(z) \\
&\qquad + \frac{(1-z)^2 (\varpi_i^{\mathrm{tr}})^2 \eta_i}{n} \operatorname{tr} \mathbf{F}_i^{-1}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N)\mathbf{F}_i^{-1}(z) \bigg) I(\mathcal{Q}_i) + \mathrm{o}_p(1).
\end{aligned}$$

Check that

$$\begin{aligned}
-\frac{\mathrm{d}(1-z)\varpi_i^{\mathrm{tr}}(z)\eta_i(z)}{\mathrm{d}z} &= -(1-z)\varpi_i^{\mathrm{tr}}(z) K_i(z) + \big(\varpi_i^{\mathrm{tr}}(z)\big)^2 \eta_i(z) \\
&\quad + \frac{(1-z)^2 (\varpi_i^{\mathrm{tr}}(z))^2 \eta_i(z)}{n} \operatorname{tr} \mathbf{F}_i^{-2}(z)(\mathbf{S}_i + \alpha_n \mathbf{T}_N),
\end{aligned}$$

which implies

$$\begin{aligned}
&\frac{1}{2\pi\mathrm{i}} \int_{\mathcal{C}^t} f(z) p\big(s_n(z) - \mathbb{E}s_n(z)\big) I(\mathcal{Q}_i)\,\mathrm{d}z \\
&\quad = -\frac{1}{2\pi\mathrm{i}} \sum_{i=1}^{n} \int_{\mathcal{C}^t} f(z) \mathbb{E}_i I(\mathcal{Q}_i)\,\mathrm{d}(1-z)\varpi_i^{\mathrm{tr}}(z)\eta_i(z) + \mathrm{o}_p(1).
\end{aligned}$$

Apparently, $\{\mathbb{E}_i I(\mathcal{Q}_i)\,\mathrm{d}(1-z)\varpi_i^{\mathrm{tr}}(z)\eta_i(z)/\mathrm{d}z\}$ is a martingale difference sequence so we can resort to the CLT for martingale (see Theorem 35.12 in [7]). By Lemma 5.4 and (4.9), we can get

$$\mathbb{E}\big|K_i(z) I(\mathcal{Q}_i)\big|^4 \le \frac{C\delta_n^4}{n},$$

which together with (4.10) and (4.8) implies

$$\sum_{k=1}^{n} \mathbb{E}\big|I(\mathcal{Q}_i)\,\mathrm{d}(1-z)\varpi_i^{\mathrm{tr}}(z)\eta_i(z)/\mathrm{d}z\big|^4 = \mathrm{O}(\delta_n) \to 0.$$

This ensures the Lyapunov condition. Thus, it is sufficient to investigate the limit of the following covariance function

$$-\frac{1}{4\pi^2} \int_{\mathcal{C}_1^t} \int_{\mathcal{C}_2^t} f(z_1) f(z_2) \frac{\partial^2}{\partial z_1\,\partial z_2} \mathcal{G}_n(z_1, z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2, \tag{4.40}$$

where

$$\mathcal{G}_n(z_1, z_2) = \sum_{i=1}^{n} \mathbb{E}_{i-1}\big[\mathbb{E}_i\big((1-z_1)\varpi_i^{\mathrm{tr}}(z_1)\eta_i(z_1) I(\mathcal{Q}_i)\big)\mathbb{E}_i\big((1-z_2)\varpi_i^{\mathrm{tr}}(z_2)\eta_i(z_2) I(\mathcal{Q}_i)\big).$$

From the arguments in [4], we need to show $\mathcal{G}_n(z_1, z_2)$ converges in probability. Applying (4.8), (4.10), (4.14) and the fact $\varpi_i^{\mathrm{tr}} = \varpi_i^{\mathbb{E}} - \varpi_i^{\mathrm{tr}}\varpi_i^{\mathbb{E}}\xi_i$, we have

$$\begin{aligned}&\mathcal{G}_n(z_1, z_2)\\&\quad = (1-z_1)(1-z_2)\sum_{i=1}^{n}\varpi_1^{\mathbb{E}}(z_1)\varpi_1^{\mathbb{E}}(z_2)\mathbb{E}_{i-1}\big[\mathbb{E}_i\big(\eta_i(z_1)I(\mathcal{Q}_i)\big)\mathbb{E}_i\big(\eta_i(z_2)I(\mathcal{Q}_i)\big)\big] + \mathrm{o}_p(1).\end{aligned}$$

By Lemma 5.5, we have

$$\begin{aligned}&\mathbb{E}_{i-1}\big[\mathbb{E}_i\big(\eta_i(z_1)I(\mathcal{Q}_i)\big)\mathbb{E}_i\big(\eta_i(z_2)I(\mathcal{Q}_i)\big)\big]\\&\quad = \frac{\mathbb{E}|x_{11}|^4 - |\mathbb{E}x_{11}^2| - 2}{n^2}\mathbb{E}_{i-1}\sum_{j=1}^{n}\big[\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_1)I(\mathcal{Q}_i)\big)_{jj}\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)I(\mathcal{Q}_i)\big)_{jj}\big] \qquad (4.41)\\&\qquad + \frac{\mathbb{E}x_{11}^2 + 1}{n^2}\mathbb{E}_{i-1}\operatorname{tr}\big[\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_1)I(\mathcal{Q}_i)\big)\mathbb{E}_i\big(\mathbf{F}_i^{-1}(z_2)I(\mathcal{Q}_i)\big)\big]. \qquad (4.42)\end{aligned}$$

Using (4.37), we have

$$(4.42) = \frac{\mathfrak{m}_x - \mathfrak{t} - 2}{n^2}\sum_{j=1}^{n}\big[\big(\mathbf{H}_1^{-1}(z_1)\big)_{jj}\big(\mathbf{H}_1^{-1}(z_2)\big)_{jj}\big] + \mathrm{o}_p(1).$$

It is worthy to remind the reader that in order to satisfy the condition in the last subsection we used here the fact

$$\mathbb{P}\big(I(\mathcal{Q}_i) \neq I(\mathcal{Q}_{ij})\big) \le n^{-l}.$$

And by (4.38), we have

$$(4.42) = \frac{\mathfrak{t}+1}{n}\frac{(1/n)\operatorname{tr}\mathbf{H}_1^{-1}(z_2)\mathbf{H}_1^{-1}(z_1)}{1 - (((i-1)(1-z_1)(1-z_2)\varpi_1^{\mathbb{E}}(z_1)\varpi_1^{\mathbb{E}}(z_2))/n^2)\operatorname{tr}\mathbf{H}_1^{-1}(z_2)\mathbf{H}_1^{-1}(z_1)} + \mathrm{o}_p(1).$$

From the arguments of the next part, we can conclude that for $z \in \mathcal{C}^t$

$$\mathbb{E}_0 s_n(z) = s_0(z) + \mathrm{O}\big(n^{-1}\big) \xrightarrow{\text{i.p.}} s(z).$$

Thus we get in probability

$$\begin{aligned}&\mathcal{G}_n(z_1, z_2)\\&\quad \to (\mathfrak{t}+1)\int\bigg[\bigg(\int\frac{y(1-z_1)(1-z_2)\varpi(z_1)\varpi(z_2)}{((1-z_1)\varpi - z_1\alpha t)((1-z_2)\varpi - z_2\alpha t)}\,\mathrm{d}F_{\mathrm{mp}}^{Y}(t)\bigg)\\&\qquad\qquad \times\bigg(1 - w\int\frac{y(1-z_1)(1-z_2)\varpi(z_1)\varpi(z_2)}{((1-z_1)\varpi - z_1\alpha t)((1-z_2)\varpi - z_2\alpha t)}\,\mathrm{d}F_{\mathrm{mp}}^{Y}(t)\bigg)^{-1}\bigg]\mathrm{d}w\\&\qquad + (\mathfrak{m}_x - \mathfrak{t} - 2)y\int\frac{(1-z_1)\varpi(z_1)}{(1-z_1)\varpi - z_1\alpha t}\,\mathrm{d}F_{\mathrm{mp}}^{Y}(t)\int\frac{(1-z_2)\varpi(z_2)}{(1-z_2)\varpi - z_2\alpha t}\,\mathrm{d}F_{\mathrm{mp}}^{Y}(t),\end{aligned}$$

which is (3.6).

In addition, by definition of $S_{n1}^{(1)}$ we get

$$\mathbb{E}\big|S_{n1}^{(1)}(z_1) - S_{n1}^{(1)}(z_2)\big|^2 = |z_1 - z_2|^2 \mathbb{E}\big|\operatorname{tr}\mathbf{D}^{-1}(z_1)\mathbf{D}^{-1}(z_2) - \mathbb{E}_0 \operatorname{tr}\mathbf{D}^{-1}(z_1)\mathbf{D}^{-1}(z_2)\big|^2 I(\mathcal{Q}).$$

Therefore using (4.12), Lemmas 4.1, 4.5 and the fact that

$$\mathbf{D}^{-1}(z) = (1-z)^{-1}\big(\mathbf{I} + \alpha_n \mathbf{T}_N \mathbf{F}^{-1}(z)\big),$$

we can easily check that

$$\mathbb{E}\big|S_{n1}^{(1)}(z_1) - S_{n1}^{(1)}(z_2)|^2 \le C|z_1 - z_2\big|^2, \qquad z_1, z_2 \in \mathcal{C}^t, \tag{4.43}$$

which implies the sequence $\{S_{n1}^{(1)}(\cdot)\}$ forms a tight sequence on $\mathcal{C}^t$.

4.2.2. *The mean part*

From the definition of the Stieltje transform of $s_n(z)$, we have

$$\begin{aligned} s_n(z) = s_{F^{\mathbf{B}_n}} &= \frac{1}{p}\operatorname{tr}\mathbf{D}^{-1} = \frac{1}{p}\operatorname{tr}(\mathbf{S}_n + \alpha_n \mathbf{T}_N)\mathbf{F}^{-1}(z) \\ &= \left(1 + \frac{1-z}{z}\right)\frac{1}{p}\operatorname{tr}\mathbf{S}_n\mathbf{F}^{-1}(z) - \frac{1}{z} = \frac{1}{zp}\operatorname{tr}\mathbf{S}_n\mathbf{F}^{-1}(z) - \frac{1}{z}. \end{aligned}$$

Using (4.11) we get that

$$\mathbf{S}_n\mathbf{F}^{-1}(z) = \sum_{i=1}^{n} \varpi_i \mathbf{r}_i \mathbf{r}_i^* \mathbf{F}_i^{-1}, \tag{4.44}$$

which implies

$$\frac{1}{p}\operatorname{tr}\mathbf{S}_n\mathbf{F}^{-1}(z) = \frac{n}{p(1-z)}\left(1 - \frac{1}{n}\sum_{i=1}^{n}\varpi_i\right).$$

Thus we have

$$\frac{1}{n}\sum_{i=1}^{n}\varpi_i = 1 - y_n(1-z)(zs_n + 1) \tag{4.45}$$

and

$$\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1) = 1 - y_n(1-z)\big(z\mathbb{E}_0 s_n I(\mathcal{Q}_1) + 1\big). \tag{4.46}$$

Denote $\mathbf{A}_n = \mathbb{E}_0(\varpi_1 I(\mathcal{Q}_1))(\mathbb{E}_0(\varpi_1 I(\mathcal{Q}_1))\mathbf{I} + \alpha_n \mathbf{T}_N)^{-1}$,

$$\mathbf{C}_n = \mathbf{A}_n - z\mathbf{I}\Delta(z) = \mathbb{E}_0 \quad \text{and} \quad s_n(z)I(\mathcal{Q}_1) - p^{-1}\operatorname{tr}\mathbf{C}_n.$$

Then we obtain that

$$p^{-1}\operatorname{tr}\mathbf{C}_n=\int\frac{\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)+\alpha_n t}{(1-z)\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)-z\alpha_n t}\,\mathrm{d}F^{\mathbf{T}_N}(t). \tag{4.47}$$

Recalling the definition of ϖ_0 and (2.9) we have

$$\frac{1-\varpi_0}{zy(1-z)}-\frac{1}{z}=\frac{1}{1-z}+\frac{1}{1-z}\int\frac{\alpha_n t}{(1-z)\varpi_0-z\alpha_n t}\,\mathrm{d}F^{\mathbf{T}_N}(t),$$

which implies

$$\varpi_0=\left(1+y\int\frac{(1-z)}{(1-z)\varpi_0-z\alpha_n t}\,\mathrm{d}F^{\mathbf{T}_N}(t)\right)^{-1}.$$

According to (4.46) and (4.47), we get that

$$\begin{aligned}&\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)\\&=\left(1+y\int\frac{(1-z)}{(1-z)\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)-z\alpha_n t}\,\mathrm{d}F^{\mathbf{T}_N}(t)+\left(\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)\right)^{-1}zy(1-z)\Delta_n\right)^{-1}.\end{aligned}$$

The difference of the above two identities yields

$$\begin{aligned}\varpi_0-\mathbb{E}\varpi_1 I(\mathcal{Q}_1)=&\int\frac{\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)\varpi_0 y(1-z)^2(\varpi_0-\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1))}{[(1-z)\varpi_0-z\alpha_n t][(1-z)\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)-z\alpha_n t]}\,\mathrm{d}F^{\mathbf{T}_N}(t)\\&+\varpi_0 zy(1-z)\Delta_n.\end{aligned}$$

Thus we use (4.39) to obtain that

$$\begin{aligned}&\mathbb{E}_0 s_n(z)I(\mathcal{Q})-s_0(z)\\&=\varpi_0\Delta_n\left(1-\int\frac{y_n(1-z)^2\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)\varpi_0}{[(1-z)\varpi_0-z\alpha_n t][(1-z)\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)-z\alpha_n t]}\,\mathrm{d}F^{\mathbf{T}_N}(t)\right)^{-1}.\end{aligned} \tag{4.48}$$

We will use the following lemma.

Lemma 4.10. *For $z\in\mathcal{C}_t$*

$$\begin{aligned}p\Delta(z)=&\frac{(\mathfrak{m}_x-\mathfrak{t}-2)\alpha_n(1-z)(\varpi_1^{\mathbb{E}})^2}{pn}\operatorname{tr}\mathbf{H}_0^{-1}(z)\operatorname{tr}\mathbf{H}_0^{-2}(z)\mathbf{T}_N\\&+\frac{\mathfrak{t}(1-z)(\varpi_1^{\mathbb{E}})^2\alpha_n}{n}\operatorname{tr}\mathbf{H}_0^{-3}(z)\mathbf{T}_N+\mathrm{o}(1).\end{aligned}$$

Proof. It follows from the definition of $\mathbf{D}_n$ and $\mathbf{C}_n$ that

$$\mathbf{D}_n^{-1}-\mathbf{C}_n^{-1}=\mathbf{C}_n^{-1}(\mathbf{A}_n-\mathbf{B}_n)\mathbf{D}_n^{-1}=\mathbf{C}_n^{-1}\mathbf{A}_n\mathbf{D}_n^{-1}-\mathbf{C}_n^{-1}\mathbf{B}_n\mathbf{D}_n^{-1}.$$

Using (4.44), we have

$$\mathbf{C}_n^{-1}\mathbf{B}_n\mathbf{D}_n^{-1}=\mathbf{C}^{-1}\sum_{i=1}^{n}\varpi_i\mathbf{r}_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)$$

and

$$\mathbf{C}_n^{-1}\mathbf{A}(\mathbf{B}_n-z\mathbf{I})^{-1}=\mathbf{C}_n^{-1}\mathbf{A}\sum_{i=1}^{n}\varpi_i\mathbf{r}_i\mathbf{r}_i^*\mathbf{F}_i^{-1}(z)+\alpha_n\mathbf{C}_n^{-1}\mathbf{A}\mathbf{T}_N\mathbf{F}^{-1}(z).$$

Then from the definition of $\Delta(z)$ and (4.11) we have

$$\begin{aligned}p\Delta_n&=n\mathbb{E}_0\varpi_1\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\mathbf{r}_1I(\mathcal{Q}_1)\\&\quad+\mathbb{E}_0\alpha_n\operatorname{tr}\mathbf{A}\mathbf{T}_N\mathbf{F}^{-1}(z)\mathbf{C}^{-1}I(\mathcal{Q}_1)-n\mathbb{E}_0\varpi_1\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{r}_iI(\mathcal{Q}_1)\\&=d_1+d_2+d_3+d_4,\end{aligned}$$

where

$$\begin{aligned}d_1&=n\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\mathbf{r}_1-\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A},\\d_2&=\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}-\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}^{-1}(z)\mathbf{C}^{-1}\mathbf{A},\\d_3&=\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}-n\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{r}_1,\\d_4&=\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}^{-1}(z)\mathbf{C}^{-1}-\mathbb{E}_0\varpi_1I(\mathcal{Q}_1)\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}.\end{aligned}$$

First, consider d_1. We apply (4.3) and (4.2) to represent d_1 as

$$\begin{aligned}d_1&=-n(1-z)\big(\varpi_1^{\mathbb{E}}\big)^2\mathbb{E}_0\eta_1\big(\mathbf{r}_1^*\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}\mathbf{r}_1-n^{-1}\operatorname{tr}\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}\big)I(\mathcal{Q}_1)&(4.49)\\&\quad+(1-z)\big(\varpi_1^{\mathbb{E}}\big)^2\mathbb{E}_0\xi_1\big(\operatorname{tr}\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}-\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}\big)I(\mathcal{Q}_1)&(4.50)\\&\quad+n(1-z)^2\big(\varpi_1^{\mathbb{E}}\big)^2\big(\mathbb{E}_0\varpi_1\gamma_1^2\mathbf{r}_1^*\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}\mathbf{r}_1\\&\qquad-n^{-1}\mathbb{E}_0\varpi_1\gamma_1^2\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}\mathbf{C}^{-1}\mathbf{A}\big)I(\mathcal{Q}_1).&(4.51)\end{aligned}$$

Note that similar to (4.18) we can get that $\|\mathbf{C}^{-1}\|$ and $\|\mathbf{A}\mathbf{C}^{-1}\|$ are both bounded when $z\in\mathcal{C}^t$. Thus by Lemma 4.1 and Lemma 5.4 we obtain that

$$\mathbb{E}\big|\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}-\mathbb{E}_0\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\big|^2=\mathrm{O}(1),$$

$$\mathbb{E}\big|\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\mathbf{r}_1-n^{-1}\operatorname{tr}\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\big|^2=\mathrm{O}\big(n^{-1}\big).$$

These together with (4.10), (4.8), (4.14) and Hölder's inequalityimply that

$$(4.50)=\mathrm{O}_p\big(n^{-1/2}\big)\quad\text{and}\quad(4.51)=\mathrm{O}_p\big(\delta_n^2\big).$$

Using Lemma (5.5), we have

$$(4.49) = -(\mathfrak{m}_x - \mathfrak{t} - 2)(1-z)\big(\varpi_1^{\mathbb{E}}\big)^2 y_n \mathbb{E}_0\big(\mathbf{F}_1^{-1}(z)\big)_{11}\big(\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\big)_{11} I(\mathcal{Q}_1) - \frac{(\mathfrak{t}+1)(1-z)(\varpi_1^{\mathbb{E}})^2}{n}\mathbb{E}_0 \operatorname{tr} \mathbf{F}_1^{-2}(z)\mathbf{C}^{-1}\mathbf{A} I(\mathcal{Q}_1).$$

For d_2, we use (4.12) to get

$$\begin{aligned} d_2 &= (1-z)\mathbb{E}_0\varpi_1\mathbb{E}_0\big(\varpi_1 I(\mathcal{Q}_1)\mathbf{r}_1^*\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\mathbf{A}\mathbf{F}_1^{-1}(z)I(\mathcal{Q}_1)\mathbf{r}_1\big) \\ &= \frac{(1-z)(\varpi_1^{\mathbb{E}})^2}{n} \operatorname{tr} \mathbf{F}_1^{-2}(z)\mathbf{C}^{-1}\mathbf{A} I(\mathcal{Q}_1) + \mathrm{o}_p(1). \end{aligned}$$

Similarly, we can get

$$\begin{aligned} d_3 &= (\mathfrak{m}_x - \mathfrak{t} - 2)(1-z)\big(\varpi_1^{\mathbb{E}}\big)^2 y_n \mathbb{E}_0\big(\mathbf{F}_1^{-1}(z)\big)_{11}\big(\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}\big)_{11} I(\mathcal{Q}_1) \\ &\quad + \frac{(\mathfrak{t}+1)(1-z)(\varpi_1^{\mathbb{E}})^2}{n}\mathbb{E}_0 \operatorname{tr} \mathbf{F}_1^{-2}(z)\mathbf{C}^{-1} I(\mathcal{Q}_1) + \mathrm{o}_p(1) \end{aligned}$$

and

$$d_4 = -\frac{(1-z)(\varpi_1^{\mathbb{E}})^2}{n} \operatorname{tr} \mathbf{F}_1^{-2}(z)\mathbf{C}^{-1} + \mathrm{o}_p(1).$$

Therefore combining the above four equations, we conclude that

$$\begin{aligned} &\mathbb{E}_0 \operatorname{tr} \mathbf{D}_n^{-1} I(\mathcal{Q}_1) - \operatorname{tr} \mathbf{C}_n^{-1} \\ &\quad = \frac{(\mathfrak{m}_x - \mathfrak{t} - 2)(1-z)(\varpi_1^{\mathbb{E}})^2 p}{n}\mathbb{E}_0\big(\mathbf{F}_1^{-1}(z)\big)_{11}\big(\mathbf{F}_1^{-1}(z)\mathbf{C}^{-1}(\mathbf{I}-\mathbf{A})\big)_{11} I(\mathcal{Q}_1) \\ &\qquad + \frac{\mathfrak{t}(1-z)(\varpi_1^{\mathbb{E}})^2}{n}\mathbb{E}_0 \operatorname{tr} \mathbf{F}_1^{-2}(z)\mathbf{C}^{-1}(\mathbf{I}-\mathbf{A}) I(\mathcal{Q}_1) + \mathrm{o}_p(1). \end{aligned}$$

By Lemmas 4.3–4.5 and the fact that

$$\big\|\mathbf{C}^{-1}(\mathbf{I}-\mathbf{A})\big\| = \big|\alpha_n \mathbf{T}_N\big((1-z)\mathbb{E}_0\varpi_1 I(\mathcal{Q}_1)\mathbf{I} - z\alpha_n\mathbf{T}_N\big)^{-1}\big| \le C,$$

we have that

$$\begin{aligned} \mathbb{E}_0 \operatorname{tr} \mathbf{D}_n^{-1} - \operatorname{tr} \mathbf{C}_n^{-1} &= \frac{(\mathfrak{m}_x - \mathfrak{t} - 2)\alpha_n(1-z)(\varpi_1^{\mathbb{E}})^2}{pn} \operatorname{tr} \mathbf{H}_0^{-1}(z) \operatorname{tr} \mathbf{H}_0^{-2}(z)\mathbf{T}_N \\ &\quad + \frac{\mathfrak{t}(1-z)(\varpi_1^{\mathbb{E}})^2\alpha_n}{n} \operatorname{tr} \mathbf{H}_0^{-3}(z)\mathbf{T}_N + \mathrm{o}_p(1), \end{aligned}$$

which complete the proof of this lemma. □

Noting the transform $\ddot{s}_n(z) = \frac{y}{(1+z)^2} s_n(\frac{z}{1+z}) - \frac{1-y+z}{z(1+z)}$, $\ddot{s}_0(z) = \frac{y}{(1+z)^2} s_0(\frac{z}{1+z}) - \frac{1-y+z}{z(1+z)}$ and (3.12) in [3] we have that for $z \in \mathcal{C}_t$

$$\left\| \left(1 - \int \frac{\alpha_n y_n z(1-z)t}{[(1-z)\mathbb{E}\varpi_1 - z\alpha_n t][(1-z)\varpi_0 - z\alpha_n t]} \, \mathrm{d}F^{\mathbf{T}_N}(t)\right)^{-1} \right\| \le C_\theta.$$

Thus we have $\mathbb{E}s_n = s_0 + \mathrm{O}(n^{-1}) \to s$, which combined with (4.48) gives (3.5).

We so far have proved Lemma 3.1 under the condition that $z \in \mathcal{C}^t$. It is easy to check that the above arguments evidently work when z belongs to the bottom line due to symmetry.

When z belongs to the left vertical line of the contour, that is $z \in \mathcal{C}^l = \{\Re z = c_l - \theta, \Im z \in [-\theta, \theta]\}$, we split $\mathcal{C}^l$ into two parts $\mathcal{C}^l_1 + \mathcal{C}^l_2$ where

$$\mathcal{C}^l_1 = \left\{\Re z = c_l - \theta, n^{-1}\varepsilon_n < |\Im z| < \theta\right\} \quad \text{and} \quad \mathcal{C}^l_2 = \left\{\Re z = c_l - \theta, |\Im z| < n^{-1}\varepsilon_n\right\}$$

with $\varepsilon_n = n^{-\beta}$ for some $\beta \in (0, 1)$. We truncate s_n at each part, that is

$$\hat{s}_n(z) = \begin{cases} s_n(z), & z \in \mathcal{C}^l_1; \\ s_n\big(\Re z + \mathrm{i}n^{-1}\varepsilon_n\big), & z \in \mathcal{C}^l_2. \end{cases}$$

Then from a similar argument in [4] we can get that the limit of $p(\hat{s}_n(z)I(\mathcal{Q}) - s_0)$ has the same form as Lemma 3.1 provided. Here $\mathcal{Q} = \{\|(\mathbf{S}_n + \alpha_n \mathbf{T}_N)^{-1}\| \le C\} \cap \{\lambda_1^{\mathbf{S}_n} > c_l - \iota\}$ with small enough $\iota > 0$. And the situation is the same if z belongs to the right vertical line of the contour due to symmetry. We omit the details.

5. Some basic lemmas

In this section, we give some basic lemmas which are used in the paper.

Lemma 5.1 (Lemma 6.1 in [16]).

$$\begin{aligned} z &= -\frac{\ddot{s}(z)(\ddot{s}(z) + 1 - y)}{(\ddot{s}(z) + 1/(1-Y))(1-Y)}, \\ \ddot{s}(z) &= \frac{(\ddot{s}(z) + 1/(1-Y))(1-Y)}{\ddot{s}(z)(\ddot{s}(z) + 1)}, \\ \big(\ddot{s}(z)\big)' &= -\frac{(\ddot{s}(z) + 1/(1-Y))^2(1-Y)^2}{(1-Y)\ddot{s}(z)^2 + 2\ddot{s}(z) + 1 - y}, \\ \int \frac{1}{\ddot{s}(z) + t} \, \mathrm{d}F^Y_{\mathrm{mp}}(t) &= \frac{\ddot{s}(z)}{(\ddot{s}(z) + 1/(1-Y))(1-Y)}, \\ \int t\big(\ddot{s}(z) + t\big)^{-2} \, \mathrm{d}F^Y_{\mathrm{mp}}(t) &= \frac{(\ddot{s}(z))^2}{(1-Y)\ddot{s}(z)^2 + 2\ddot{s}(z) + 1}, \end{aligned}$$

$$\ddot{s}'(z) = -\frac{(1-Y)\ddot{s}(z)^2 + 2\ddot{s}(z) + 1}{(\ddot{s}(z))^2(\ddot{s}(z)+1)^2}\big(\ddot{s}(z)\big)' = -\big(1 - Y\big(\ddot{s}(z)\big)^2\big(\ddot{s}(z)+1\big)^2\big)\ddot{s}^{-2}(z)\big(\ddot{s}(z)\big)',$$

$$\frac{2y\int \alpha t(\ddot{s}(z))^3(\ddot{s}(z)+t)^{-3}\,\mathrm{d}F^Y_{\mathrm{mp}}(t)}{(1 - y\int(\ddot{s}(z))^2(\ddot{s}(z)+t)^{-2}\,\mathrm{d}F^Y_{\mathrm{mp}}(t))^2} = -\left(\log\frac{(1-Y)\ddot{s}(z)^2 + 2\ddot{s}(z) + 1 - y}{(1-Y)\ddot{s}(z)^2 + 2\ddot{s}(z)+1}\right)',$$

$$\frac{2Y\ddot{s}'(z)(\ddot{s}(z))^3(\ddot{s}(z)+t)^{-3}}{((1-Y(\ddot{s}(z))^2(\ddot{s}(z)+1)^2))^2} = \big(\log\big(1 - Y\big(\ddot{s}(z)\big)^2\big(\ddot{s}(z)+1\big)^2\big)\big)'.$$

Lemma 5.2 (Lemma 2.3 in [13]). *Let x, y be arbitrary non-negative numbers. For $\mathbf{A}$ and $\mathbf{B}$ square matrices of the same size,*

$$F^{\sqrt{(\mathbf{AB})(\mathbf{AB})^*}}\big\{(xy,\infty)\big\} \le F^{\sqrt{\mathbf{AA}^*}}\big\{(x,\infty)\big\} + F^{\sqrt{\mathbf{BB}^*}}\big\{(y,\infty)\big\}.$$

Lemma 5.3 (Lemma A.45 and Corollary A.41 in [2]). *Let $\mathbf{A}$ and $\mathbf{B}$ be two $n\times n$ Hermitian matrices. Then*

$$L\big(F^{\mathbf{A}}, F^{\mathbf{B}}\big) \le \|\mathbf{A}-\mathbf{B}\| \quad \text{and} \quad L^3\big(F^{\mathbf{A}}, F^{\mathbf{B}}\big) \le \frac{1}{n}\operatorname{tr}(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*,$$

where $L(\cdot,\cdot)$ denotes the Lévy distance and $\|\cdot\|$ denotes the spectral norm.

Lemma 5.4 (Lemma 9.1 of [2]). *Let $\mathbf{A}$ be an $n\times n$ nonrandom matrix bounded in norm by M, and $X = (x_1,\dots,x_n)^*$ be a random vector of independent entries. Assume that $\mathbb{E}x_i = 0$, $\mathbb{E}|x_i|^2 = 1$, $\mathbb{E}|x_j|^4 < \infty$ and $|x_i| \le \delta_n\sqrt{n}$ with $\delta_n \to 0$ slowly. Then for any given $2 \le l \le b\log(n\delta_n^2)$ with some $b > 1$, there exists a constant C such that*

$$\mathbb{E}\big|X^*\mathbf{A}X - \operatorname{tr}\mathbf{A}\big|^l \le n^l\big(n\delta_n^4\big)^{-1}\big(MC\delta_n^2\big)^l.$$

Lemma 5.5 ((1.15) of [4]). *Let $\mathbf{A} = (a_{ij})_{p\times p}$ and $\mathbf{B} = (b_{ij})_{p\times p}$ be nonrandom matrices and $X = (x_1,\dots,x_n)^*$ be a random vector of independent entries. Assume that $\mathbb{E}x_i = 0$ and $\mathbb{E}|x_i|^2 = 1$. Then we have,*

$$\mathbb{E}\big(X^*\mathbf{A}X - \operatorname{tr}\mathbf{A}\big)\big(X^*\mathbf{B}X - \operatorname{tr}\mathbf{B}\big) = \sum_{i=1}^{p}\big(\mathbb{E}|x_i|^4 - |\mathbb{E}x_i^2|^2 - 2\big)a_{ii}b_{ii} + \operatorname{tr}\mathbf{A}_x\mathbf{B}_x^T + \operatorname{tr}\mathbf{AB}, \tag{5.1}$$

where $\mathbf{A}_x = (\mathbb{E}x_i^2 a_{ij})_{p\times p}$, $\mathbf{B}_x = (\mathbb{E}x_i^2 b_{ij})_{p\times p}$ and the superscript ^{T} is the transpose of a matrix.

Lemma 5.6 (Burkholder inequality). *Let $\{X_k\}$ be a complex martingale difference sequence with respect to the increasing σ-field $\mathcal{F}_k$, and let $\mathbb{E}_k$ denote conditional expectation with respect to $\mathcal{F}_k$. Then we have*

(a) *for $p > 1$,*

$$\mathbb{E}\left|\sum_{k=1}^{n} X_k\right|^p \le K_p\mathbb{E}\left(\sum_{k=1}^{n}|X_k|^2\right)^{p/2}; \tag{5.2}$$

(b) *for* $p \geq 2$,

$$\mathbb{E}\left|\sum_{k=1}^{n} X_k\right|^p \leq K_p^* \left(\mathbb{E}\left(\sum_{k=1}^{n} \mathbb{E}_{k-1}|X_k|^2\right)^{p/2} + \mathbb{E}\sum_{k=1}^{n}|X_k|^p\right). \tag{5.3}$$

Lemma 5.7. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $n \times n$ *non-negative definite Hermitian matrices.* $\lambda_i^{\mathbf{A}}$ *and* $\lambda_i^{\mathbf{B}}$ *denote the* i*th smallest eigenvalue of* $\mathbf{A}$ *and* $\mathbf{B}$, *respectively. Then we have*

$$\lambda_1^{\mathbf{A}}\lambda_i^{\mathbf{B}} \leq \lambda_i^{\mathbf{AB}} \leq \lambda_n^{\mathbf{A}}\lambda_i^{\mathbf{B}} \quad \textit{and} \quad \lambda_i^{\mathbf{A}}\lambda_1^{\mathbf{B}} \leq \lambda_i^{\mathbf{AB}} \leq \lambda_i^{\mathbf{A}}\lambda_n^{\mathbf{B}}, \qquad i = 1, \dots, n.$$

Acknowledgements

The authors would like to thank the referee for many constructive comments. Zhidong Bai was partially supported by CNSF 11171057, Fundamental Research Funds for the Central Universities, NUS Grant R-155-000-141-112. Jiang Hu was partially supported by CNSF 11301063 and Fundamental Research Funds for the Central Universities. Guangming Pan and Wang Zhou were partially supported by the Ministry of Education, Singapore, under grant # ARC 14/11. Wang Zhou was also partially supported by a grant R-155-000-116-112 at the National University of Singapore.

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR1990662

[2] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2567175

[3] Bai, Z.D. and Silverstein, J.W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051

[4] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792

[5] Bai, Z.D., Yin, Y.Q. and Krishnaiah, P.R. (1987). On limiting empirical distribution function of the eigenvalues of a multivariate F matrix. *Teor. Veroyatn. Primen.* **32** 537–548. MR0914942

[6] Bai, Z.D. and Zhang, L.X. (2010). The limiting spectral distribution of the product of the Wigner matrix and a nonnegative definite matrix. *J. Multivariate Anal.* **101** 1927–1949. MR2671192

[7] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1324786

[8] Dumitriu, I. and Paquette, E. (2012). Global fluctuations for linear statistics of Beta–Jacobi ensembles. *Random Matrices Theory Appl.* **1** 1250013, 60. MR3039374

[9] Fujikoshi, Y., Ulyanov, V.V. and Shimizu, R. (2010). *Multivariate Statistics*. *High-Dimensional and Large-Sample Approximations*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR2640807

[10] Marčenko, V.A. and Pastur, L.A. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sb.* **1** 457–483.

[11] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. *Wiley Series in Probability and Mathematical Statistics* **42**. New York: Wiley. MR0652932

[12] Silverstein, J.W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal*. **55** 331–339. MR1370408

[13] Silverstein, J.W. and Bai, Z.D. (1995). On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal*. **54** 175–192. MR1345534

[14] Silverstein, J.W. and Choi, S.-I. (1995). Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal*. **54** 295–309. MR1345541

[15] Yin, Y.Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal*. **20** 50–68. MR0862241

[16] Zheng, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Ann. Inst. Henri Poincaré Probab. Stat*. **48** 444–476. MR2954263

Received August 2012 and revised November 2013