The Annals of Applied Probability
2011, Vol. 21, No. 5, 1860–1910
DOI: 10.1214/10-AAP742

CENTRAL LIMIT THEOREM FOR HOTELLING'S T^2 STATISTIC UNDER LARGE DIMENSION

BY G. M. PAN[1] AND W. ZHOU[2]

Nanyang Technological University and National University of Singapore

Dedicated to Z. D. Bai on the occasion of his 65th birthday

In this paper we prove the central limit theorem for Hotelling's T^2 statistic when the dimension of the random vectors is proportional to the sample size.

1. Introduction and main results. Since the famous Marčenko and Pastur law was found in [16], the theory of large sample covariance matrices has been further developed. Among others, we mention Jonsson [14], Yin [24], Silverstein [18], Watcher [22], Yin, Bai and Krishanaiah [25]. Lately, Johnstone [13] discovered the law of the largest eigenvalue of the Wishart matrix, Bai and Silverstein [5] established the central limit theorems (CLT) of linear spectral statistics and Bai, Miao and Pan [3] derived CLT for functionals of the eigenvalues and eigenvectors. We also refer to [9, 12, 21] for CLT on linear statistics of eigenvalues of other classes of random matrices.

The sample covariance matrix is defined by

$$\mathcal{S} = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{s}_j - \bar{\mathbf{s}})(\mathbf{s}_j - \bar{\mathbf{s}})^T,$$

where $\bar{\mathbf{s}} = n^{-1}\sum_{j=1}^{n}\mathbf{s}_j$ and $\mathbf{s}_j = (X_{1j}, \dots, X_{pj})^T$. Here $\{X_{ij}\}$, $i, j = \cdots$, is a double array of independent and identically distributed (i.i.d.) real r.v.'s with $EX_{11} = 0$ and $EX_{11}^2 = 1$. However, in the large random matrices theory (RMT) the commonly used sample covariance matrix is

$$\mathbf{S} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{s}_j\mathbf{s}_j^T = \frac{1}{n}\mathbf{X}_n\mathbf{X}_n^T,$$

where $\mathbf{X}_n = (\mathbf{s}_1, \dots, \mathbf{s}_n)$.

Received November 2009; revised September 2010.

[1]Supported in part by Grant M58110052 at the Nanyang Technological University.

[2]Supported in part by Grant R-155-000-106-112 at the National University of Singapore.

MSC2010 subject classifications. Primary 15B52, 60F15, 62E20; secondary 60F17.

Key words and phrases. Hotelling's T^2 statistic, sample means, sample covariance matrices, central limit theorem, Stieltjes transform.

Note that $\mathcal{S} = \mathbf{S} - \bar{\mathbf{s}}\bar{\mathbf{s}}^T$ and thus, by the rank inequality, there is no difference when one is only concerned with the limiting empirical spectral distribution (ESD) of the eigenvalues in large random matrices. Therefore, the limiting ESD of $\mathcal{S}$ is Marčenko and Pastur's law $F_c(x)$ (see [14] and [16]) when $\lim \frac{p}{n} = c > 0$ which has a density function

$$p_c(x) = \begin{cases} (2\pi cx)^{-1}\sqrt{(b-x)(x-a)}, & a \le x \le b, \\ 0, & \text{otherwise,} \end{cases}$$

and has point mass $1 - c^{-1}$ at the origin if $c > 1$, where $a = (1-\sqrt{c})^2$ and $b = (1+\sqrt{c})^2$. The Stieljes transform $m(z)$ of $F_c(x)$ satisfies the equation (see [20])

$$m(z) = \frac{1}{1 - c - czm(z) - z}, \tag{1.1}$$

where the Stieljes transform for any function $G(x)$ is defined by

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

Observe that the spectra of $n^{-1}\mathbf{X}_n\mathbf{X}_n^T$ and $n^{-1}\mathbf{X}_n^T\mathbf{X}_n$ are identical except for zero eigenvalues. This leads to the equality

$$\underline{m}_n^{\mathbf{S}}(z) = -\frac{1 - p/n}{z} + \frac{p}{n} m_n^{\mathbf{S}}(z) \tag{1.2}$$

and therefore,

$$z = -\frac{1}{\underline{m}(z)} + \frac{c}{1 + \underline{m}(z)}, \tag{1.3}$$

where $m_n^{\mathbf{S}}(z)$ and $\underline{m}_n^{\mathbf{S}}(z)$ denote, respectively, the Stieljes transform of the ESD of $n^{-1}\mathbf{X}_n\mathbf{X}_n^T$ and $n^{-1}\mathbf{X}_n^T\mathbf{X}_n$ and, correspondingly, $\underline{m}(z)$ is the limit of $\underline{m}_n^{\mathbf{S}}(z)$.

Sample covariance matrices are also of essential importance in multivariate statistical analysis because many test statistics involve their eigenvalues and/or eigenvectors. The typical example is T^2 statistic which was proposed by Hotelling [10]. We refer to [1] and [15] for various uses of the T^2 statistic.

The T^2 statistic, which is the origin of multivariate linear hypothesis tests and the associated confidence sets, is defined by

$$T^2 = n(\bar{\mathbf{s}} - \boldsymbol{\mu}_0)^T \mathcal{S}^{-1}(\bar{\mathbf{s}} - \boldsymbol{\mu}_0), \tag{1.4}$$

whose distribution is invariant if each $\mathbf{s}_j$ is replaced by $\boldsymbol{\Sigma}^{1/2}\mathbf{s}_j$ with $\boldsymbol{\Sigma}$ any nonsingular p by p matrix when $\boldsymbol{\mu}_0 = 0$. If $\{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ is a sample from the p-dimensional population $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $[T^2/(n-1)][(n-p)/p]$ follows a noncentral F distribution and moreover, the F distribution is central if $\boldsymbol{\mu} = \boldsymbol{\mu}_0$. When p is fixed, the limiting distribution of T^2 for $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ is the χ^2-distribution even if the parent distribution is not normal.

In the recent three or four decades in many research areas, including signal processing, network security, image processing, genetics, stock marketing and other economic problems, people are interested in the case where p is quite large or proportional to the sample size. Thus, it will be desirable if one can obtain the asymptotic distribution of the famous Hotelling T^2 statistic when the dimension of the random vectors is proportional to the sample size. It is the aim of this work. In addition, we would like to point out that some discussions about the two-sample T^2 statistic under the assumption that the underlying r.v.'s are normal were presented in [2].

The main results are presented in the following theorems.

THEOREM 1. *Suppose that*:

(1) *for each* n $X_{ij} = X_{ij}^n, i, j = 1, 2, \ldots,$ *are i.i.d. real r.v.'s with* $EX_{11} = \mu, EX_{11}^2 = 1$ *and* $EX_{11}^4 < \infty$.

(2) $p \le n, c_n = p/n \to c \in (0, 1)$ *as* $n \to \infty$.

Then, when $\boldsymbol{\mu}_0 = (\mu, \ldots, \mu)^T$,

$$\frac{\sqrt{n}}{\sqrt{2c_n(1-c_n)^{-3}}}\left(\frac{T^2}{n} - c_n(1-c_n)^{-1}\right) \xrightarrow{D} N(0,1),$$

where $F_{c_n}(x)$ *denotes* $F_c(x)$ *by substituting* c_n *for* c.

REMARK 1. *When* $X_{ij} \sim N(0,1)$, *it is well known that* $(n-p)T^2/(np)$ *follows* F *distribution with degrees of freedom* p *and* $n-p$, *respectively. As* $n \to \infty$ *and* $p/n \to c$, *it follows from strong law of large numbers and CLT that*

$$\frac{(n-p)T^2/(np) - 1}{\sqrt{2/p + 2/(n-p)}} \longrightarrow N(0,1).$$

This is consistent with Theorem 1.

REMARK 2. *Since* $\int x^{-1}\,dF_c(x) = (1-c)^{-1}$ *and* $\int x^{-2}\,dF_c(x) = (1-c)^{-3}$ *which are derived through differentiating the following identity* [*the Stieljes transform* $m(z)$ *of* $F_c(x)$],

$$\int (x-z)^{-1} p_c(x)\,dx = \frac{-(z+c-1) + (z+c-1)\sqrt{1-4zc(z+c-1)^{-2}}}{2cz},$$

we actually prove that

$$\frac{\sqrt{n}}{\sqrt{2c_n\int x^{-2}\,dF_{c_n}(x)}}\left(\frac{T^2}{n} - c_n\int\frac{dF_{c_n}(x)}{x}\right) \xrightarrow{D} N(0,1).$$

One typical application of Theorem 1 lies in making inference on the large-dimensional mean vector of the multivariate model

$$\mathbf{Z}_j = \Gamma \mathbf{s}_j + \boldsymbol{\mu}, \qquad E\mathbf{s}_j = 0, \qquad j = 1, \ldots, n,$$

where Γ is an m by p matrix, $m \le p$. This model means that each $\mathbf{Z}_j$ is a linear transformation of some p-variate random vector $\mathbf{s}_j$. It can generate a rich collection of $\mathbf{Z}_j$ from $\mathbf{s}_j$ with the given covariance matrix $\boldsymbol{\Sigma} = \Gamma\Gamma^T$. Most important, it includes the multivariate normal model.

We will prove Theorem 1 by establishing Theorem 2 which presents asymptotic distributions of random quadratic forms involving sample means and sample covariance matrices.

For any analytic function $f(\cdot)$, define

$$f(\boldsymbol{\mathcal{S}}) = \mathbf{U}^T \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_p))\mathbf{U},$$

where $\mathbf{U}^T \operatorname{diag}(\lambda_1, \ldots, \lambda_p)\mathbf{U}$ denotes the spectral decomposition of the matrix $\boldsymbol{\mathcal{S}}$.

THEOREM 2. *In addition to the assumption* (1) *of Theorem* 1, *suppose that* $c_n = p/n \to c > 0$, $EX_{11} = 0$, $g(x)$ *is a function with a continuous first derivative in a neighborhood of* c *and* $f(x)$ *is analytic on an open region containing the interval*

$$[I_{(0,1)}(c)(1-\sqrt{c})^2, (1+\sqrt{c})^2]. \tag{1.5}$$

Then,

$$\left(\sqrt{n}\left[\frac{\bar{\mathbf{s}}^T f(\boldsymbol{\mathcal{S}})\bar{\mathbf{s}}}{\|\bar{\mathbf{s}}\|^2} - \int f(x)\,dF_{c_n}(x)\right], \sqrt{n}\big(g(\bar{\mathbf{s}}^T\bar{\mathbf{s}}) - g(c_n)\big)\right) \xrightarrow{D} (X, Y),$$

where $Y \sim N(0, 2c(g'(c))^2)$, *which is independent of* X, *a Gaussian r.v. with* $EX = 0$ *and*

$$\operatorname{Var}(X) = \frac{2}{c}\left(\int f^2(x)\,dF_c(x) - \left(\int f(x)\,dF_c(x)\right)^2\right). \tag{1.6}$$

REMARK 3. *Let* $\mathbf{x}_n = (x_{n1}, \ldots, x_{np})^T \in \mathbb{R}^p$, $\|\mathbf{x}_n\| = 1$ *where* $\|\cdot\|$ *denotes the Euclidean norm. Note that*, *when* $\max_i x_{ni} \to 0$ (*see* [17], (1.16), *or* [19]),

$$\sqrt{n}\left[\mathbf{x}_n^T f(\mathbf{S})\mathbf{x}_n - \int f(x)\,dF_{c_n}(x)\right] \xrightarrow{D} X. \tag{1.7}$$

This suggests that $\bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$ *can be viewed as a fixed unit vector* $\mathbf{x}_n$ *when dealing with* $\bar{\mathbf{s}}^T f(\boldsymbol{\mathcal{S}})\bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|^2$ *even if* $\bar{\mathbf{s}}$ *is not independent of* $\boldsymbol{\mathcal{S}}$.

Theorem 2 relies on Lemma 1 below which deals with the asymptotic joint distribution of

$$X_n(z) = \sqrt{n}\left[\frac{\bar{\mathbf{s}}^T(\boldsymbol{\mathcal{S}} - z\mathbf{I})^{-1}\bar{\mathbf{s}}}{\|\bar{\mathbf{s}}\|^2} - m_n(z)\right], \qquad Y_n = \sqrt{n}\big(g(\bar{\mathbf{s}}^T\bar{\mathbf{s}}) - g(c_n)\big),$$

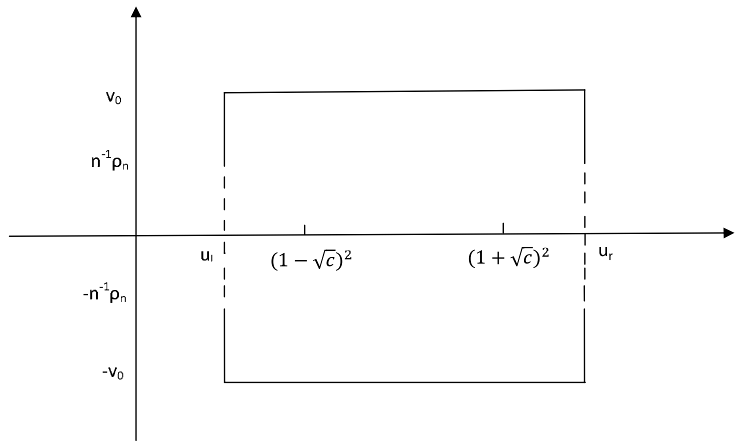

FIG. 1. *Contour $\mathcal{C}$ when $c < 1$.*

where $m_n(z) = \int (x-z)^{-1} dF_{c_n}(x)$. The stochastic process $X_n(z)$ is defined on a contour $\mathcal{C}$, given below. Let $v_0 > 0$ be arbitrary and set $\mathcal{C}_u = \{u + iv_0, u \in [u_l, u_r]\}$, where u_l is any negative number if the left endpoint of (1.5) is zero, otherwise u_l is any positive number smaller than the left endpoint of (1.5) and u_r any number larger than the right endpoint of (1.5). Then define

$$\mathcal{C}^+ = \{u_l + iv : v \in [0, v_0]\} \cup \mathcal{C}_u \cup \{u_r + iv : v \in [0, v_0]\}$$

and let $\mathcal{C}^-$ be the symmetric part of $\mathcal{C}^+$ about the real axis. Then set $\mathcal{C} = \mathcal{C}^+ \cup \mathcal{C}^-$. See Figures 1 and 2 for a picture of the contour $\mathcal{C}$ when $c < 1$ and $c \geq 1$, respectively.

Let $\mathbf{A}^{-1}(z) = (\mathbf{S} - zI)^{-1}$. Since it is difficult to control the spectral norm of $(\boldsymbol{\mathcal{S}} - z\mathbf{I})^{-1}$ or $\mathbf{A}^{-1}(z)$ on the whole contour $\mathcal{C}$, especially for $v = 0$, we further

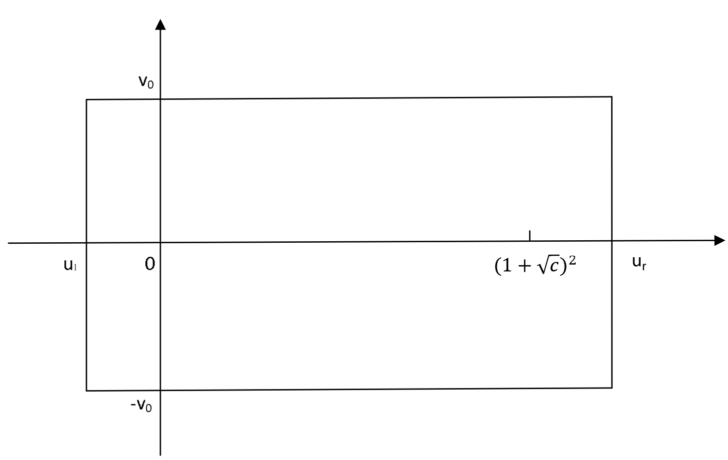

FIG. 2. *Contour $\mathcal{C}$ when $c \geq 1$.*

define $\hat{X}_n(z)$, a truncated version of $X_n(z)$, as in [5]. Select a sequence of positive numbers ρ_n satisfying for some $\beta \in (0, 1)$,

$$\rho_n \downarrow 0, \qquad \rho_n \geq n^{-\beta}. \tag{1.8}$$

Let

$$\mathcal{C}_l = \begin{cases} \{u_l + iv : v \in [n^{-1}\rho_n, v_0]\}, & \text{if } u_l > 0, \\ \{u_l + iv : v \in [0, v_0]\}, & \text{if } u_l < 0, \end{cases}$$

and

$$\mathcal{C}_r = \{u_r + iv : v \in [n^{-1}\rho_n, v_0]\}.$$

Write $\mathcal{C}_n^+ = \mathcal{C}_l \cup \mathcal{C}_u \cup \mathcal{C}_r$. We can now define the truncated process for $z = u + iv \in \mathcal{C}$ by

$$\hat{X}_n(z) = \begin{cases} X_n(z), & \text{if } z \in \mathcal{C}_n^+ \cup \mathcal{C}_n^-, \\ \dfrac{nv + \rho_n}{2\rho_n} X_n(z_{r1}) + \dfrac{\rho_n - nv}{2\rho_n} X_n(z_{r2}), & \\ & \text{if } u = u_r, v \in [-n^{-1}\rho_n, n^{-1}\rho_n], \\ \dfrac{nv + \rho_n}{2\rho_n} X_n(z_{l1}) + \dfrac{\rho_n - nv}{2\rho_n} X_n(z_{l2}), & \\ & \text{if } u = u_l > 0, v \in [-n^{-1}\rho_n, n^{-1}\rho_n], \end{cases} \tag{1.9}$$

where $z_{r1} = u_r + in^{-1}\rho_n$, $z_{r2} = u_r - in^{-1}\rho_n$, $z_{l1} = u_l + in^{-1}\rho_n$, $z_{l2} = u_l - in^{-1}\rho_n$ and $\mathcal{C}_n^-$ denotes the symmetric part of $\mathcal{C}_n^+$ about the real axis. A picture of $\mathcal{C}_n^+ \cup \mathcal{C}_n^-$ is the rectangle in Figure 1 with the dash line removed. The advantage of $\hat{X}_n(z)$ over $X_n(z)$ is that the spectral norm of $\mathbf{A}^{-1}(z)$ involved in $\hat{X}_n(z)$ may be well controlled on the contour $\mathcal{C}$. Indeed, loosely speaking, all eigenvalues of $\mathbf{S}$ are located inside the interval (1.5) with a high probability. Therefore, the spectral norm of $\mathbf{A}^{-1}(z)$ corresponding to this case is bounded on $\mathcal{C}$. If some eigenvalues run outside of the interval (1.5) then, at least, we will still have an upper bound $n\rho_n^{-1}$ for the spectral norm of $\mathbf{A}^{-1}(z)$ on $\mathcal{C}$. But, the probability that some eigenvalues run outside of the interval (1.5) is very small, which can offset $n\rho_n^{-1}$ and even more. This is crucial to establish tightness of $\hat{X}_n(z)$ on the contour $\mathcal{C}$. On the other hand, such a truncation does not change the weak limit given in Theorem 2 because the truncation has been made only at the intervals of the length $2\rho_n/n$.

Note that $\hat{X}_n(z)$ may be viewed as a random element in the metric space $C(\mathcal{C}, \mathbb{R}^2)$ of continuous functions from $\mathcal{C}$ to $\mathbb{R}^2$. We are now in a position to state Lemma 1.

LEMMA 1. *Under the assumptions of Theorem 2, we have for $z \in \mathcal{C}$,*

$$(\hat{X}_n(z), Y_n) \xrightarrow{D} (X(z), Y),$$

where $Y \sim N(0, 2c(g'(c))^2)$, *which is independent of* $X(z)$, *a Gaussian stochastic process with mean zero and covariance function* $\operatorname{Cov}(X(z_1), X(z_2))$ *equal to*

$$\frac{2}{cz_1z_2[(1+\underline{m}(z_1))(1+\underline{m}(z_2)) - c\underline{m}(z_1)\underline{m}(z_2)]} - \frac{2m(z_1)m(z_2)}{c}. \tag{1.10}$$

REMARK 4. *Also, note that* $X(z)$ *is exactly the weak limit of the stochastic process* $\sqrt{n}(\mathbf{x}_n^T(\mathbf{S} - z\mathbf{I})^{-1})\mathbf{x}_n - m_n(z))$ *when* $\max_i x_{ni} \to 0$, *whose covariance function is*

$$\operatorname{Cov}(X(z_1), X(z_2)) = \frac{2(z_2\underline{m}(z_2) - z_1\underline{m}(z_1))^2}{c^2z_1z_2(z_1 - z_2)(\underline{m}(z_1) - \underline{m}(z_2))}$$

(*see* [3] *and* [17]).

We conclude this section by presenting the structure of this work. In Section 2, we present a simulation study to identify when the asymptotic normality "kicks in." Then we turn to the proof. To transfer Lemma 1 to Theorem 2 we introduce a new empirical distribution function

$$F_2^{\mathbf{S}}(x) = \sum_{i=1}^{p} t_i^2 I(\lambda_i \le x), \tag{1.11}$$

where $\mathbf{t} = (t_1, \dots, t_n)^T = \mathbf{U}\bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$ and $\mathbf{U}$ is the eigenvector matrix of $\mathcal{S}$. It turns out that $F_2^{\mathbf{S}}(x)$ and the ESD of $\mathbf{S}$ have the same limit, that is, $F_2^{\mathbf{S}}(x) \xrightarrow{\text{i.p.}} F_c(x)$. Thus, by analyticity of $f(x)$, $\bar{\mathbf{s}}^T f(\boldsymbol{\mathcal{S}})\bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|^2$ in Theorem 2 is transferred to the Stieljes transform of $F_2^{\mathbf{S}}(x)$, $\bar{\mathbf{s}}^T(\boldsymbol{\mathcal{S}} - z\mathbf{I})^{-1}\bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|^2$. Moreover, note that

$$\frac{\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}}{1 - \bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}} = \bar{\mathbf{s}}^T(\boldsymbol{\mathcal{S}} - z\mathbf{I})^{-1}\bar{\mathbf{s}}. \tag{1.12}$$

Indeed, this is from the identity (see [20], (2.1))

$$\mathbf{r}^T(\mathbf{B} + a\mathbf{r}\mathbf{r}^T)^{-1} = \frac{r^T\mathbf{B}^{-1}}{1 + a\mathbf{r}^T\mathbf{B}^{-1}\mathbf{r}}, \tag{1.13}$$

where $\mathbf{B}$ and $\mathbf{B} + a\mathbf{r}\mathbf{r}^T$ are both invertible, $\mathbf{r} \in \mathbb{R}^p$ and $a \in \mathbb{R}$. The stochastic process $X_n(z)$ in Lemma 1 is then transferred to the stochastic process $M_n(z)$, where

$$M_n(z) = \sqrt{n}\left(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \frac{c_nm_n(z)}{1 + c_nm_n(z)}\right).$$

The convergence of the stochastic process $M_n(z)$ is given in Sections 3 and 4. The proofs of Theorems 1 and 2, Lemma 1 and Remark 4 are included in Section 5. The last section picks up the truncation of the underlying r.v.'s and some useful lemmas. At this point we would like to point out that both this paper and [5] deal with Stieljes transform of random variables of interest and use martingale method to establish CLT. But the random variable of interest in this paper is a kind of random quadratic forms while [5] is concerned with the trace of random matrices.

Throughout this paper, to save notation, $\mathfrak{M}$ may denote different constants on different occasions.

2. Simulation study. In this section, we provide a simulation study to investigate the performance of normal approximations in Theorem 1. We consider three different populations, the standard normal distribution, the exponential distribution with parameter 1 and the Poisson distribution with parameter 1. From each population we generate 5000 samples of order 100×200, 200×400 and 400×800 matrices, respectively, by routines in R. Each $p \times n$ matrix can be regarded as a collection of n observations of p-dimensional vectors $\mathbf{s}$, so we can calculate T^2 for each matrix. Based on 5000 samples, we have 5000 observed T^2 which give us an estimator of the probability

$$P\left(\frac{\sqrt{n}}{\sqrt{2c_n(1-c_n)^{-3}}}\left(\frac{T^2}{n} - c_n(1-c_n)^{-1}\right) \le x\right)$$

by

$$5000^{-1}\sum I\left(\frac{\sqrt{n}}{\sqrt{2c_n(1-c_n)^{-3}}}\left(\frac{T^2}{n} - c_n(1-c_n)^{-1}\right) \le x\right).$$

In Figures 3–11, there are nine curves. In each figure the horizontal axis means theoretical quantiles of the standard normal distribution and the vertical axis indicates sample quantiles of the normalized Hotelling's T^2 statistics. Every curve represents the quantile-quantile plot for each sampled matrix. From these pictures we see that the quantiles of T^2 get closer to the standard normal one as the sample size and the dimension increase. Actually, when $p = 100$ and $n = 200$, normal distributions already "kick in."

3. Weak convergence of the finite-dimensional distributions. For $z \in \mathcal{C}_n^+$, let $M_n(z) = M_n^{(1)}(z) + M_n^{(2)}(z)$, where

$$M_n^{(1)}(z) = \sqrt{n}(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - E\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}})$$

and

$$M_n^{(2)}(z) = \sqrt{n}\left(E\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \frac{c_n m_n(z)}{1 + c_n m_n(z)}\right).$$

In this section the aim is to prove that for any positive integer r and complex numbers $a_1, \ldots, a_r$,

$$\sum_{i=1}^r a_i M_n^{(1)}(z_i), \qquad \Im z_i \neq 0,$$

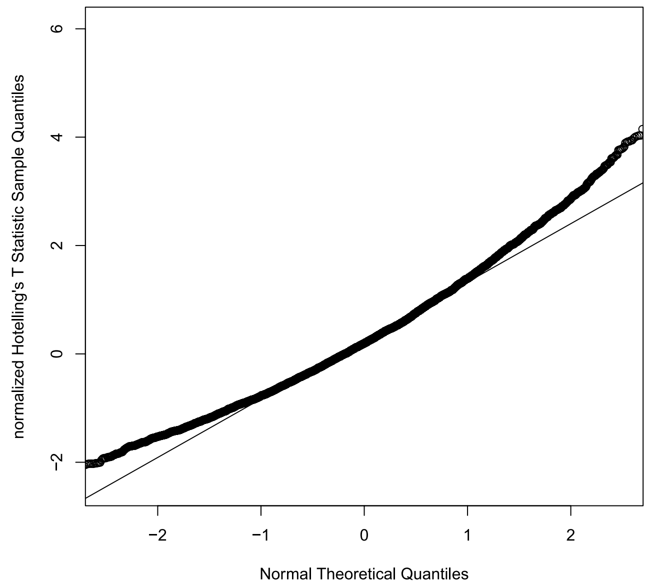

FIG. 3. *Q–Q plot for normal data when* $p = 100$.

converges in distribution to a Gaussian r.v. and to derive the asymptotic covariance function. Before proceeding, r.v.'s need to be truncated. However, we shall post-

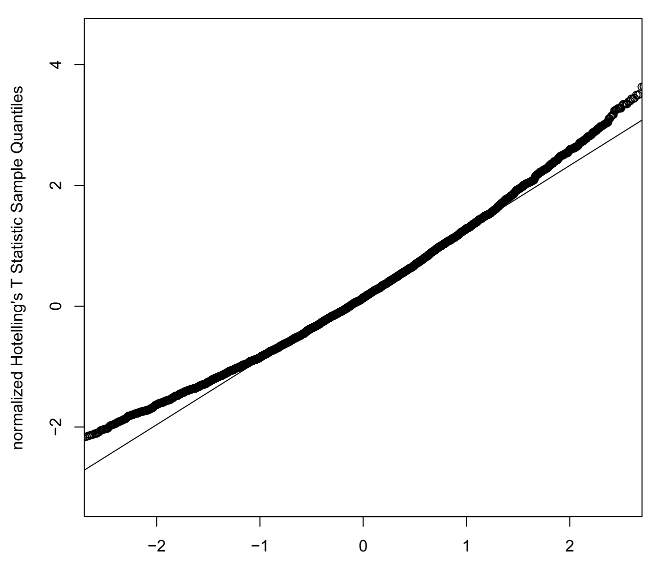

FIG. 4. *Q–Q plot for normal data when* $p = 200$.

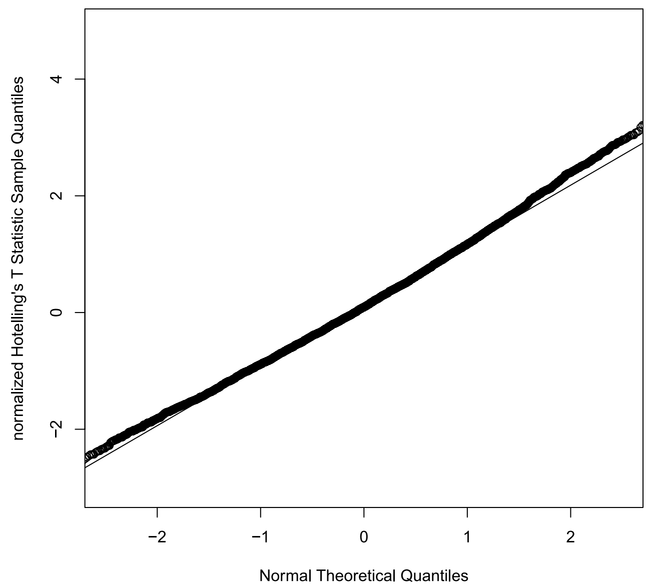

FIG. 5. *Q–Q plot for normal data when* $p = 400$.

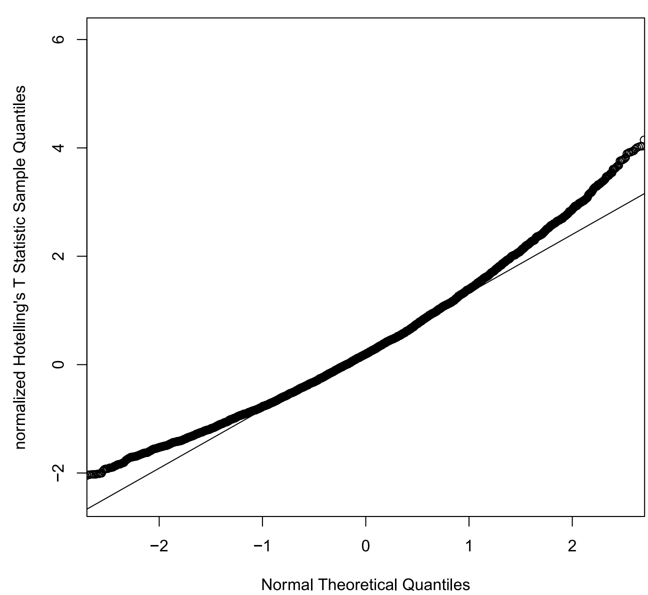

FIG. 6. *Q–Q plot for exponential data when* $p = 100$.

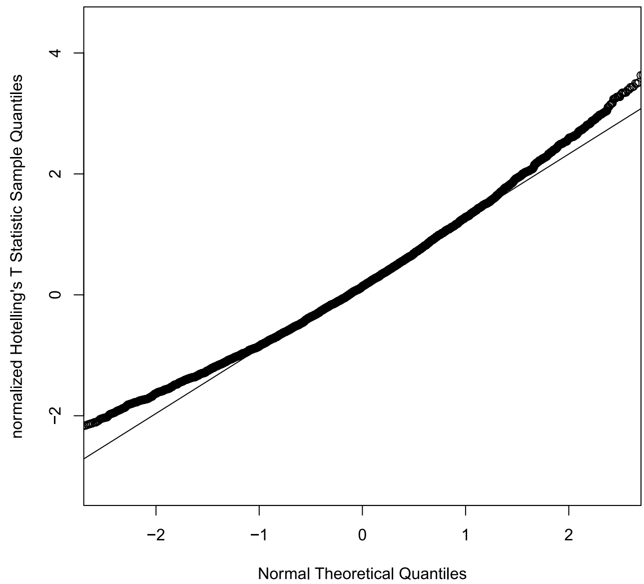

FIG. 7. *Q–Q plot for exponential data when* $p = 200$.

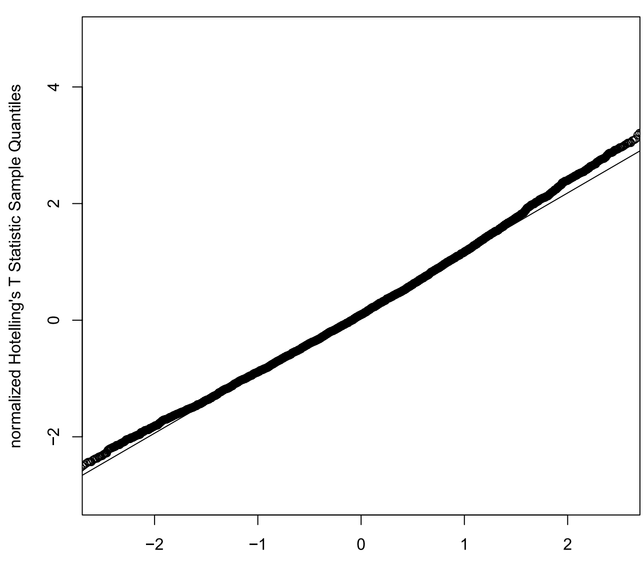

FIG. 8. *Q–Q plot for exponential data when* $p = 400$.

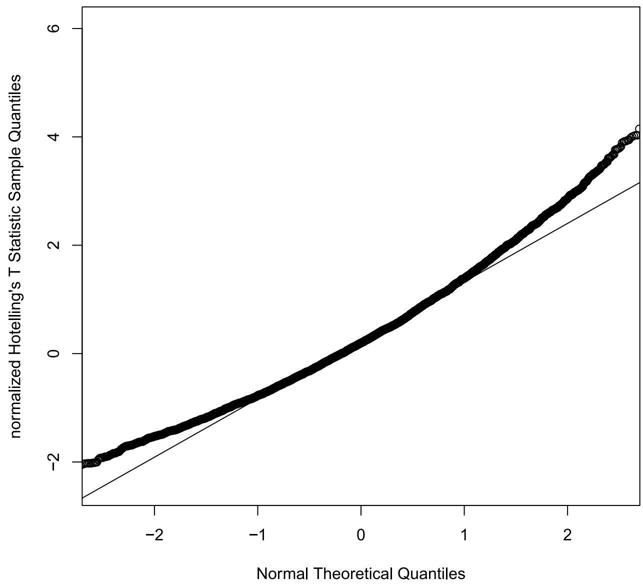

FIG. 9. *Q–Q plot for Poisson data when* $p = 100$.

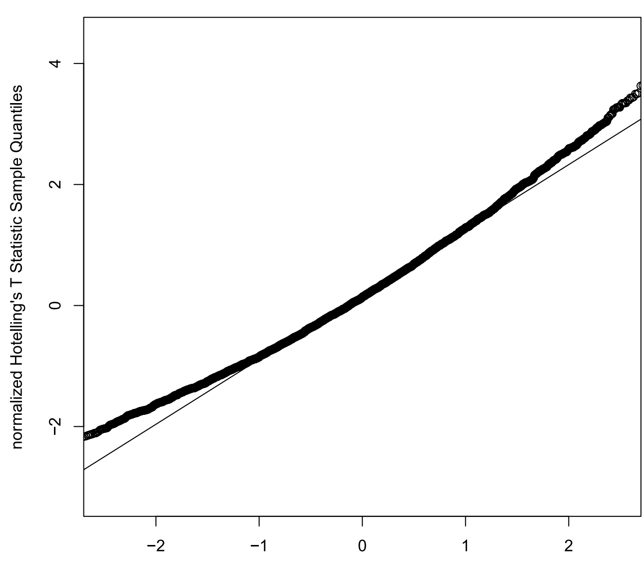

FIG. 10. *Q–Q plot for Poisson data when* $p = 200$.

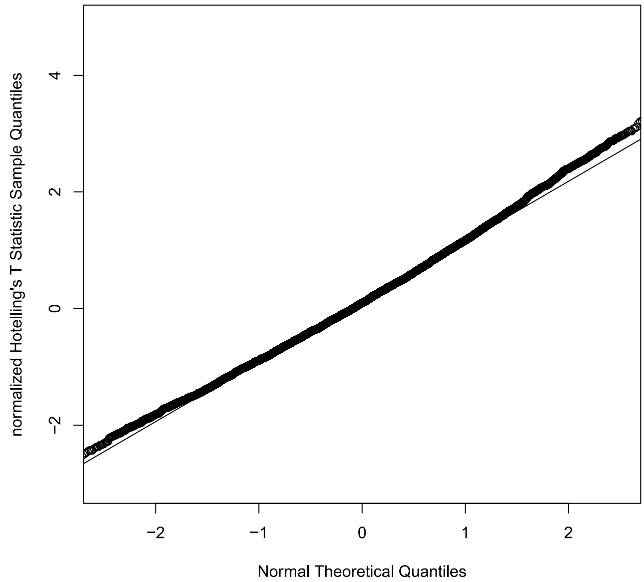

FIG. 11. *Q–Q plot for Poisson data when* $p = 400$.

pone the truncation of r.v.'s until the last section. As a consequence of Lemma 7, we assume that the underlying r.v.'s satisfy

$$|X_{ij}| \le \varepsilon_n \sqrt{n}, \qquad EX_{11} = 0, \qquad E|X_{11}|^2 = 1, \qquad E|X_{11}|^4 < \infty, \tag{3.1}$$

where ε_n is a positive sequence which converges to zero as n goes to infinity.

3.1. *Outline of the argument.* The underlying idea is to write $M_n^{(1)}(z)$ as a sum of martingale difference sequences and to apply Lemma 3, CLT for martingale. Define the σ-field $\mathcal{F}_j = \sigma(\mathbf{s}_1, \dots, \mathbf{s}_j)$ and let $E_j(\cdot) = E(\cdot|\mathcal{F}_j)$ and $E_0(\cdot)$ be the unconditional expectation. We first simplify the martingale representation of $M_n^{(1)}(z)$ as $\sum_{j=1}^n Y_j(z) + o_p(1)$, where $Y_j(z) = -2z\underline{m}(z)E_j(\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j) + z\underline{m}(z)\sqrt{n}E_j(\alpha_j(z))$ and $\alpha_j(z)$ and $\bar{\mathbf{s}}_j$ are defined in the next subsection. Condition (ii) in Lemma 3 is relatively easy to verify. Subsequently, to identify the asymptotic covariance function of $M_n^{(1)}(z)$, the following limits in probability need to be determined:

$$\frac{1}{n}\sum_{j=1}^n E_{j-1}[E_j(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j)], \tag{3.2}$$

$$\sum_{j=1}^n E_{j-1}[E_j(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\alpha_j(z_2))], \tag{3.3}$$

$$n\sum_{j=1}^{n} E_{j-1}[E_j(\alpha_j(z_1))E_j(\alpha_j(z_2))]. \tag{3.4}$$

As for (3.2), note that

$$E_{j-1}[E_j(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j)] = E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)$$

and $\bar{\mathbf{s}}_j$ is an average value of all $\mathbf{s}_1,\dots,\mathbf{s}_n$ without $\mathbf{s}_j$. Intuitively, the product of two conditional expectations in the right-hand side of the above formula should be a multiple of $\frac{1}{n}\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_1))E_j(\mathbf{A}_j^{-1}(z_2))]$. This turns out to be true. For (3.4), a direct calculation indicates that $E_{j-1}[E_j(\alpha_j(z_1))E_j(\alpha_j(z_2))]$ involves $\operatorname{tr}[E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2))]$ [$\mathbf{D}_j(z)$ is defined in the next subsection]. Then our aim is to transfer it to $[E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)]^2$ so that the limit of (3.2) may be used. Essentially, we expect that (3.2) and (3.4) could be reduced to something like

$$\frac{1}{n}\sum_{j=1}^{n} h\left(\frac{j-1}{n}\right)$$

for some function $h(x)$. Finally, since the number of $\mathbf{s}_j$ involved in (3.3) is odd and $\mathbf{s}_j$ is independent of $\bar{\mathbf{s}}_j$ we expect that (3.3) $\xrightarrow{\text{i.p.}} 0$.

3.2. *Notation and estimates*. We first introduce some notation. Let

$$\begin{aligned}
\mathbf{A}_j^{-1}(z) &= (\mathbf{S} - n^{-1}\mathbf{s}_j\mathbf{s}_j^T - z\mathbf{I})^{-1},\\
\mathbf{A}_{ij}^{-1}(z) &= (\mathbf{S} - n^{-1}\mathbf{s}_i\mathbf{s}_i^T - n^{-1}\mathbf{s}_j\mathbf{s}_j^T - z\mathbf{I})^{-1},\\
\bar{\mathbf{s}}_j &= \bar{\mathbf{s}} - n^{-1}\mathbf{s}_j,\\
\mathbf{D}_j(z) &= \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z),\\
\beta_j(z) &= \frac{1}{1+(1/n)\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j},\\
\beta_j^{\mathrm{tr}}(z) &= \frac{1}{1+(1/n)\operatorname{tr}\mathbf{A}_j^{-1}(z)},\\
b_1(z) &= \frac{1}{1+(1/n)E\operatorname{tr}\mathbf{A}_1^{-1}(z)},\\
\gamma_j(z) &= \frac{1}{n}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j - \frac{1}{n}\operatorname{tr}\mathbf{A}_j^{-1}(z),\\
\xi_j(z) &= \frac{1}{n}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j - \frac{1}{n}E\operatorname{tr}\mathbf{A}_j^{-1}(z),
\end{aligned}$$

$$\alpha_j(z) = \frac{1}{n}\mathbf{s}_j^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j \bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z)\mathbf{s}_j - \frac{1}{n}\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-2}(z)\bar{\mathbf{s}}_j,$$

$$\beta_{ij}(z) = \frac{1}{1+(1/n)\mathbf{s}_i^T A_{ij}^{-1}(z)\mathbf{s}_i},$$

$$\beta_{ij}^{\mathrm{tr}}(z) = \frac{1}{1+(1/n)\operatorname{tr} A_{ij}^{-1}(z)},$$

$$b_{12}(z) = \frac{1}{1+(1/n)E\operatorname{tr} A_{12}^{-1}(z)}$$

and

$$\xi_{ij}(z) = \frac{1}{n}\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z)\mathbf{s}_i - \frac{1}{n}E\operatorname{tr}\mathbf{A}_{12}^{-1}(z),$$

$$\gamma_{ij}(z) = \frac{1}{n}\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z)\mathbf{s}_i - (1/n)\operatorname{tr}\mathbf{A}_{ij}^{-1}(z).$$

We next list some results to be used later. A direct calculation indicates that the following equalities are true:

$$E(\mathbf{s}_1^T\mathbf{A}\mathbf{s}_1 - \operatorname{tr}\mathbf{A})(\mathbf{s}_1^T\mathbf{B}\mathbf{s}_1 - \operatorname{tr}\mathbf{B}) = (EX_{11}^4 - |EX_{11}^2|^2 - 2)\sum_{i=1}^p a_{ii}b_{ii} + |EX_{11}^2|^2\operatorname{tr}\mathbf{A}\mathbf{B}^T + \operatorname{tr}\mathbf{A}\mathbf{B}; \tag{3.5}$$

$$E[(\mathbf{s}_1^T\mathbf{A}\mathbf{s}_1 - \operatorname{tr}\mathbf{A})\mathbf{s}_1^T\mathbf{B}\mathbf{r}] = EX_{11}^3\sum_{i=1}^p a_{ii}\mathbf{e}_i^T\mathbf{B}\mathbf{r}, \tag{3.6}$$

where $\mathbf{B} = (b_{ij})_{p\times p}$ and $\mathbf{A} = (a_{ij})_{p\times p}$ are deterministic complex matrices and $\mathbf{r}$ is a deterministic vector. Here $\mathbf{e}_i$ is the vector with the ith element being 1 and zero otherwise. In what follows, to facilitate the analysis in the subsequent subsections, we shall assume $v = \Im z > 0$. Note that $\beta_j(z), \beta_j^{\mathrm{tr}}(z), \beta_{ij}(z), \beta_{ij}^{\mathrm{tr}}(z), b_1(z), b_{12}(z)$ are bounded in absolute value by $|z|/v$ [see [4], (3.4)]. From (1.13) we have

$$\mathbf{A}^{-1}(z) - \mathbf{A}_j^{-1}(z) = \mathbf{A}^{-1}(z)\big(\mathbf{A}_j(z) - \mathbf{A}(z)\big)\mathbf{A}_j^{-1}(z) = -\frac{1}{n}\tilde{\mathbf{A}}_j(z)\beta_j(z), \tag{3.7}$$

where $\tilde{\mathbf{A}}_j(z) = \mathbf{A}_j^{-1}(z)\mathbf{s}_j\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)$. From Lemma 2.10 of [4], for any matrix $\mathbf{B}$,

$$\big|\operatorname{tr}\big[\big(\mathbf{A}^{-1}(z) - \mathbf{A}_j^{-1}(z)\big)\mathbf{B}\big]\big| \le \frac{\|\mathbf{B}\|}{v}, \tag{3.8}$$

where $\|\cdot\|$ denotes the spectral norm of a matrix. Moreover, Section 4 in [4] shows that

$$n^{-k}E|\operatorname{tr}\mathbf{A}_1^{-1}(z) - E\operatorname{tr}\mathbf{A}_1^{-1}(z)|^k = O(n^{-k/2}), \qquad k \ge 2. \tag{3.9}$$

One should also note that (3.9) is still true when $\mathbf{A}_1^{-1}(z)$ is replaced by $\mathbf{A}_{12}^{-1}(z)$.

From now on, we calculate estimates. To simplify the statements, assume that the spectral norms of nonrandom $\mathbf{B}, \mathbf{B}_i, \mathbf{A}_i, \mathbf{C}$ involved in the equalities (3.10)–(3.18) below are all bounded above by a constant. For $k \ge 2$, it follows from Lemma 4, (3.1) and (3.9) that

$$\begin{aligned} n^{-k}E|\mathbf{s}_1^T\mathbf{B}\mathbf{s}_1 - \operatorname{tr}\mathbf{B}|^k &= O(\varepsilon_n^{2k-4}n^{-1}), \\ E|\xi_1(z)|^k &= O(\varepsilon_n^{2k-4}n^{-1}) \end{aligned} \tag{3.10}$$

and that

$$\begin{aligned} &n^{-k}E|\mathbf{s}_1^T\mathbf{B}\mathbf{e}_i\mathbf{e}_j^T\mathbf{C}\mathbf{s}_1|^k \\ &\quad \le \mathfrak{M}n^{-k}[E|\mathbf{s}_1^T\mathbf{B}\mathbf{e}_i\mathbf{e}_j^T\mathbf{C}\mathbf{s}_1 - \operatorname{tr}(\mathbf{B}\mathbf{e}_i\mathbf{e}_j^T\mathbf{C})|^k + E|\mathbf{e}_j^T\mathbf{C}\mathbf{B}\mathbf{e}_i|^k] \\ &\quad = O(\varepsilon_n^{2k-4}n^{-2}). \end{aligned} \tag{3.11}$$

We shall establish the estimates (3.12)–(3.14) below:

$$\begin{aligned} E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1|^k &= O\big(n^{(k/2-2)}\varepsilon_n^{k-4}\big), \qquad k \ge 4, \\ E|\alpha_1(z)|^k &= O(n^{-2}\varepsilon_n^{2k-4}), \qquad k \ge 2, \end{aligned} \tag{3.12}$$

$$E|\mathbf{s}_1^T\mathbf{B}\mathbf{s}_2|^k = O(n^{k-2}\varepsilon_n^{k-4}), \qquad k \ge 4, \tag{3.13}$$

and for $m \ge 0, q \ge 1, 0 \le r \le 2$,

$$E\left|\prod_{i=1}^m \frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1 \prod_{j=1}^q \frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)(\mathbf{s}_1^T\mathbf{C}_l\bar{\mathbf{s}}_1)^r\right| = O\big(n^{-1/2}\varepsilon_n^{(q-2)\vee 0}\big). \tag{3.14}$$

One should note that (3.12) and (3.13) also give the estimates for $k = 2$. For example,

$$E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1|^2 \le (E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1|^4)^{1/2} = O(1). \tag{3.15}$$

In addition, from (3.10) and (3.13) we also conclude that

$$\begin{aligned} E|n^{-1}\mathbf{s}_1^T\mathbf{B}\mathbf{s}_1\mathbf{s}_1^T\mathbf{C}\mathbf{s}_2|^4 &\le \mathfrak{M}E|n^{-1}(\mathbf{s}_1^T\mathbf{B}\mathbf{s}_1 - \operatorname{tr}\mathbf{B})\mathbf{s}_1^T\mathbf{C}\mathbf{s}_2|^4 \\ &\quad + \mathfrak{M}E|\mathbf{s}_1^T\mathbf{C}\mathbf{s}_2|^4 \\ &= O(n^{5/2}). \end{aligned} \tag{3.16}$$

Consider (3.12) first. Note that for $k \ge 4$

$$\begin{aligned} E|\bar{\mathbf{s}}_1^T\bar{\mathbf{s}}_1|^k &\le \frac{\mathfrak{M}}{n^{2k}}\left[E\left|\sum_{i=2}^n \mathbf{s}_i^T\mathbf{s}_i\right|^k + E\left|\sum_{i_1\ne i_2, i_1>1, i_2>1} \mathbf{s}_{i_1}^T\mathbf{s}_{i_2}\right|^k\right] \\ &= O(1). \end{aligned} \tag{3.17}$$

Indeed, applying Lemma 2 twice gives

$$\begin{aligned}
E\left|\sum_{i=2}^{n}\mathbf{s}_i^T\mathbf{s}_i\right|^k &\le \mathfrak{M}E\left|\sum_{i=2}^{n}\big(\mathbf{s}_i^T\mathbf{s}_i - E(\mathbf{s}_i^T\mathbf{s}_i)\big)\right|^k + \mathfrak{M}\left|\sum_{i=2}^{n}E(\mathbf{s}_i^T\mathbf{s}_i)\right|^k \\
&\le \mathfrak{M}\left(\sum_{i=2}^{n}E\big(\mathbf{s}_i^T\mathbf{s}_i - E(\mathbf{s}_i^T\mathbf{s}_i)\big)^2\right)^{k/2} \\
&\quad + \mathfrak{M}\sum_{i=2}^{n}E|\mathbf{s}_i^T\mathbf{s}_i - E(\mathbf{s}_i^T\mathbf{s}_i)|^k + \mathfrak{M}n^{2k} \\
&\le \mathfrak{M}n^k + \mathfrak{M}n\left[\left(\sum_{m=1}^{p}E(X_{m2}^2-1)^2\right)^{k/2} + \sum_{m=1}^{p}E|X_{m2}^2-1|^k\right] \\
&\quad + \mathfrak{M}n^{2k} \\
&\le \mathfrak{M}n^{2k},
\end{aligned}$$

while using Lemma 2 three times we obtain

$$\begin{aligned}
E\left|\sum_{i_1\ne i_2, i_1>1, i_2>1}\mathbf{s}_{i_1}^T\mathbf{s}_{i_2}\right|^k &\le n^{k-1}\sum_{i_1>1}E\left|\sum_{i_2>1, i_1\ne i_2}\mathbf{s}_{i_1}^T\mathbf{s}_{i_2}\right|^k \le n^k E\left|\sum_{i=3}^{n}\mathbf{s}_2^T\mathbf{s}_i\right|^k \\
&\le \mathfrak{M}n^k E\left|\sum_{i=3}^{n}E[(\mathbf{s}_2^T\mathbf{s}_i)^2|\mathcal{G}_{i-1}]\right|^{k/2} + \mathfrak{M}n^k\sum_{i=3}^{n}E|\mathbf{s}_2^T\mathbf{s}_i|^k \\
&\le \mathfrak{M}[n^{(3/2)k}E|\mathbf{s}_2^T\mathbf{s}_2 - E\mathbf{s}_2^T\mathbf{s}_2|^{k/2} + n^{2k} + n^{k+1}E|\mathbf{s}_2^T\mathbf{s}_3|^k] \\
&= O(n^{2k}),
\end{aligned}$$

where $\mathcal{G}_i = \sigma(\mathbf{s}_2, \dots, \mathbf{s}_i)$. It follows from (3.17) that for $k \ge 4$

$$E|\bar{\mathbf{s}}_1^T\mathbf{B}\bar{\mathbf{s}}_1|^k = E\|\bar{\mathbf{s}}_1^T\mathbf{B}\bar{\mathbf{s}}_1\|^k \le E(\|\bar{\mathbf{s}}_1^T\|\|\mathbf{B}\|\|\bar{\mathbf{s}}_1\|)^k \le \mathfrak{M}E|\bar{\mathbf{s}}_1^T\bar{\mathbf{s}}_1|^k \le \mathfrak{M}, \tag{3.18}$$

where $\|\cdot\|$ denotes the spectral norm of a matrix. This, together with Lemma 4, ensures that for $k \ge 4$

$$\begin{aligned}
E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1|^k &= E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1\bar{\mathbf{s}}_1^T\mathbf{B}^*\mathbf{s}_1|^{k/2} \\
&\le \mathfrak{M}E|\mathbf{s}_1^T\mathbf{B}\bar{\mathbf{s}}_1\bar{\mathbf{s}}_1^T\mathbf{B}^*\mathbf{s}_1 - \bar{\mathbf{s}}_1^T\mathbf{B}^*\mathbf{B}\bar{\mathbf{s}}_1|^{k/2} + \mathfrak{M}E|\bar{\mathbf{s}}_1^T\mathbf{B}^*\mathbf{B}\bar{\mathbf{s}}_1|^{k/2} \\
&\le [\mathfrak{M}n^{k/2-2}\varepsilon_n^{k-4} + \mathfrak{M}]E|\bar{\mathbf{s}}_1^T\mathbf{B}^*\mathbf{B}\bar{\mathbf{s}}_1|^{k/2} + \mathfrak{M} \\
&\le \mathfrak{M}n^{k/2-2}\varepsilon_n^{k-4},
\end{aligned}$$

which gives the first estimate in (3.12) as well as the order of $E|\alpha_1(z)|^k$.

Second, consider (3.13). Let $\mathbf{y} = (y_1, \dots, y_p)^T = \mathbf{B}\mathbf{s}_2$ and then, by Lemma 2 and (3.10), for $k \ge 4$,

$$
\begin{aligned}
E|\mathbf{s}_1^T\mathbf{y}|^k &\le \mathfrak{M}E\left(\sum_{m=1}^{p}|y_m|^2\right)^{k/2} + M\sum_{m=1}^{p}E|X_{m1}|^kE|y_m|^k \\
&\le \mathfrak{M}E|\mathbf{y}^*\mathbf{y}|^{k/2} + \mathfrak{M}n^{k/2-2}\varepsilon_n^{k-4}E|\mathbf{y}^*\mathbf{y}|^{k/2} \\
&\le \mathfrak{M}(1+n^{k/2-2}\varepsilon_n^{k-4})E|\mathbf{s}_2^T\mathbf{B}^*\mathbf{B}\mathbf{s}_2 - \operatorname{tr}(\mathbf{B}^*\mathbf{B})|^{k/2} \\
&\quad + \mathfrak{M}n^{k/2} + \mathfrak{M}n^{k-2}\varepsilon_n^{k-4} \\
&= O(n^{k-2}\varepsilon_n^{k-4}),
\end{aligned} \tag{3.19}
$$

where we also use the fact that for $k \ge 4$

$$
\sum_m |y_m|^k \le \left(\sum_m |y_m|^2\right)^{k/2}.
$$

As for (3.14), if $m = 0$ and $r = 0$, then (3.14) directly follows from (3.10) and the Hölder inequality. If $m \ge 1$ and $r = 0$, then by induction on m we have

$$
\begin{aligned}
&E\left|\prod_{i=1}^{m}\frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)\right| \\
&\quad \le E\left|\prod_{i=1}^{m-1}\frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1\frac{1}{n}(\mathbf{s}_1^T\mathbf{A}_m\mathbf{s}_1 - \operatorname{tr}\mathbf{A}_m)\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)\right| \\
&\qquad + \mathfrak{M}E\left|\prod_{i=1}^{m-1}\frac{1}{n}\operatorname{tr}\mathbf{A}_i\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)\right| \\
&\quad = O\big(n^{-1/2}\varepsilon_n^{(q-2)\vee 0}\big).
\end{aligned}
$$

Repeating the argument above gives

$$
E\left|\prod_{i=1}^{m}\frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)\right|^2 = O\big(n^{-1}\varepsilon_n^{(2q-4)\vee 0}\big)
$$

[$m = 0$ by (3.10) and $m \ge 1$ by induction]. Thus, for the case $m \ge 1$ and $2 \ge r \ge 1$, by (3.12) we obtain

$$
\begin{aligned}
&E\left|\prod_{i=1}^{m}\frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)(\mathbf{s}_1^T\mathbf{C}_1\bar{\mathbf{s}}_1)^r\right| \\
&\quad \le \left(E\left|\prod_{i=1}^{m}\frac{1}{n}\mathbf{s}_1^T\mathbf{A}_i\mathbf{s}_1\prod_{j=1}^{q}\frac{1}{n}(\mathbf{s}_1^T\mathbf{B}_j\mathbf{s}_1 - \operatorname{tr}\mathbf{B}_j)\right|^2E|\mathbf{s}_1^T\mathbf{C}_1\bar{\mathbf{s}}_1|^{2r}\right)^{1/2} \\
&\quad = O\big(n^{-1/2}\varepsilon_n^{(q-2)\vee 0}\big).
\end{aligned}
$$

When $m = 0$ and $2 \geq r \geq 1$, (3.14) can be obtained similarly. Thus, we have proved (3.14).

3.3. *The simplification of* $M_n^{(1)}(z)$. To develop CLT for $M_n^{(1)}(z)$, we write it as a sum of martingale difference sequences. When simplifying such a martingale representation, a well-known trick is to use the fact that

$$E_j[h(\operatorname{tr}\mathbf{A}_j^{-1}(z))] = E_{j-1}[h(\operatorname{tr}\mathbf{A}_j^{-1}(z))], \tag{3.20}$$

where $h(x)$ is some function. For example, when $h(x) = 1/(1 + n^{-1}x)$, (3.20) becomes $E_j(\beta_j^{\mathrm{tr}}) = E_{j-1}(\beta_j^{\mathrm{tr}})$.

Notice that $E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j) = E_{j-1}(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)$. We then write

$$\begin{aligned}
M_n^{(1)}(z) &= \sqrt{n}\sum_{j=1}^{n}[E_j(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}) - E_{j-1}(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}})] \\
&= \sqrt{n}\sum_{j=1}^{n}\big[E_j\big(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\big) \\
&\qquad\qquad - E_{j-1}\big(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\big)\big] \\
&= \sqrt{n}\sum_{j=1}^{n}[(E_j - E_{j-1})(a_{n1} + a_{n2} + a_{n3})],
\end{aligned} \tag{3.21}$$

where

$$a_{n1} = (\bar{\mathbf{s}} - \bar{\mathbf{s}}_j)^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}, \qquad a_{n2} = \bar{\mathbf{s}}_j^T\big(\mathbf{A}^{-1}(z) - \mathbf{A}_j^{-1}(z)\big)\bar{\mathbf{s}},$$
$$a_{n3} = \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)(\bar{\mathbf{s}} - \bar{\mathbf{s}}_j).$$

The above sum involving a_{n1} and a_{n2} will be further simplified below.

First, splitting $\mathbf{A}^{-1}(z)$ into the sum of $\mathbf{A}^{-1}(z) - \mathbf{A}_j^{-1}(z)$ and $\mathbf{A}_j^{-1}(z)$ and splitting $\bar{\mathbf{s}}$ into the sum of $\bar{\mathbf{s}}_j$ and $\mathbf{s}_j/n$, by (3.7) we then have

$$a_{n1} = a_{n1}^{(1)} + a_{n1}^{(2)} + a_{n1}^{(3)} + a_{n1}^{(4)}, \tag{3.22}$$

where

$$a_{n1}^{(1)} = -\frac{1}{n^3}(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j)^2\beta_j(z), \qquad a_{n1}^{(2)} = -\frac{1}{n^2}\mathbf{s}_j^T\tilde{\mathbf{A}}_j(z)\bar{\mathbf{s}}_j\beta_j(z)$$

and

$$a_{n1}^{(3)} = \frac{1}{n^2}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j, \qquad a_{n1}^{(4)} = \frac{1}{n}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j.$$

Using (3.20) and

$$\beta_j(z) = \beta_j^{\mathrm{tr}}(z) - \beta_j(z)\beta_j^{\mathrm{tr}}(z)\gamma_j(z), \tag{3.23}$$

we have

$$
\begin{aligned}
&(E_j - E_{j-1})\big(a_{n1}^{(1)}\big) \\
&\quad = (E_j - E_{j-1})\left[\frac{1}{n^3}(\mathbf{s}_j^T \mathbf{A}_j^{-1}(z)\mathbf{s}_j)^2 \beta_j^{\mathrm{tr}}(z)\right] - \varsigma_n \\
&\quad = (E_j - E_{j-1})\left[\frac{1}{n}\gamma_j^2(z)\beta_j^{\mathrm{tr}}(z)\right] \\
&\qquad + (E_j - E_{j-1})\left[\frac{2}{n^2}\gamma_j(z)\beta_j^{\mathrm{tr}}(z)\operatorname{tr}\mathbf{A}_j^{-1}(z)\right] - \varsigma_n,
\end{aligned}
$$

where $\varsigma_n = (E_j - E_{j-1})\frac{1}{n^3}(\mathbf{s}_j^T \mathbf{A}_j^{-1}(z)\mathbf{s}_j)^2 \beta_j(z)\beta_j^{\mathrm{tr}}(z)\gamma_j(z)$. This, together with (3.14), shows that

$$
\begin{aligned}
&E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})\big(a_{n1}^{(1)}\big)\right|^2 \\
&\quad = n\sum_{j=1}^{n} E\big|(E_j - E_{j-1})\big(a_{n1}^{(1)}\big)\big|^2 \\
&\quad \le \mathfrak{M}E|\gamma_1(z)|^4 + E|\gamma_1(z)|^2 + \mathfrak{M}E\left|\gamma_1(z)\frac{1}{n^2}(\mathbf{s}_1^T \mathbf{A}_1^{-1}(z)\mathbf{s}_1)^2\right|^2 \\
&\quad = O(n^{-1/2}),
\end{aligned}
$$

which gives

$$
\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})\big(a_{n1}^{(1)}\big) \xrightarrow{\text{i.p.}} 0.
$$

By (3.10) it is a simple matter to verify that

$$
\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})\big(a_{n1}^{(3)}\big) \xrightarrow{\text{i.p.}} 0.
$$

Appealing to (3.14) we have

$$
E\left|\sum_{j=1}^{n}(E_j - E_{j-1})\gamma_j(z)\frac{1}{\sqrt{n}}\mathbf{s}_j^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j \beta_j^{\mathrm{tr}}(z)\right|^2 = O(n^{-1/2})
$$

and

$$
E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})\frac{1}{n^2}\mathbf{s}_j^T \tilde{\mathbf{A}}_j(z)\bar{\mathbf{s}}_j \beta_j(z)\gamma_j(z)\beta_j^{\mathrm{tr}}(z)\right|^2 = O(n^{-1/2}),
$$

which, together with (3.23), leads to

$$\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})\big(a_{n1}^{(2)}\big)$$
$$= -\sum_{j=1}^{n} E_j\bigg[\big(1-\beta_j^{\mathrm{tr}}(z)\big)\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\bigg] + o_p(1).$$

This ensures that

$$\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})(a_{n1})$$
$$= \sum_{j=1}^{n} E_j\bigg(\beta_j^{\mathrm{tr}}(z)\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\bigg) + o_p(1) \tag{3.24}$$
$$= -z\underline{m}(z)\sum_{j=1}^{n} E_j\bigg(\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\bigg) + o_p(1),$$

because, by (2.17) in [5], (3.9) and (3.12),

$$E\big|\big(\beta_j^{\mathrm{tr}}(z) + z\underline{m}(z)\big)\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\big|^2 = o(1). \tag{3.25}$$

Second, splitting $\bar{\mathbf{s}}$ into the sum of $\bar{\mathbf{s}}_j$ and $\mathbf{s}_j/n$ further gives

$$a_{n2} = -\frac{1}{n^2}\bar{\mathbf{s}}_j^T\tilde{\mathbf{A}}_j(z)\mathbf{s}_j\beta_j(z) - \frac{1}{n}\bar{\mathbf{s}}_j^T\tilde{\mathbf{A}}_j(z)\bar{\mathbf{s}}_j\beta_j(z)$$

and thus, as in treating $a_{n1}^{(2)}$, we have

$$\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})(a_{n2})$$
$$= -\sum_{j=1}^{n}(E_j - E_{j-1})\bigg[\big(1-\beta_j^{\mathrm{tr}}(z)\big)\frac{1}{\sqrt{n}}\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j\bigg]$$
$$- \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(E_j - E_{j-1})[\bar{\mathbf{s}}_j^T\tilde{\mathbf{A}}_j(z)\bar{\mathbf{s}}_j\beta_j^{\mathrm{tr}}(z)] + o_p(1)$$
$$= -\big(1+z\underline{m}(z)\big)\sum_{j=1}^{n} E_j\bigg(\frac{1}{\sqrt{n}}\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)\mathbf{s}_j\bigg)$$
$$+ z\underline{m}(z)\sum_{j=1}^{n}\sqrt{n}E_j(\alpha_j(z)) + o_p(1),$$

where in the last step we also use the estimate

$$\begin{aligned}E\big|\big(\beta_j^{\mathrm{tr}}(z)+z\underline{m}(z)\big)\alpha_j(z)\big|^2&=E\big[E\big(\big|\big(\beta_j^{\mathrm{tr}}(z)+z\underline{m}(z)\big)\alpha_j(z)\big|^2|\sigma(\mathbf{s}_i,i\neq j)\big)\big]\\&=E\big[|\beta_j^{\mathrm{tr}}(z)+z\underline{m}(z)|^2E\big(|\alpha_j(z)|^2|\sigma(\mathbf{s}_i,i\neq j)\big)\big]\\&=o(n^{-2}),\end{aligned}$$

which is from (2.17) in [5], (3.9) and (3.12).

Recalling $Y_j(z)=-2z\underline{m}(z)E_j(\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)+z\underline{m}(z)\sqrt{n}E_j(\alpha_j(z))$, so far we have proved

$$M_n^{(1)}(z)=\sum_{j=1}^n Y_j(z)+o_p(1).$$

Consequently, for finite dimension convergence of $M_n^{(1)}(z)$, we need consider only the sum

$$\sum_{i=1}^r a_i\sum_{j=1}^n Y_j(z_i)=\sum_{j=1}^n\sum_{i=1}^r a_iY_j(z_i). \tag{3.26}$$

Next we verify condition (ii) of Lemma 3. Recalling $\mathbf{D}_j(z)=\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\bar{\mathbf{s}}_j^T\times\mathbf{A}_j^{-1}(z)$, write

$$\alpha_j(z)=\alpha_j^{(1)}(z)+\alpha_j^{(2)}(z)+\alpha_j^{(3)}(z),$$

where

$$\alpha_j^{(3)}(z)=\frac{1}{n}\sum_{h\neq l}\mathbf{e}_h^T\mathbf{D}_j(z)\mathbf{e}_lX_{hj}X_{lj},$$

$$\alpha_j^{(2)}(z)=\frac{1}{n}\sum_{h=1}^p\mathbf{e}_h^T\mathbf{D}_j(z)\mathbf{e}_h[X_{hj}^2I(|X_{hj}|\le\log n)-EX_{hj}^2I(|X_{hj}|\le\log n)]$$

and

$$\alpha_j^{(1)}(z)=\frac{1}{n}\sum_{h=1}^p\mathbf{e}_h^T\mathbf{D}_j(z)\mathbf{e}_h[X_{hj}^2I(|X_{hj}|>\log n)-EX_{hj}^2I(|X_{hj}|>\log n)].$$

Lemma 5 and (3.18) show that $E|\alpha_j^{(3)}(z)|^4=O(n^{-4})$. Lemma 2 and (3.18) give $E|\alpha_j^{(2)}(z)|^4=O(n^{-4}(\log n)^4)$ because

$$\begin{aligned}\sum_{h=1}^n|\mathbf{e}_h^T\mathbf{D}_j(z)\mathbf{e}_h|^k&\le\left|\sum_{h=1}^n\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(\bar{z})\mathbf{e}_h\mathbf{e}_h^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j\right|^k\\&=(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(\bar{z})\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)^k,\end{aligned}\tag{3.27}$$

where $k = 2$ or 4 and $\mathbf{A}_j^{-1}(\bar{z})$ denotes the complex conjugate of $\mathbf{A}_j^{-1}(z)$. We conclude from (3.27) and $EX_{11}^4 I(|X_{11}| > \log n) \to 0$ that $E|\alpha_j^{(1)}(z)|^2 = o(n^{-2})$. Therefore, we obtain

$$\begin{aligned}\sum_{j=1}^n E\left|\sum_{i=1}^r a_i Y_j(z_i)\right|^2 I\left(\left|\sum_{i=1}^r a_i Y_j(z_i)\right| \ge \varepsilon\right)\\ \le 4\sum_{j=1}^n \sum_{h=1}^4 E\left|\sum_{i=1}^r a_i Y_j^{(h)}(z_i)\right|^2 I\left(\left|\sum_{i=1}^r a_i Y_j^{(h)}(z_i)\right| \ge \varepsilon/4\right)\\ \le \frac{\mathfrak{M}}{\varepsilon^2}\sum_{j=1}^n \sum_{h=2}^4 E\left|\sum_{i=1}^r a_i Y_j^{(h)}(z_i)\right|^4 + 4\sum_{j=1}^n E\left|\sum_{i=1}^r a_i Y_j^{(1)}(z_i)\right|^2 \to 0,\end{aligned}$$

where $Y_j^{(h)}(z) = z\underline{m}(z)\sqrt{n}E_j(\alpha_j^{(h)}(z)), h = 1, 2, 3$ and $Y_j^{(4)}(z) = -2z\underline{m}(z) \times E_j(\frac{1}{\sqrt{n}}\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)$. Here we also use $E|Y_j^{(4)}(z)|^4 = O(n^{-2})$ by (3.12). Thus, the condition (ii) of Lemma 3 is satisfied. Hence, the next task is to find, for $z_1, z_2 \in \mathbb{C} \setminus \mathbb{R}$, the limit in probability of

$$\sum_{j=1}^n E_{j-1}(Y_j(z_1)Y_j(z_2)). \tag{3.28}$$

To this end, it is enough to find the limits in probability for (3.2), (3.3) and (3.4).

The limits of (3.2)–(3.4) and finally (3.28) will be determined in the subsequent subsections.

3.4. *The limit of* (3.2). Our aim is to prove that

$$\begin{aligned}(3.2) = \frac{z_1 z_2\underline{m}(z_1)\underline{m}(z_2)}{n}\sum_{j=1}^n \frac{j-1}{n^2}\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)))\\ + o_p(1).\end{aligned} \tag{3.29}$$

The strategy is to first replace $\bar{\mathbf{s}}_j$ by $\frac{1}{n}\sum_{i\neq j}^n \mathbf{s}_i$, then replace the resulting quadratic forms in terms of $\mathbf{s}_i$ by its corresponding trace and $\beta_{ij}(z_2)$ by its corresponding limit.

To this end, introduce $\underline{\mathbf{A}}_j^{-1}(z)$ and $\underline{\bar{\mathbf{s}}}_j$ like $\mathbf{A}_j^{-1}(z)$ and $\bar{\mathbf{s}}_j$, respectively, but $\underline{\mathbf{A}}_j^{-1}(z)$ and $\underline{\bar{\mathbf{s}}}_j$ are now defined by $\mathbf{s}_1, \dots, \mathbf{s}_{j-1}, \underline{\mathbf{s}}_{j+1}, \dots, \underline{\mathbf{s}}_n$ instead of $\mathbf{s}_1, \dots, \mathbf{s}_{j-1}, \mathbf{s}_{j+1}, \dots, \mathbf{s}_n$. Here $\{\underline{\mathbf{s}}_{j+1}, \dots, \underline{\mathbf{s}}_n\}$ are i.i.d. copies of $\mathbf{s}_1$ and independent of $\{\mathbf{s}_j, j = 1, \dots, n\}$. Therefore, (3.2) is equal to

$$\frac{1}{n}\sum_{j=1}^n \operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))] = \frac{1}{n}\sum_{j=1}^n E_j[\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j].$$

Applying $\bar{\mathbf{s}}_j = \frac{1}{n}\sum_{i\neq j}^n \mathbf{s}_i$ and (1.13) further gives

$$E_j[\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j] = \frac{1}{n}\sum_{i\neq j}^n E_j[\beta_{ij}(z_2)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j]. \tag{3.30}$$

The next aim is to replace $\beta_{ij}(z_2)$ in the equality above by $\beta_{ij}^{\mathrm{tr}}(z_2)$. To this end, consider the case $i > j$ first. By (3.14)

$$E\big|E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j\big]\big| = O(n^{-1/2}). \tag{3.31}$$

Second, when $i < j$, break $\underline{\mathbf{A}}_j^{-1}(z_1)$ into the sum of $\underline{\mathbf{A}}_{ij}^{-1}(z_1)$ and $\underline{\mathbf{A}}_j^{-1}(z_1) - \underline{\mathbf{A}}_{ij}^{-1}(z_1)$, $\underline{\bar{\mathbf{s}}}_j$ into the sum of $\underline{\bar{\mathbf{s}}}_{ij}$ and $\underline{\bar{\mathbf{s}}}_j - \underline{\bar{\mathbf{s}}}_{ij}$, where $\underline{\mathbf{A}}_{ij}(z_1) = \underline{\mathbf{A}}_j(z_1) - n^{-1}\mathbf{s}_i\mathbf{s}_i^T$ and $\underline{\bar{\mathbf{s}}}_{ij} = \underline{\bar{\mathbf{s}}}_j - \mathbf{s}_i/n$. Then, when $i < j$, with notation

$$\underline{\beta}_{ij}(z) = \frac{1}{1 + (1/n)\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z)\mathbf{s}_i},$$

we have

$$E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j\big] = c_{n1} + c_{n2} + c_{n3} + c_{n4}, \tag{3.32}$$

where

$$c_{n1} = E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\underline{\bar{\mathbf{s}}}_{ij}\big],$$

$$c_{n2} = \frac{1}{n}E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i\big],$$

$$c_{n3} = -\frac{1}{n}E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\underline{\beta}_{ij}(z_1)\underline{\bar{\mathbf{s}}}_{ij}\big]$$

and

$$c_{n4} = -\frac{1}{n^2}E_j\big[\big(\beta_{ij}(z_2) - \beta_{ij}^{\mathrm{tr}}(z_2)\big)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\underline{\beta}_{ij}(z_1)\mathbf{s}_i\big].$$

It follows from (3.14) that $E|c_{nj}| \le \mathfrak{M}n^{-1/2}$, $j = 1, 2, 3, 4$. Thus, $\beta_{ij}(z_2)$ in (3.30) can be replaced by $\beta_{ij}^{\mathrm{tr}}(z_2)$, as expected.

In what follows we use the notation $o_{L_1}(1)$ to denote convergence to zero in L_1. Moreover, note that $E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j] = 0$ when $i > j$. This, together with (3.31) and (3.32), implies that

$$\begin{aligned}
&E_j[\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j] \\
&\quad = \frac{1}{n}\sum_{i\neq j}^n E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j] + o_{L_1}(1) \\
&\quad = \frac{1}{n}\sum_{i<j} E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1)\underline{\bar{\mathbf{s}}}_j] + o_{L_1}(1) \\
&\quad = d_{n1} + d_{n2} + d_{n3} + o_{L_1}(1),
\end{aligned} \tag{3.33}$$

where

$$d_{n1} = \frac{1}{n^2}\sum_{i<j} E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i\underline{\beta}_{ij}(z_1)],$$

$$d_{n2} = \frac{1}{n}\sum_{i<j} E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}]$$

and

$$d_{n3} = -\frac{1}{n^2}\sum_{i<j} E_j[\beta_{ij}^{\mathrm{tr}}(z_2)\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i\mathbf{s}_i^T\underline{\mathbf{A}}_{ij}^{-1}(z_1)\underline{\beta}_{ij}(z_1)\bar{\underline{\mathbf{s}}}_{ij}].$$

Here, in the last step, we apply $\bar{\underline{\mathbf{s}}}_j = \mathbf{s}_i/n + \bar{\underline{\mathbf{s}}}_{ij}$ first, then use (1.13) and finally split $\underline{\mathbf{A}}_j^{-1}(z_1)$ into two parts as before.

We claim that the terms d_{n2} and d_{n3} are both negligible. To see this, we first prove the following estimate:

$$E\left|\frac{1}{n}\sum_{i<j}\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}\right|^2 = o(1). \tag{3.34}$$

Indeed, the left-hand side of (3.34) may be expanded as

$$\frac{1}{n^2}\sum_{i_1<j,i_2<j} E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1j}\mathbf{s}_{i_2}^T\mathbf{A}_{i_2j}^{-1}(\bar{z}_2)\underline{\mathbf{A}}_{i_2j}^{-1}(\bar{z}_1)\bar{\underline{\mathbf{s}}}_{i_2j}). \tag{3.35}$$

From (3.12), the term corresponding to $i_1 = i_2$ in (3.35) is bounded by

$$\frac{1}{n^2}\sum_{i_1<j} E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1j}|^2 = O\left(\frac{1}{n}\right).$$

To treat the case $i_1 \neq i_2$, we need to further split $\mathbf{A}_{i_1j}^{-1}(z_2)$ as the sum of $\mathbf{A}_{i_1i_2j}^{-1}(z_2)$ and $\mathbf{A}_{i_1j}^{-1}(z_2) - \mathbf{A}_{i_1i_2j}^{-1}(z_2)$, where $\mathbf{A}_{i_1i_2j}(z_2) = \mathbf{A}_{i_1j}(z_2) - n^{-1}\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T$. Moreover, both $\underline{\mathbf{A}}_{i_1j}^{-1}(z_1)$ and $\bar{\underline{\mathbf{s}}}_{i_1j}$ are also needed to be similarly split. To simplify notation, define

$$\beta_{i_1i_2j}(z) = \frac{1}{1+(1/n)\mathbf{s}_{i_2}^T\mathbf{A}_{i_1i_2j}^{-1}(z)\mathbf{s}_{i_2}},$$

$$\underline{\beta}_{i_1i_2j}(z) = \frac{1}{1+(1/n)\mathbf{s}_{i_2}^T\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z)\mathbf{s}_{i_2}}$$

and

$$\underline{\mathbf{A}}_{i_1i_2j}(z) = \underline{\mathbf{A}}_{i_1j}(z) - \mathbf{s}_{i_2}\mathbf{s}_{i_2}^T, \qquad \bar{\underline{\mathbf{s}}}_{i_1i_2j} = \bar{\underline{\mathbf{s}}}_{i_1j} - \frac{\mathbf{s}_{i_2}}{n},$$

$$\zeta_{i_2j} = \mathbf{s}_{i_2}^T\mathbf{A}_{i_2j}^{-1}(\bar{z}_2)\underline{\mathbf{A}}_{i_2j}^{-1}(\bar{z}_1)\bar{\underline{\mathbf{s}}}_{i_2j}.$$

By (1.13), (3.12), (3.13) and (3.16) we have

$$
\begin{aligned}
&\frac{1}{n}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1j}^{-1}(z_1)\mathbf{s}_{i_2}\zeta_{i_2j})|\\
&\quad=\frac{1}{n}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\underline{\beta}_{i_1i_2j}(z_1)\mathbf{s}_{i_2}\zeta_{i_2j})|\\
&\quad\le\frac{\mathfrak{M}}{n}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\underline{\beta}_{i_1i_2j}(z_1)\mathbf{s}_{i_2}\zeta_{i_2j})|\\
&\qquad+\frac{\mathfrak{M}}{n^2}\big|E\big(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\beta_{i_1i_2j}(z_2)\\
&\qquad\qquad\times\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\underline{\beta}_{i_1i_2j}(z_1)\mathbf{s}_{i_2}\zeta_{i_2j}\big)\big|\\
&\quad\le\frac{\mathfrak{M}}{n}(E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}|^2E|\zeta_{i_2j}|^2)^{1/2}\\
&\qquad+\frac{\mathfrak{M}}{n^2}(E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}|^2E|\zeta_{i_2j}|^2)^{1/2}\\
&\quad=O(n^{-3/8});\\
&\frac{1}{n}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\beta_{i_1i_2j}(z_2)\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}\zeta_{i_2j})|\\
&\quad\le\frac{\mathfrak{M}}{n}(E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}|^4E|\mathbf{s}_{i_2}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}|^4)^{1/4}(E|\zeta_{i_2j}|^2)^{1/2}\\
&\quad=O(n^{-1/2});\\
&\frac{1}{n}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\underline{\beta}_{i_1i_2j}(z_1)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}\zeta_{i_2j})|\\
&\quad\le\frac{\mathfrak{M}}{n}(E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}|^4E|\mathbf{s}_{i_2}^T\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}|^4)^{1/4}(E|\zeta_{i_2j}|^2)^{1/2}\\
&\quad=O(n^{-1/2});\\
&\frac{1}{n^2}|E(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\beta_{i_1i_2j}(z_2)\mathbf{A}_{i_1i_2j}^{-1}(z_2)\\
&\quad\times\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\underline{\beta}_{i_1i_2j}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}\zeta_{i_2j})|\\
&\quad\le\frac{\mathfrak{M}}{n^2}(E|\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\mathbf{s}_{i_2}\mathbf{s}_{i_2}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\mathbf{s}_{i_2}|^4E|\mathbf{s}_{i_2}^T\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{i_1i_2j}|^4)^{1/4}\\
&\qquad\times(E|\zeta_{i_2j}|^2)^{1/2}=O(n^{-3/8}).
\end{aligned}
$$

The above four estimates, together with the fact that

$$E\big(\mathbf{s}_{i_1}^T\mathbf{A}_{i_1i_2j}^{-1}(z_2)\underline{\mathbf{A}}_{i_1i_2j}^{-1}(z_1)\times\bar{\underline{\mathbf{s}}}_{i_1i_2j}\zeta_{i_2j}\big)=0,\qquad i_1\ne i_2,$$

imply that all terms in (3.35) corresponding to $i_1 \neq i_2$ are bounded in absolute value by $\mathfrak{M}n^{-3/8}$, which ensures (3.34).

Consider the term d_{n2} now. In view of (3.9) and (3.14) we may substitute $b_{12}(z_2)$ for $\beta_{ij}^{\mathrm{tr}}(z_2)$ in the term d_{n2} first and then apply (3.34) to conclude that $E|d_{n2}| = o(1)$. As for the term d_{n3}, it follows from (3.9) and (3.14) that $\beta_{ij}^{\mathrm{tr}}(z_2), \underline{\beta}_{ij}(z_1)$ and $\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)\mathbf{s}_i$ can be replaced by $b_{12}(z_2), \underline{b}_{12}(z_1)$ and $\frac{1}{n}\operatorname{tr}\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)$, respectively, where

$$\underline{b}_{12}(z) = \frac{1}{1 + (1/n)E\operatorname{tr}\underline{\mathbf{A}}_{12}^{-1}(z)}$$

[note: $\underline{b}_{12}(z) = b_{12}(z)$]. Moreover, by an inequality similar to (3.8) we have

$$\left|E_j\left[\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}\frac{1}{n}\big(\operatorname{tr}(\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)) - \operatorname{tr}(\mathbf{A}_j^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1))\big)\right]\right|$$
$$\leq \mathfrak{M}\frac{E_j|\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}|}{n}.$$

Therefore, from (3.12) we obtain

$$d_{n3} = -\frac{b_{12}(z_2)\underline{b}_{12}(z_1)}{n^2}E_j\left[\operatorname{tr}(\mathbf{A}_j^{-1}(z_2)\underline{\mathbf{A}}_j^{-1}(z_1))\sum_{i<j}\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}\right] + o_{L_1}(1).$$

As in (3.34) we may prove that (even simpler)

$$E\left|\frac{1}{n}\sum_{i<j}\mathbf{s}_i^T \underline{\mathbf{A}}_{ij}^{-1}(z_1)\bar{\underline{\mathbf{s}}}_{ij}\right|^2 = o(1), \tag{3.36}$$

which then implies that $E|d_{n3}| = o(1)$.

As for d_{n1}, we conclude from (3.9), (3.14) and (3.8) that

$$d_{n1} = \frac{b_{12}(z_2)b_{12}(z_1)}{n^2}\sum_{i<j}\operatorname{tr}E_j[\mathbf{A}_{ij}^{-1}(z_2)\underline{\mathbf{A}}_{ij}^{-1}(z_1)] + o_{L_1}(1)$$
$$= \frac{b_{12}(z_2)b_{12}(z_1)}{n^2}(j-1)\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1))] + o_{L_1}(1).$$

Summarizing the above, we have thus proved that

$$\begin{aligned}
&E_j(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)\\
&\quad = \frac{j-1}{n^2}b_{12}(z_2)b_{12}(z_1)\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1))] + o_{L_1}(1)\\
&\quad = \frac{j-1}{n^2}z_1 z_2\underline{m}(z_1)\underline{m}(z_2)\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1))] + o_{L_1}(1),
\end{aligned} \tag{3.37}$$

using the fact that, by (2.17) in [5] and (3.8),

$$b_{12}(z) \to -z\underline{m}(z). \tag{3.38}$$

This implies (3.29).

3.5. *The limit of* (3.3). Our goal is to show that

$$(3.3) \xrightarrow{\text{i.p.}} 0. \tag{3.39}$$

In view of (3.6) we have

$$(3.3) = \frac{EX_{11}^3}{n} \sum_{j=1}^{n} \sum_{i=1}^{p} [E_j(\mathbf{D}_j(z_2))]_{ii} [E_j(\mathbf{e}_i^T \mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)]. \tag{3.40}$$

We first prove that $\mathbf{e}_i^T \mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j$ above may be replaced by $E(\mathbf{e}_i^T \mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)$. Using martingale decompositions as in (3.21) and the fact that $\mathbf{e}_i^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j = \bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z)\mathbf{e}_i$, we obtain that

$$\begin{aligned} &\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\mathbf{e}_i E_j[\theta_{ij}(z_1)] \\ &\quad = [\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\mathbf{e}_i - E(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\mathbf{e}_i)] E_j[\theta_{ij}(z_1)] \\ &\qquad + E(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\mathbf{e}_i) E_j[\theta_{ij}(z_1)] \\ &\quad = \theta_{ij}(z_2) \times E_j[\theta_{ij}(z_1)] + E(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2)\mathbf{e}_i) E_j[\theta_{ij}(z_1)], \end{aligned} \tag{3.41}$$

where

$$\theta_{ij}(z) = \mathbf{e}_i^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j - E(\mathbf{e}_i^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j) = \sum_{m \neq j}^{n} (E_m - E_{m-1})(\theta_{ijm}(z))$$

and

$$\begin{aligned} \theta_{ijm}(z) &= \mathbf{e}_i^T \mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j - \mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\bar{\mathbf{s}}_{jm} \\ &= \biggl[-\frac{1}{n^2}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z_1)\mathbf{s}_m \mathbf{s}_m^T \mathbf{A}_{jm}^{-1}(z)\beta_{mj}(z)\mathbf{s}_m \\ &\qquad - \frac{1}{n}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m \mathbf{s}_m^T \mathbf{A}_{jm}^{-1}(z)\beta_{mj}(z)\bar{\mathbf{s}}_{jm} + \frac{1}{n}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m\biggr]. \end{aligned}$$

As in (3.19), one can verify that

$$\begin{aligned} &E|n^{-1}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m|^k = O(n^{-k}), \qquad k = 2 \text{ or } 4, \\ &E|n^{-1}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m|^8 = O(n^{-6}). \end{aligned} \tag{3.42}$$

Thus, for $k = 2$ or 4, via (3.12),

$$E\left|\frac{1}{n}\mathbf{e}_i^T \mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m \mathbf{s}_m^T \mathbf{A}_{jm}^{-1}(z)\bar{\mathbf{s}}_{jm}\right|^k = O(n^{-2}\varepsilon_n^{k-2})$$

and, via (3.10),

$$E\left|\frac{1}{n^2}\mathbf{e}_i^T\mathbf{A}_{jm}^{-1}(z_1)\mathbf{s}_m\mathbf{s}_m^T\mathbf{A}_{jm}^{-1}(z)\mathbf{s}_m\right|^k = O\big(n^{-2-3(k-2)/4}\big).$$

These yield that $E|\theta_{ijm}(z)|^2 = O(n^{-2})$, $E|\theta_{ijm}(z)|^4 = O(n^{-2}\varepsilon_n)$ and then

$$\begin{aligned} &E|\theta_{ij}(z)|^2 = O(n^{-1}),\\ &E|\theta_{ij}(z)|^4 = O(n^{-1}\varepsilon_n). \end{aligned} \tag{3.43}$$

Therefore,

$$\begin{aligned} &\left[E\sum_{i=1}^{p}|[E_j(\mathbf{D}_j(z_2))]_{ii}E_j(\theta_{ij}(z_1))|\right]^2\\ &\quad\le \sum_{i=1}^{p}E|\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j|^2\sum_{i=1}^{p}E|\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{e}_iE_j(\theta_{ij}(z_1))|^2\\ &\quad\le \mathfrak{M}\sum_{i=1}^{p}[E|\theta_{ij}(z_2)|^4E|\theta_{ij}(z_1)|^4]^{1/2}\\ &\qquad+\mathfrak{M}\sum_{i=1}^{p}|E(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{e}_i)|^2E|\theta_{ij}(z_1)|^2\\ &\quad= O(\varepsilon_n). \end{aligned} \tag{3.44}$$

Here, by (3.18)

$$\begin{aligned} \sum_{i=1}^{p}|E(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{e}_i)|^2E|\theta_{ij}(z_2)|^2 &\le \frac{\mathfrak{M}}{n}\sum_{i=1}^{p}|E(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{e}_i)|^2\\ &\le \frac{\mathfrak{M}}{n}E(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{A}_j^{-1}(\bar{z}_2)\bar{\mathbf{s}}_j)\\ &\le \frac{\mathfrak{M}}{n}. \end{aligned}$$

Thus, $\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j$ involved in (3.40) may be replaced by $E(\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)$, as expected.

In addition, by (3.18) and (A.2)

$$\begin{aligned} &E\sum_{i=1}^{p}|[E_j(\mathbf{D}_j(z_2))]_{ii}E(\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)|\\ &\quad\le E\sum_{i=1}^{p}[E_j(\mathbf{A}_j^{-1}(\bar{z}_2)\bar{\mathbf{s}}_j\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))]_{ii}|E(\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)|\\ &\quad\le \max_i|E(\mathbf{e}_i^T\mathbf{A}_1^{-1}(z_1)\bar{\mathbf{s}}_1)|E(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2)\mathbf{A}_j^{-1}(\bar{z}_2)\bar{\mathbf{s}}_j)\to 0. \end{aligned} \tag{3.45}$$

It follows from (3.44) and (3.45) that

$$E\sum_{i=1}^{p}|[E_j(\mathbf{D}_j(z_2))]_{ii}E_j(\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)| \to 0, \tag{3.46}$$

which then ensures (3.39).

3.6. *The limit of* (3.4). The goal in this section is to prove that

$$\begin{aligned}(3.4) &= \frac{2z_1^2z_2^2\underline{m}^2(z_1)\underline{m}^2(z_2)}{n}\\ &\quad\times\sum_{j=1}^{n}\frac{(j-1)^2}{n^4}[\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)))]^2+o_p(1).\end{aligned}\tag{3.47}$$

First, (3.5) shows that (3.4) is equal to

$$\begin{aligned}&\frac{E|X_{11}|^4-3}{n}\sum_{j=1}^{n}\sum_{i=1}^{p}E_j(\mathbf{D}_j(z_1))_{ii}E_j(\mathbf{D}_j(z_2))_{ii}\\ &\qquad+\frac{2}{n}\sum_{j=1}^{n}\operatorname{tr}[E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2))].\end{aligned}\tag{3.48}$$

To prove (3.47), the strategy is to substitute $E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z))$ for each $\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z)$ involved in $E_j(\mathbf{D}_j(z))$ by a martingale method. As we shall see, the above first term converges to zero in probability and the second term has a close connection with (3.2).

Consider the second term of (3.48) first. Write

$$\begin{aligned}&\operatorname{tr}[E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2))]\\ &\quad=E_j[\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2)\underline{\bar{\mathbf{s}}}_j\underline{\bar{\mathbf{s}}}_j^T\underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j]\\ &\quad=E_j[\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2)\underline{\bar{\mathbf{s}}}_j\underline{\bar{\mathbf{s}}}_j^T\underline{\mathbf{A}}_j^{-1}(z_2)E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)]+f_n,\end{aligned}\tag{3.49}$$

where

$$f_n=E_j\big[\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2)\underline{\bar{\mathbf{s}}}_j\underline{\bar{\mathbf{s}}}_j^T\underline{\mathbf{A}}_j^{-1}(z_2)\big(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j-E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)\big)\big].$$

We claim that

$$E|f_n|=o(1). \tag{3.50}$$

To see this, let $\underline{E}_{ij} = E(\cdot|\mathbf{s}_1, \ldots, \mathbf{s}_i, \underline{\mathbf{s}}_{j+1}, \ldots, \underline{\mathbf{s}}_n)$. Then, recalling the definitions of $\underline{\mathbf{A}}_j^{-1}(z)$ and $\bar{\underline{\mathbf{s}}}_j$ as before, we obtain a martingale decomposition

$$\begin{aligned}
&\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\big(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j - E_{jj}(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)\big) \\
&\quad = \sum_{i=j+1}^{n} \big(\underline{E}_{ij}[\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j] - \underline{E}_{(i-1)j}[\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j]\big) \\
&\quad = \sum_{i=j+1}^{n} \big(\underline{E}_{ij} - \underline{E}_{(i-1)j}\big)[\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j - \bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\bar{\mathbf{s}}_{ij}] \\
&\quad = f_{n1} + f_{n2},
\end{aligned}$$

where

$$f_{n1} = \frac{1}{n}\sum_{i=j+1}^{n} \big(\underline{E}_{ij} - \underline{E}_{(i-1)j}\big)[\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\mathbf{s}_i \beta_{ij}(z_1)]$$

and

$$f_{n2} = -\frac{1}{n}\sum_{i=j+1}^{n} \big(\underline{E}_{ij} - \underline{E}_{(i-1)j}\big)[\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\mathbf{s}_i\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_1)\bar{\mathbf{s}}_{ij}\beta_{ij}(z_1)].$$

Note that $\bar{\underline{\mathbf{s}}}_j$ is independent of $\mathbf{s}_i$ for $i > j$. Then applying (3.12) yields

$$E|f_{n1}|^2 \le \frac{\mathfrak{M}}{n^2}\sum_{i=j+1}^{n} E|\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\mathbf{s}_i|^2 = O\left(\frac{1}{n}\right)$$

and

$$\begin{aligned}
E|f_{n2}|^2 &\le \frac{\mathfrak{M}}{n^2}\sum_{i=j+1}^{n} E|\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\mathbf{s}_i\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_1)\bar{\mathbf{s}}_{ij}|^2 \\
&\le \frac{\mathfrak{M}}{n^2}\sum_{i=j+1}^{n} \big(E|\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\mathbf{A}_{ij}^{-1}(z_1)\mathbf{s}_i|^4 E|\mathbf{s}_i^T \mathbf{A}_{ij}^{-1}(z_1)\bar{\mathbf{s}}_{ij}|^4\big)^{1/2} \\
&= O\left(\frac{1}{n}\right),
\end{aligned}$$

which ensures that

$$E\big|\bar{\underline{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2)\big(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j - E_{jj}(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)\big)\big|^2 = O\left(\frac{1}{n}\right).$$

So (3.50) follows from the above estimate and

$$E|\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2)\underline{\bar{\mathbf{s}}}_j|^2 = O(1),$$

which may be obtained immediately by checking the argument of (3.18).

As in (3.50) we may also prove that

$$\begin{aligned} &E\big|E_j[\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2) \\ &\quad\times \underline{\bar{\mathbf{s}}}_j\big(\underline{\bar{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2) - E_j(\underline{\bar{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2))\big)E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)]\big| = o(1). \end{aligned} \tag{3.51}$$

Therefore, combining (3.49)–(3.51) with (3.37) we have

$$\begin{aligned} &\operatorname{tr}[E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2))] \\ &\quad= E_j[\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\underline{\mathbf{A}}_j^{-1}(z_2)\underline{\bar{\mathbf{s}}}_j E_j(\underline{\bar{\mathbf{s}}}_j^T \underline{\mathbf{A}}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)] \\ &\qquad + o_{L_1}(1) \\ &\quad= E_j(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1))E_j(\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j) \\ &\qquad + o_{L_1}(1) \\ &\quad= \frac{(j-1)^2}{n^4} z_1^2 z_2^2 \underline{m}^2(z_1)\underline{m}^2(z_2)[\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j\underline{\mathbf{A}}_j^{-1}(z_1))]^2 \\ &\qquad + o_{L_1}(1). \end{aligned} \tag{3.52}$$

We now turn to the first term in (3.48) and claim that

$$\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{p} E_j(\mathbf{D}_j(z_1))_{ii} E_j(\mathbf{D}_j(z_2))_{ii} \xrightarrow{\text{i.p.}} 0. \tag{3.53}$$

Indeed, it follows from (3.41) that

$$\begin{aligned} &E\left|\sum_{i=1}^{p} E_j(\mathbf{D}_j(z_2))_{ii} E_j(\theta_{ij}(z_1)\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\mathbf{e}_i)\right| \\ &\quad\le \sum_{i=1}^{p} E|E_j(\mathbf{D}_j(z_2))_{ii} E_j(\theta_{ij}(z_1))^2| \\ &\qquad + \sum_{i=1}^{p} E|E_j(\mathbf{D}_j(z_2))_{ii} E_j(\theta_{ij}(z_1))E(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\mathbf{e}_i)|. \end{aligned}$$

The second term above is not greater than

$$\max_i |E(\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\mathbf{e}_i)| \sum_{i=1}^{p} E|E_j(\mathbf{D}_j(z_2))_{ii} E_j(\theta_{ij}(z_1))|,$$

which converges to zero by (3.44) and (A.2). Moreover, by (3.18) and (3.43)

$$\left(\sum_{i=1}^{p} E|E_j(\mathbf{D}_j(z_2))_{ii}E_j(\theta_{ij}(z_1))^2|\right)^2$$
$$\leq \sum_{i=1}^{p} E|(\mathbf{D}_j(z_2))_{ii}|^2 \sum_{i=1}^{p} E|\theta_{ij}(z_1)|^4$$
$$\leq E\left(\sum_{i=1}^{p} \bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(\bar{z}_1)\mathbf{e}_i\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j\right)^2 \sum_{i=1}^{p} E|\theta_{ij}(z_1)|^4$$
$$= O(\varepsilon_n).$$

In addition, it follows from Lemma 6 and (3.46) that

$$E\left|\sum_{i=1}^{p} E_j(\mathbf{D}_j(z_2))_{ii}E(\mathbf{e}_i^T\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{e}_i)\right|$$
$$\leq \max_i |E(\mathbf{e}_i^T\mathbf{A}_1^{-1}(z_1)\bar{\mathbf{s}}_1)| \sum_{i=1}^{p} E|E_j(\mathbf{D}_j(z_2))_{ii}E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{e}_i)| \to 0.$$

Consequently, the proof of (3.53) is complete. Thus, (3.47) follows from (3.52), (3.53) and (3.48).

3.7. *The limit of* (3.28). Note that (see [5], (2.18))

$$\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)))\left[1 - \frac{(j-1)p}{n^2}\frac{\underline{m}_n(z_1)\underline{m}_n(z_2)}{(1+\underline{m}_n(z_1))(1+\underline{m}_n(z_2))}\right]$$
$$= \frac{p}{z_1z_2(1+\underline{m}_n(z_1))(1+\underline{m}_n(z_2))} + l_n, \tag{3.54}$$

where $E|l_n| \leq \mathfrak{M}\sqrt{n}$ and $\underline{m}_n(z)$ is defined like $m_n(z)$, but corresponding to $\underline{m}(z)$. Obviously, $\underline{m}_n(z) \to \underline{m}(z)$. This implies that

$$\frac{(j-1)z_1z_2\underline{m}(z_1)\underline{m}(z_2)}{n^2}\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)))$$
$$= \frac{z_1z_2(1+\underline{m}(z_1))(1+\underline{m}(z_2))}{p}\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\mathbf{A}_j^{-1}(z_1)))$$
$$- 1 + o_{L_1}(1),$$

which, together with (3.37) and (3.52), leads to

$$4\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))] + 2\operatorname{tr}(E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2)))$$
$$= \frac{4(j-1)z_1z_2\underline{m}(z_1)\underline{m}(z_2)}{n^2}\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1)))$$

$$+\frac{2(j-1)^2 z_1^2 z_2^2 \underline{m}^2(z_1)\underline{m}^2(z_2)}{n^4}[\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1)))]^2$$
$$+ o_{L_1}(1)$$
$$= -2 + 2z_1^2 z_2^2 \big(1+\underline{m}(z_1)\big)^2\big(1+\underline{m}(z_2)\big)^2 \frac{[\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_2))E_j(\underline{\mathbf{A}}_j^{-1}(z_1)))]^2}{p^2}$$
$$+ o_{L_1}(1).$$

Further, we conclude from (3.54) that

$$\frac{1}{np^2}\sum_{j=1}^{n}[\operatorname{tr}(E_j(\mathbf{A}_j^{-1}(z_1))E_j(\mathbf{A}_j^{-1}(z_2)))]^2$$
$$= \frac{1}{z_1^2 z_2^2(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}$$
$$\times \frac{1}{n}\sum_{j=1}^{n}\frac{1}{(1-(j-1)p/n^2(\underline{m}(z_1)\underline{m}(z_2)/((1+\underline{m}(z_1))(1+\underline{m}(z_2)))))^2}$$
$$+ o_p(1)$$
$$\xrightarrow{\text{i.p.}} \frac{1}{z_1^2 z_2^2(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}$$
$$\times \int_0^1 \frac{dx}{(1-x(c\underline{m}(z_1)\underline{m}(z_2)/((1+\underline{m}(z_1))(1+\underline{m}(z_2)))))^2}$$
$$= \frac{1}{z_1^2 z_2^2(1+\underline{m}(z_1))(1+\underline{m}(z_2))[(1+\underline{m}(z_1))(1+\underline{m}(z_2)) - c\underline{m}(z_1)\underline{m}(z_2)]}.$$

It follows that

$$\begin{aligned}(3.28) = {} & z_1 z_2 \underline{m}(z_1)\underline{m}(z_2)\frac{1}{n}\sum_{j=1}^{n}\big[4\operatorname{tr}[E_j(\mathbf{A}_j^{-1}(z_1)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_2))] \\ & \qquad + 2\operatorname{tr}[E_j(\mathbf{D}_j(z_1))E_j(\mathbf{D}_j(z_2))]\big] \\ & + o_p(1) \\ & \xrightarrow{\text{i.p.}} \frac{2cz_1 z_2 \underline{m}^2(z_1)\underline{m}^2(z_2)}{(1+\underline{m}(z_1))(1+\underline{m}(z_2)) - c\underline{m}(z_1)\underline{m}(z_2)}.\end{aligned} \tag{3.55}$$

4. Tightness of $\hat{M}_n^{(1)}(z)$ and convergence of $M_n^{(2)}(z)$. First, we proceed to prove the tightness of $\hat{M}_n^{(1)}(z)$ for $z \in \mathcal{C}$, which is a truncated version of $M_n(z)$ as in (1.9). By (3.12) we have

$$E\left|\sum_{i=1}^{m} a_i \sum_{j=1}^{n} Y_j(z_i)\right|^2 = \sum_{j=1}^{n} E\left|\sum_{i=1}^{m} a_i Y_j(z_i)\right|^2 \le \mathfrak{M}, \qquad v_0 = \Im z_i,$$

which ensures that condition (i) of Theorem 12.3 in [6] is satisfied, as pointed out in [5]. Here $Y_j(z)$ is defined in (3.26). Condition (ii) of Theorem 12.3 in [6] will be verified if the following holds:

$$E\frac{|M_n^{(1)}(z_1) - M_n^{(1)}(z_2)|^2}{|z_1 - z_2|^2} \le \mathfrak{M} \qquad \text{for } z_1, z_2 \in \mathcal{C}_n^+ \cup \mathcal{C}_n^-. \tag{4.1}$$

In the sequel, since $\mathcal{C}_n^+$ and $\mathcal{C}_n^-$ are symmetric, we shall prove the above inequality on $\mathcal{C}_n^+$ only. Throughout this section, all bounds including $O(\cdot)$ and $o(\cdot)$ expressions hold uniformly for $z \in \mathcal{C}_n^+$.

In view of our truncation steps, (1.9a) and (1.9b) in [5] apply to our case as well, that is, for any $\eta_1 > (1+\sqrt{c})^2$, $0 < \eta_2 < I(0,1)(c)(1-\sqrt{c})^2$ and any positive l

$$P(\|\mathbf{S}\| \ge \eta_1) = o(n^{-l}), \qquad P\big(\lambda_{\min}(\mathbf{S}) \le \eta_2\big) = o(n^{-l}). \tag{4.2}$$

Note that when either $z \in \mathcal{C}_u$ or $z \in \mathcal{C}_l$ and $u_l < 0$, $\|\mathbf{A}_j^{-1}(z)\|$ is bounded in n. But this is not the case for $z \in \mathcal{C}_r$ or $z \in \mathcal{C}_l$ and $u_l > 0$. In general, for $z \in \mathcal{C}_n^+$, we have

$$\|\mathbf{A}_j^{-1}(z)\| \le M + v^{-1} I\big(\|\mathbf{A}_j\| \ge h_r \text{ or } \lambda_{\min}(\mathbf{A}_j) \le h_l\big). \tag{4.3}$$

Here, $\mathbf{A}_j = \mathbf{S} - \mathbf{s}_j\mathbf{s}_j^T$, $h_r \in ((1+\sqrt{c})^2, u_r)$ and $h_l \in (u_l, (1-\sqrt{c})^2)$.

Note that $\mathbf{A}^{-1}(z_1) - \mathbf{A}^{-1}(z_2) = (z_2 - z_1)\mathbf{A}^{-1}(z_1)\mathbf{A}^{-1}(z_2)$. As in Section 3.3, we then write

$$\begin{aligned} &\frac{M_n^{(1)}(z_1) - M_n^{(1)}(z_2)}{z_1 - z_2} \\ &\quad = -\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})[\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z_1)\mathbf{A}^{-1}(z_2)\bar{\mathbf{s}} \\ &\qquad\qquad - \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j]. \end{aligned} \tag{4.4}$$

Moreover, expanding the above difference we get

$$\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z_1)\mathbf{A}^{-1}(z_2)\bar{\mathbf{s}} - \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j = q_{n1} + q_{n2} + q_{n3},$$

where

$$q_{n1} = (\bar{\mathbf{s}}^T - \bar{\mathbf{s}}_j^T)\mathbf{A}^{-1}(z_1)\mathbf{A}^{-1}(z_2)\bar{\mathbf{s}},$$

$$q_{n2} = \bar{\mathbf{s}}_j^T\big(\mathbf{A}^{-1}(z_1)\mathbf{A}^{-1}(z_2) - \mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\big)\bar{\mathbf{s}}$$

and

$$q_{n3} = \bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)(\bar{\mathbf{s}} - \bar{\mathbf{s}}_j).$$

It follows from (1.8), (3.12), (4.3) and (4.2) that

$$E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})q_{n3}\right|^2 \le \frac{1}{n}\sum_{j=1}^{n} E|\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j|^2$$
$$\le \mathfrak{M} + \mathfrak{M}n^8\rho_n^{-4}P\big(\|\mathbf{A}_1\| \ge h_r \text{ or } \lambda_{\min}(\mathbf{A}_1) \le h_l\big)$$
$$\le \mathfrak{M},$$

where we use, on the event $(\|\mathbf{A}_j\| \ge h_r$ or $\lambda_{\min}(\mathbf{A}_j) \le h_l)$, by (3.1),

$$|\bar{\mathbf{s}}_j^T \mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j| \le \|\bar{\mathbf{s}}_j\|\|\mathbf{s}_j\|\|\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\| \le \mathfrak{M}v^{-2}n^2 \le \mathfrak{M}n^4\rho_n^{-2}. \tag{4.5}$$

For q_{n2}, expanding its difference term by term we have

$$q_{n2} = q_{n2}^{(1)} + \cdots + q_{n2}^{(6)},$$

where

$$q_{n2}^{(1)} = \frac{1}{n^2}\bar{\mathbf{s}}_j^T \beta_j(z_1)\beta_j(z_2)\tilde{\mathbf{A}}_j(z_1)\tilde{\mathbf{A}}_j(z_2)\bar{\mathbf{s}}_j,$$
$$q_{n2}^{(2)} = -\frac{1}{n}\bar{\mathbf{s}}_j^T \beta_j(z_1)\tilde{\mathbf{A}}_j(z_1)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j,$$
$$q_{n2}^{(3)} = -\frac{1}{n}\bar{\mathbf{s}}_j^T \beta_j(z_2)\mathbf{A}_j^{-1}(z_1)\tilde{\mathbf{A}}_j(z_2)\bar{\mathbf{s}}_j,$$
$$q_{n2}^{(4)} = \frac{1}{n^3}\bar{\mathbf{s}}_j^T \beta_j(z_1)\beta_j(z_2)\tilde{\mathbf{A}}_j(z_1)\tilde{\mathbf{A}}_j(z_2)\mathbf{s}_j$$

and

$$q_{n2}^{(5)} = -\frac{1}{n^2}\bar{\mathbf{s}}_j^T \beta_j(z_1)\tilde{\mathbf{A}}_j(z_1)\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j,$$
$$q_{n2}^{(6)} = -\frac{1}{n^2}\bar{\mathbf{s}}_j^T \beta_j(z_2)\mathbf{A}_j^{-1}(z_1)\tilde{\mathbf{A}}_j(z_2)\mathbf{s}_j.$$

We conclude from (3.14), (4.2), (4.3) and (4.5) that

$$E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})q_{n2}^{(6)}\right|^2 \le \mathfrak{M} + \mathfrak{M}v^{-8}n^8P\big(\|\mathbf{S}\| \ge h_r \text{ or } \lambda_{\min}(\mathbf{A}_1) \le h_l\big) \le \mathfrak{M},$$

where we use, on the event $(\|\mathbf{S}\| \ge h_r$ or $\lambda_{\min}(\mathbf{A}_1) \le h_l)$,

$$|\beta_j(z)| = |1 - n^{-1}\mathbf{s}_j^T \mathbf{A}^{-1}(z)\mathbf{s}_j| \le 1 + n^{-1}v^{-1}\|\mathbf{s}_j\|^2 \le \mathfrak{M}v^{-1}n \tag{4.6}$$

by (3.7). Similar argument shows that

$$E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})q_{n2}^{(6)}\right|^2 = O(1), \qquad j = 2, \ldots, 5.$$

Moreover, write $q_{n1} = q_{n1}^{(1)} + q_{n1}^{(2)} + q_{n1}^{(3)}$, where

$$q_{n1}^{(1)} = \frac{1}{n^2}\beta_j(z_1)\beta_j(z_2)\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\mathbf{s}_j,$$
$$q_{n1}^{(2)} = \frac{1}{n}\beta_j(z_1)\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j$$

and

$$q_{n1}^{(3)} = -\frac{1}{n^2}\beta_j(z_1)\beta_j(z_2)\mathbf{s}_j^T\mathbf{A}_j^{-1}(z_1)\tilde{\mathbf{A}}_j(z_2)\bar{\mathbf{s}}_j.$$

The argument for $q_{n2}^{(6)}$ also works for $q_{n1}^{(j)}$, $j = 1, 2, 3$, and thus,

$$E\left|\sqrt{n}\sum_{j=1}^{n}(E_j - E_{j-1})q_{n1}\right|^2 \le \mathfrak{M}.$$

The proof of (4.1) is complete.

Next, consider $M_n^{(2)}(z)$. By $\bar{\mathbf{s}} = n^{-1}\sum_{i=1}^n \mathbf{s}_i$, (1.13) and an equality similar to (A.3) we obtain

$$\begin{aligned}\sqrt{n}E(\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}E(\beta_i(z)\mathbf{s}_i^T\mathbf{A}_i^{-1}(z)\bar{\mathbf{s}}) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}E(\beta_i(z)\mathbf{s}_i^T\mathbf{A}_i^{-1}(z)\bar{\mathbf{s}}_i) \\ &\quad + \frac{1}{n^{3/2}}\sum_{i=1}^{n}E(\beta_i(z)\mathbf{s}_i^T\mathbf{A}_i^{-1}(z)\mathbf{s}_i) \\ &= \frac{b_1(z)}{\sqrt{n}}E(\operatorname{tr}\mathbf{A}_1^{-1}(z)) + b_1(z)t_{n1} + b_1(z)t_{n2},\end{aligned}$$

where

$$t_{n1} = -\sqrt{n}E(\beta_1(z)\xi_1(z)\mathbf{s}_1^T\mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1),$$
$$t_{n2} = -\frac{1}{\sqrt{n}}E(\beta_1(z)\xi_1(z)\mathbf{s}_1^T\mathbf{A}_1^{-1}(z)\mathbf{s}_1).$$

Again, using an equality similar to (A.3) further gives

$$t_{n1} = t_{n1}^{(1)} + t_{n1}^{(2)}, \qquad t_{n2} = t_{n2}^{(1)} + t_{n2}^{(2)},$$

where

$$t_{n1}^{(1)} = -\sqrt{n}b_1(z)E(\xi_1(z)\mathbf{s}_1^T\mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1),$$
$$t_{n1}^{(2)} = \sqrt{n}b_1(z)E(\beta_1(z)\xi_1^2(z)\mathbf{s}_1^T\mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1)$$

and

$$t_{n2}^{(1)} = -\frac{b_1(z)}{\sqrt{n}} E(\xi_1(z)\mathbf{s}_1^T \mathbf{A}_1^{-1}(z)\mathbf{s}_1),$$

$$t_{n2}^{(2)} = \frac{b_1(z)}{\sqrt{n}} E(\beta_1(z)\xi_1^2(z)\mathbf{s}_1^T \mathbf{A}_1^{-1}(z)\mathbf{s}_1).$$

Note that $|b_1(z)| \le \mathfrak{M}$ for $z \in \mathcal{C}_n$ (see [5], three lines below (3.6)). It follows from (3.9), (3.10), (3.12), (4.2), (4.3) and (4.6) that

$$|t_{n1}^{(2)}| \le \mathfrak{M}\varepsilon_n + \mathfrak{M}n^{10}\rho_n^{-4} P(\|\mathbf{S}\| \ge h_r \text{ or } \lambda_{\min}(\mathbf{A}_1) \le h_l) \le \mathfrak{M}\varepsilon_n,$$

because $|\beta_i(z)\xi_i^2(z)\mathbf{s}_i^T \mathbf{A}_i^{-1}(z)\bar{\mathbf{s}}_i| \le n^5 v^{-4}$ on the event $(\|\mathbf{S}\| \ge h_r$ or $\lambda_{\min}(\mathbf{A}_1) \le h_l)$. This argument clearly applies to $t_{n2}^{(2)}$ as well and so $|t_{n2}^{(2)}| \le \mathfrak{M}\varepsilon_n$. Notice that $\frac{1}{n}E[\operatorname{tr}(\mathbf{A}_1^{-1}(z))] = E[\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\mathbf{e}_m]$. This and (3.6) show that

$$\begin{aligned}
|t_{n1}^{(1)}| &= \left| -\frac{b_1(z)EX_{11}^3}{\sqrt{n}} \sum_{m=1}^{p} E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\mathbf{e}_m \mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1) \right| \\
&= \left| -\frac{b_1(z)EX_{11}^3(1/n)E\operatorname{tr}(\mathbf{A}_1^{-1}(z))}{\sqrt{n}} \sum_{m=1}^{p} E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1) \right| + o(1) \\
&\le \mathfrak{M}\left| b_1(z)EX_{11}^3 \frac{1}{n} E\operatorname{tr}(\mathbf{A}_1^{-1}(z)) \right| \max_m \sqrt{n}|E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1)| + o(1) \\
&= o(1),
\end{aligned}$$

where we make use of the facts that by (A.2), (4.2) and (4.3),

$$\max_m \sqrt{n}|E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1)| = o(1), \qquad \frac{1}{n}E\operatorname{tr}(\mathbf{A}_1^{-1}(z)) = O(1)$$

and that by (A.7), (3.43), (4.2) and (4.3),

$$\begin{aligned}
&E\big|\big(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\mathbf{e}_m - E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\mathbf{e}_m)\big)\big(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1 - E(\mathbf{e}_m^T \mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1)\big)\big| \\
&\qquad \le \mathfrak{M}n^{-1} + \mathfrak{M}v^{-2} n P(\|\mathbf{A}_1\| \ge h_r \text{ or } \lambda_{\min}(\mathbf{A}_1) \le h_l) = O(n^{-1}).
\end{aligned}$$

Note that $n^{-1}E(\xi_i(z)\mathbf{s}_i^T \mathbf{A}_i^{-1}(z)\mathbf{s}_i) = E\gamma_i^2(z) + n^{-2}E(\operatorname{tr}\mathbf{A}_i^{-1}(z) - E\operatorname{tr}\mathbf{A}_i^{-1}(z))^2$ and then applying (3.9), (3.10), (4.2) and (4.3) gives $t_{n2}^{(1)} = O(n^{-1/2})$.

Summarizing the above we obtain

$$\sqrt{n}E(\bar{\mathbf{s}}^T \mathbf{A}^{-1}(z)\bar{\mathbf{s}}) = \frac{b_1(z)}{n^{1/2}} E(\operatorname{tr}\mathbf{A}_1^{-1}(z)) + o(1).$$

Moreover, it is proven in [5], Section 4, that $n(E\operatorname{tr}\mathbf{A}^{-1}(z)/n - c_n m_n(z))$ is bounded for $z \in \mathcal{C}_n$. In addition, by (3.8), (4.2) and (4.3) we have

$$\sqrt{n}\left| \frac{E\operatorname{tr}\mathbf{A}_1^{-1}(z)}{n} - \frac{E\operatorname{tr}\mathbf{A}^{-1}(z)}{n} \right| \le \frac{\mathfrak{M}}{\sqrt{n}}.$$

It follows that $n(E\operatorname{tr}\mathbf{A}_1^{-1}(z)/n - c_n m_n(z))$ is bounded. This, together with the boundedness of $b_1(z)$, shows that

$$\sup_{z\in\mathcal{C}_n}\sqrt{n}\left(E\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \frac{c_n m_n(z)}{1+c_n m_n(z)}\right) \to 0.$$

5. Proofs of Lemma 1, Theorems 1 and 2.

PROOF OF LEMMA 1. To finish Lemma 1, $\bar{\mathbf{s}}^T\bar{\mathbf{s}} - c_n$ needs to be written as a sum of martingale difference sequence so that we can get a CLT for $\bar{\mathbf{s}}^T\bar{\mathbf{s}} - c_n$ and, more importantly, obtain the asymptotic covariance between $\bar{\mathbf{s}}^T\bar{\mathbf{s}} - c_n$ and $\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}}$.

Thus, write

$$\begin{aligned}
\sqrt{n}(\bar{\mathbf{s}}^T\bar{\mathbf{s}} - c_n) &= \sqrt{n}\sum_{j=1}^n (E_j - E_{j-1})(\bar{\mathbf{s}}^T\bar{\mathbf{s}}) \\
&= \sqrt{n}\sum_{j=1}^n (E_j - E_{j-1})(\bar{\mathbf{s}}^T\bar{\mathbf{s}} - \bar{\mathbf{s}}_j^T\bar{\mathbf{s}}_j) \\
&= \sqrt{n}\sum_{j=1}^n (E_j - E_{j-1})\left(2\frac{\bar{\mathbf{s}}_j^T\mathbf{s}_j}{n} + \frac{\mathbf{s}_j^T\mathbf{s}_j}{n^2}\right) \\
&= \frac{2}{\sqrt{n}}\sum_{j=1}^n E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j) + o_p(1),
\end{aligned} \tag{5.1}$$

because

$$E\left|\sqrt{n}\sum_{j=1}^n (E_j - E_{j-1})\left(\frac{\mathbf{s}_j^T\mathbf{s}_j}{n^2}\right)\right|^2 = \frac{1}{n^3}\sum_{j=1}^n E|E_j(\mathbf{s}_j^T\mathbf{s}_j) - p|^2 = O\left(\frac{1}{n}\right).$$

From (3.12) we have

$$\begin{aligned}
\sum_{j=1}^n E\left|\frac{1}{\sqrt{n}}E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)\right|^2 I\left(\frac{1}{\sqrt{n}}E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j) \ge \varepsilon\right) &\le \frac{1}{\varepsilon^2}\sum_{j=1}^n E\left|\frac{1}{\sqrt{n}}E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)\right|^4 \\
&= O(n^{-1}),
\end{aligned}$$

which implies condition (ii) of Lemma 3. Look at condition (i) of Lemma 3 next. It is easily seen that

$$E_{j-1}[E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)]^2 = E_j(\bar{\mathbf{s}}_j^T)E_j(\bar{\mathbf{s}}_j) = \frac{1}{n^2}\sum_{k_1<j,k_2<j}\mathbf{s}_{k_1}^T\mathbf{s}_{k_2}.$$

Furthermore, for the term corresponding to $k_1 = k_2$, we have

$$E\left|\frac{1}{n^2}\sum_{k_1<j}[\mathbf{s}_{k_1}^T\mathbf{s}_{k_1} - E(\mathbf{s}_{k_1}^T\mathbf{s}_{k_1})]\right|^2 = \frac{1}{n^4}\sum_{k_1<j}E|\mathbf{s}_{k_1}^T\mathbf{s}_{k_1} - E(\mathbf{s}_{k_1}^T\mathbf{s}_{k_1})|^2 = O\left(\frac{1}{n^2}\right).$$

On the other hand, when $k_1 \neq k_2$,

$$\begin{aligned} E\left|\frac{1}{n^2}\sum_{k_1\neq k_2}\mathbf{s}_{k_1}^T\mathbf{s}_{k_2}\right|^2 &= \frac{1}{n^4}\sum_{k_1\neq k_2, h_1\neq h_2}E[\mathbf{s}_{k_1}^T\mathbf{s}_{k_2}\mathbf{s}_{h_1}^T\mathbf{s}_{h_2}] \\ &= \frac{2}{n^4}\sum_{k_1\neq k_2}E(\mathbf{s}_{k_1}^T\mathbf{s}_{k_2})^2 = O\left(\frac{1}{n}\right). \end{aligned}$$

It follows that

$$\frac{4}{n}\sum_{j=1}^{n}E_{j-1}[E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)]^2 = \frac{4}{n}\sum_{j=1}^{n}\frac{c(j-1)}{n} + o_p(1) \xrightarrow{\text{i.p.}} 4c\int_0^1 x\,dx = 2c. \tag{5.2}$$

Therefore, by Lemma 3

$$\sqrt{n}(\bar{\mathbf{s}}^T\bar{\mathbf{s}} - c_n) \xrightarrow{D} N(0, 2c). \tag{5.3}$$

We conclude from Sections 2 and 3 that $\hat{M}_n(z)$ converges weakly to a Gaussian process on $\mathcal{C}$. Moreover, $m_n(z) \to m(z)$ uniformly on $\mathcal{C}$ by (4.2) in [5] and (1.2). These, together with (1.12), (5.3), (3.26) and (5.1), give, for any constants a_1 and a_2,

$$\begin{aligned} &a_1X_n(z) + a_2\sqrt{n}\big(g(\|\bar{\mathbf{s}}\|^2) - g(c_n)\big) \\ &\quad= \tilde{a}_1(z)\sqrt{n}\left[\bar{\mathbf{s}}^T\mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \frac{c_nm_n(z)}{1+c_nm_n(z)}\right] \\ &\qquad+ \tilde{a}_2(z)\sqrt{n}(\|\bar{\mathbf{s}}\|^2 - c_n) + o_p(1) \\ &\quad= \sum_{j=1}^{n}l_j(z) + o_p(1), \end{aligned} \tag{5.4}$$

where $\tilde{a}_1(z) = a_1(1+cm(z))^2/c$, $\tilde{a}_2(z) = a_2g'(c_n) - a_1m(z)/c$ and

$$l_j(z) = \tilde{a}_1(z)Y_j(z) + \tilde{a}_2(z)\frac{2}{\sqrt{n}}E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j).$$

Here, the first $o_p(1)$ denotes convergence in probability to zero in the C space and in the first step we use the fact that $g(x) = g(c_n) + g'(a)(x-c_n) + o(|x-c_n|)$ as $x \to c_n$. Thus, tightness of $\hat{X}_n(z)$ is from that of $\hat{M}_n(z)$.

Since $b_1(z) \to 1/(1+cm(z))$ and $b_1(z) \to -z\underline{m}(z)$ by (2.17) in [5], we have

$$1/\big(1+cm(z)\big) = -z\underline{m}(z). \tag{5.5}$$

Moreover, we assume for the moment that

$$\sum_{j=1}^{n} E_{j-1}\Big[Y_j(z)\frac{2}{\sqrt{n}}E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)\Big] \xrightarrow{\text{i.p.}} \frac{2cm(z)}{(1+cm(z))^2}. \tag{5.6}$$

It follows from (3.55), (5.2), (5.6) and (5.5) that

$$\begin{aligned}
&\sum_{j=1}^{n} E_{j-1}[l_j(z_1)l_j(z_2)]\\
&\quad = \tilde{a}_1(z_1)\tilde{a}_1(z_2)\frac{2cz_1z_2\underline{m}^2(z_1)\underline{m}^2(z_2)}{(1+\underline{m}(z_1))(1+\underline{m}(z_2))-c\underline{m}(z_1)\underline{m}(z_2)}\\
&\qquad + 2c\tilde{a}_2(z_1)\tilde{a}_2(z_2)+\tilde{a}_1(z_1)\tilde{a}_2(z_2)\frac{2cm(z_1)}{(1+cm(z_1))^2}\\
&\qquad + \tilde{a}_1(z_2)\tilde{a}_2(z_1)\frac{2cm(z_2)}{(1+cm(z_2))^2}+o_p(1)\\
&\quad = a_1^2\times(1.10)+a_2^2\times 2c(g'(c))^2+o_p(1).
\end{aligned}$$

Thus, Lemma 1 follows from the above argument, Lemma 3 and Cramér–Wold's device.

Now consider (5.6). Write

$$\begin{aligned}
&E_{j-1}[E_j(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)]\\
&\quad = E_j(\bar{\mathbf{s}}_j^T)E_j(\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j) = \frac{1}{n}\sum_{i<j}E_j(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\bar{\mathbf{s}}_j\beta_{ij}(z))\\
&\quad = \frac{1}{n^2}\sum_{i<j}E_j(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\mathbf{s}_i\beta_{ij}(z)) + \frac{1}{n}\sum_{i<j}E_j(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\bar{\mathbf{s}}_{ij}\beta_{ij}(z)),
\end{aligned}$$

where we use $\bar{\mathbf{s}}_j = 1/n\sum_{i\neq j}\mathbf{s}_i$ in the second step and $\bar{\mathbf{s}}_j = \bar{\mathbf{s}}_{ij}+\mathbf{s}_i/n$ in the last step. By (3.9), (3.12) and (3.10)

$$E\left|\frac{1}{n}\sum_{i<j}E_j\big(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\bar{\mathbf{s}}_{ij}(\beta_{ij}(z))-b_{12}(z)\big)\right| = O\Big(\frac{1}{\sqrt{n}}\Big),$$

which, together with (3.36), yields

$$E\left|\frac{1}{n}\sum_{i<j}E_j(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\bar{\mathbf{s}}_{ij}\beta_{ij}(z))\right| = o(1).$$

On the other hand, appealing to (3.8), (3.9) and (3.10) ensures that

$$\frac{1}{n^2}\sum_{i<j}E_j(\mathbf{s}_i^T\mathbf{A}_{ij}^{-1}(z)\mathbf{s}_i\beta_{ij}(z)) = \frac{j-1}{n}\frac{n^{-1}E\operatorname{tr}\mathbf{A}^{-1}(z)}{1+n^{-1}E\operatorname{tr}\mathbf{A}^{-1}(z)}+o_{L_1}(1).$$

Therefore, we obtain

$$
\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n} E_{j-1}[E_j(\mathbf{s}_j^T\mathbf{A}_j^{-1}(z)\bar{\mathbf{s}}_j)E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)]\\
&\quad = \frac{n^{-1}E\,\mathrm{tr}\,\mathbf{A}^{-1}(z)}{1+n^{-1}E\,\mathrm{tr}\,\mathbf{A}^{-1}(z)}\frac{1}{n}\sum_{j=1}^{n}\frac{j-1}{n}+o_{L_1}(1) \qquad (5.7)\\
&\quad \xrightarrow{\text{i.p.}} \frac{cm(z)}{2(1+cm(z))}.
\end{aligned}
$$

Next, by the Markov inequality and the Doob inequality

$$
\begin{aligned}
P\left(\max_{i,j}\frac{1}{n}\left|\sum_{k<j}v_{ik}\right|\ge\varepsilon\right) &\le \frac{\sum_{i=1}^{n}E(\max_j(1/n)|\sum_{k<j}v_{1k}|)^4}{\varepsilon^4}\\
&\le \frac{\mathfrak{M}nE((1/n)|\sum_{k<j}v_{ik}|)^4}{\varepsilon^4}\le\frac{\mathfrak{M}}{n},
\end{aligned}
$$

which implies

$$
\max_{i,j}\left|\frac{1}{n}\sum_{k<j}v_{ik}\right|\xrightarrow{\text{i.p.}}0.
$$

This and (3.6) ensure that

$$
\begin{aligned}
&\sum_{j=1}^{n}E_{j-1}[E_j\alpha_j(z)E_j(\bar{\mathbf{s}}_j^T\mathbf{s}_j)]\\
&\quad = \frac{EX_{11}^3}{n}\sum_{j=1}^{n}\sum_{i=1}^{p}[E_j\mathbf{D}_j(z_2)]_{ii}[E_j(\mathbf{e}_i^T\bar{\mathbf{s}}_j)]\\
&\quad \le \max_{i,j}\left|\frac{1}{n}\sum_{k<j}v_{ik}\right|\frac{\mathfrak{M}}{n}\sum_{j=1}^{n}\sum_{i=1}^{p}[E_j(\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(\bar{z}_2))]_{ii} \qquad (5.8)\\
&\quad \le \max_{i,j}\left|\frac{1}{n}\sum_{k<j}v_{ik}\right|\frac{\mathfrak{M}}{n}\sum_{j=1}^{n}E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(\bar{z}_2)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j)\\
&\quad \xrightarrow{\text{i.p.}} 0,
\end{aligned}
$$

because (3.18) implies that $n^{-1}\sum_{j=1}^{n}E_j(\bar{\mathbf{s}}_j^T\mathbf{A}_j^{-1}(\bar{z}_2)\mathbf{A}_j^{-1}(z_2)\bar{\mathbf{s}}_j)$ is uniformly integrable. Based on (5.8) and (5.7) we have (5.6). $\square$

PROOF OF REMARK 4. By (1.3) we get

$$
\frac{\underline{m}(z_1)-\underline{m}(z_2)}{(z_1-z_2)}=\frac{\underline{m}(z_1)\underline{m}(z_2)(1+\underline{m}(z_1))(1+\underline{m}(z_2))}{(1+\underline{m}(z_1))(1+\underline{m}(z_2))-c\underline{m}(z_1)\underline{m}(z_2)}. \qquad (5.9)
$$

Then

$$\begin{aligned}
&\frac{2}{cz_1z_2[(1+\underline{m}(z_1))(1+\underline{m}(z_2))-c\underline{m}(z_1)\underline{m}(z_2)]}\\
&\quad=\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{cz_1z_2(z_1-z_2)\underline{m}(z_1)\underline{m}(z_2)(1+\underline{m}(z_1))(1+\underline{m}(z_2))}\\
&\quad=\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{z_1z_2(z_1-z_2)(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}\\
&\qquad+\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{cz_1z_2(z_1-z_2)(1+\underline{m}(z_1))(1+\underline{m}(z_2))}\\
&\qquad\times\left[\frac{1}{\underline{m}(z_1)\underline{m}(z_2)}-\frac{c}{(1+\underline{m}(z_1))(1+\underline{m}(z_2))}\right]\\
&\quad=\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{z_1z_2(z_1-z_2)(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}\\
&\qquad+\frac{2}{cz_1z_2(1+\underline{m}(z_1))(1+\underline{m}(z_2))}\\
&\quad=\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{z_1z_2(z_1-z_2)(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}\\
&\qquad+\frac{2m(z_1)m(z_2)}{c},
\end{aligned}$$

where in the first step and the third step we use (5.9) and in the last step we use (5.5). On the other hand, via (1.3) one can verify that

$$\frac{2(\underline{m}(z_1)-\underline{m}(z_2))}{z_1z_2(z_1-z_2)(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}=\frac{2(z_2\underline{m}(z_2)-z_1\underline{m}(z_1))^2}{c^2z_1z_2(z_1-z_2)(\underline{m}(z_1)-\underline{m}(z_2))},$$

which is exactly the covariance function in Lemma 2 of [3]. Therefore, Remark 4 holds. □

PROOF OF THEOREM 2. The idea from Lemma 1 to Theorem 2 is similar to that in [5]. First, by the Cauchy formula we have

$$\int f(x)\,dG(x)=-\frac{1}{2\pi i}\oint f(z)m_G(z)\,d(z),$$

where the contour contains the support of $G(x)$ on which $f(x)$ is analytic. Then, with probability one, we have

$$\int f(x)\,dG_n(x)=-\frac{1}{2\pi i}\oint f(z)X_n(z)\,d(z)$$

for all n large, where the complex integral is over $\mathcal{C}$ and

$$G_n(x)=\sqrt{n}\big(F_2^{\mathbf{S}}(x)-F_{c_n}(x)\big).$$

Further,

$$\left|\int f(z)\big(X_n(z)-\hat{X}_n(z)\big)\,dz\right| \le \frac{\mathfrak{M}\rho_n}{\sqrt{n}(u_r-\lambda_{\max}(\boldsymbol{\mathcal{S}}))}+\frac{\mathfrak{M}\rho_n}{\sqrt{n}(\lambda_{\min}(\boldsymbol{\mathcal{S}})-u_l)} \xrightarrow{\text{a.s.}} 0,$$

where, with probability one, $\lambda_{\max}(\boldsymbol{\mathcal{S}}) \to (1+\sqrt{c})^2$ by [11] and $\lambda_{\min}(\boldsymbol{\mathcal{S}}) \to (1-\sqrt{c})^2$ by [23]. Second, note that for any constants a_1 and a_2

$$(\hat{X}_n(z), Y_n) \to a_1 \oint f(z)\hat{X}_n(z)\,dz + a_2 Y_n$$

is a continuous mapping. Therefore, the right-hand side above converges in distribution by Lemma 1. Moreover, Remark 4 shows that (1.6) follows from (1.12) and (1.15) in [3]. □

PROOF OF THEOREM 1. By taking $f(x)=x^{-1}$ and $g(x)=x$ in Theorem 2 and noting that $c_n \to c$ as $n \to \infty$, we can complete the proof. □

APPENDIX

A.1. Some lemmas. We collect some results needed to prove Lemma 1.

LEMMA 2 (Burkholder [8]). *Let $\{Y_i\}$ be a complex martingale difference sequence with respect to the increasing σ-field $\{\mathcal{F}_i\}$. Then for $k \ge 2$*

$$E\left|\sum_i Y_i\right|^k \le \mathfrak{M}_k E\left(\sum_i E(|Y_i|^2|\mathcal{F}_{i-1})\right)^{k/2} + \mathfrak{M}_k E\left(\sum_i |Y_i|^k\right).$$

LEMMA 3 (Theorem 35.12 of Billingsley [7]). *Suppose for each n, $Y_{n,1}, Y_{n,2}, \dots, Y_{n,r_n}$ is a real martingale difference sequence with respect to the increasing σ-field $\{\mathcal{F}_{n,j}\}$ having second moments. If as $n \to \infty$*

$$\text{(i)} \qquad \sum_{j=1}^{r_n} E(Y_{n,j}^2|\mathcal{F}_{n,j-1}) \xrightarrow{i.p.} \sigma^2,$$

$$\text{(ii)} \qquad \sum_{j=1}^{r_n} E\big(Y_{n,j}^2 I_{(|Y_{n,j}|\ge\varepsilon)}\big) \to 0,$$

where σ^2 is a positive constant and ε is an arbitrary positive number, then

$$\sum_{j=1}^{r_n} Y_{n,j} \xrightarrow{D} N(0,\sigma^2).$$

LEMMA 4 ([4], Lemma 2.7). *Let* $\mathbf{Y} = (Y_1, \dots, Y_p)^T$, *where* Y_i*'s are i.i.d. real r.v.'s with mean* 0 *and variance* 1. *Let* $\mathbf{B} = (b_{ij})_{p\times p}$, *a deterministic complex matrix. Then for any* $k \ge 2$, *we have*

$$E|Y^T\mathbf{B}Y - \operatorname{tr}\mathbf{B}|^k \le \mathfrak{M}_k(EY_1^4 \operatorname{tr}\mathbf{B}\mathbf{B}^*)^{k/2} + \mathfrak{M}_k E(Y_1)^{2k}\operatorname{tr}(\mathbf{B}\mathbf{B}^*)^{k/2},$$

where $\mathbf{B}^*$ *denotes the complex conjugate transpose of* $\mathbf{B}$.

LEMMA 5. *Let* $\mathbf{C} = (c_{ij})_{p\times p}$ *be a deterministic complex matrix with* $c_{jj} = 0$ *and* $\mathbf{Y} = (Y_1, \dots, Y_p)^T$, *defined in Lemma* 4. *Then for any* $k \ge 2$,

$$E|Y^T\mathbf{C}Y|^k \le \mathfrak{M}_k(E|Y_1|^k)^2(\operatorname{tr}\mathbf{C}\mathbf{C}^*)^{k/2}. \tag{A.1}$$

Lemma 5 directly follows from the argument of Lemma A.1 in [4].

LEMMA 6. *Under the assumptions of Theorem* 1, *as* $n \to \infty$,

$$\max_i \sqrt{n}|E(\mathbf{e}_i^T\mathbf{A}_1^{-1}(z)\bar{\mathbf{s}}_1)| \to 0. \tag{A.2}$$

PROOF. We first prove that for $i \ne j$, $\sup_{i,j}\sqrt{n}|E(\mathbf{e}_j^T\mathbf{A}_1^{-1}(z)\mathbf{e}_i)| \to 0$. To this end, write

$$\mathbf{A}_1(z) + z\mathbf{I} = \frac{1}{n}\sum_{m=2}^n \mathbf{s}_m\mathbf{s}_m^T.$$

Multiplying by $\mathbf{A}_1^{-1}(z)$ from the right on both sides of the above equality gives

$$\mathbf{I} + z\mathbf{A}_1^{-1}(z) = \frac{1}{n}\sum_{m=2}^n \mathbf{s}_m\mathbf{s}_m^T\mathbf{A}_{m1}^{-1}(z)\beta_{m1}(z).$$

Using

$$\beta_{m1}(z) = b_{12}(z) - \beta_{m1}(z)b_{12}(z)\xi_{m1}(z) \tag{A.3}$$

we obtain

$$\begin{aligned}\mathbf{I} + z\mathbf{A}_1^{-1}(z) &= \frac{b_{12}(z)}{n}\sum_{m=2}^n \mathbf{s}_m\mathbf{s}_m^T\mathbf{A}_{m1}^{-1}(z)\\ &\quad - \frac{b_{12}(z)}{n}\sum_{m=2}^n \mathbf{s}_m\mathbf{s}_m^T\mathbf{A}_{m1}^{-1}(z)\beta_{m1}(z)\xi_{m1}(z).\end{aligned} \tag{A.4}$$

It follows that for $i \ne j$

$$\begin{aligned}&z\sqrt{n}E(\mathbf{e}_j^T\mathbf{A}_1^{-1}(z)\mathbf{e}_i)\\ &\quad = \frac{b_{12}(z)}{\sqrt{n}}\left(\sum_{m=2}^n E(\mathbf{e}_j^T\mathbf{A}_{m1}^{-1}(z)\mathbf{e}_i) - \sum_{m=2}^n E(\mathbf{e}_j^T\mathbf{s}_m\mathbf{s}_m^T\mathbf{A}_{m1}^{-1}(z)\beta_{m1}(z)\xi_{m1}(z)\mathbf{e}_i)\right)\\ &\quad = b_{12}(z)\sqrt{n}\big(E(\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i) - E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\beta_{21}(z)\xi_{21}(z)\mathbf{e}_i)\big).\end{aligned} \tag{A.5}$$

As in (3.10), by Lemma 4 and (3.9),

$$E|\xi_{21}(z)|^k = O(\varepsilon_n^{2k-4}n^{-1}), \qquad k \geq 2. \tag{A.6}$$

Here and in what follows (in this lemma) $O(\varepsilon_n^{2k-4}n^{-1})$ and other bounds are independent of i and j.

We conclude from (3.11) that

$$\begin{aligned}
&b_{12}(z)\sqrt{n}E(\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i)\\
&\quad= b_{12}(z)\sqrt{n}\left[E(\mathbf{e}_j^T\mathbf{A}_1^{-1}(z)\mathbf{e}_i) + E\left(\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\frac{\mathbf{s}_2\mathbf{s}_2^T}{n}\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\beta_{21}(z)\right)\right]\\
&\quad= b_{12}(z)\sqrt{n}E(\mathbf{e}_j^T\mathbf{A}_1^{-1}(z)\mathbf{e}_i) + O(n^{-1/2}).
\end{aligned}$$

For the second term in (A.5), first, by a martingale method similar to (3.21) and (3.11) we have, for $\mathbf{e}_l = \mathbf{e}_i$ or $\mathbf{e}_j$,

$$\begin{aligned}
&E|\mathbf{e}_l^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{e}_j - E(\mathbf{e}_l^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{e}_j)|^2\\
&\quad= E\left|\sum_{m=3}^n(E_m - E_{m-1})[\mathbf{e}_l^T(\mathbf{A}_{21}^{-1}(z_1) - \mathbf{A}_{m21}^{-1}(z_1))\mathbf{e}_j]\right|^2\\
&\quad\leq \frac{M}{n^2}\sum_{m=3}^n E|\mathbf{s}_m^T\mathbf{A}_{m21}^{-1}(z_1)\mathbf{e}_j\mathbf{e}_l^T\mathbf{A}_{m21}^{-1}(z_1)\mathbf{s}_m|^2 = O(n^{-1}).
\end{aligned}\tag{A.7}$$

This and (3.9) ensure that

$$\begin{aligned}
&\left|\frac{1}{n}E[\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i(\operatorname{tr}\mathbf{A}_{21}^{-1}(z) - E\operatorname{tr}\mathbf{A}_{21}^{-1}(z))]\right|\\
&\quad= \left|\frac{1}{n}E[(\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i - E\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i)(\operatorname{tr}\mathbf{A}_{21}^{-1}(z) - E\operatorname{tr}\mathbf{A}_{21}^{-1}(z))]\right|\\
&\quad\leq \frac{\mathfrak{M}}{n}(E|\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i - E\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i|^2 E|\operatorname{tr}\mathbf{A}_{21}^{-1}(z) - E\operatorname{tr}\mathbf{A}_{21}^{-1}(z)|^2)^{1/2}\\
&\quad\leq \frac{\mathfrak{M}}{n}.
\end{aligned}$$

Second, appealing to (3.5) gives

$$\begin{aligned}
&E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\gamma_{21}(z))\\
&\quad= E((\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\mathbf{e}_j^T\mathbf{s}_2 - \mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i)\gamma_{21}(z))\\
&\quad= \frac{EX_{11}^4 - 3}{n}E(\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_j) + \frac{2}{n}E(\mathbf{e}_j^T\mathbf{A}_{21}^{-2}(z)\mathbf{e}_i).
\end{aligned}$$

It follows that

$$\begin{aligned}
&\sqrt{n}E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\xi_{21}(z))\\
&\quad=\sqrt{n}E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\gamma_{21}(z))\\
&\qquad+\sqrt{n}E\biggl[\mathbf{e}_j^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\frac{1}{n}\bigl(\operatorname{tr}\mathbf{A}_{21}^{-1}(z)-E\operatorname{tr}\mathbf{A}_{21}^{-1}(z)\bigr)\biggr]\\
&\quad=O(n^{-1/2}).
\end{aligned}$$

On the other hand, in view of (3.11) and (A.6) we obtain

$$\sqrt{n}E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_2^{-1}(z)\mathbf{e}_i\beta_{21}(z)\xi_{21}^2(z))=O(\varepsilon_n).$$

Therefore, by (A.3) we find

$$\begin{aligned}
&\sqrt{n}E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\beta_{21}(z)\xi_{21}(z)\mathbf{e}_i)\\
&\quad=\sqrt{n}b_{12}(z)[E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_{21}^{-1}(z)\mathbf{e}_i\xi_{21}(z))-E(\mathbf{e}_j^T\mathbf{s}_2\mathbf{s}_2^T\mathbf{A}_2^{-1}(z)\mathbf{e}_i\beta_{21}(z)\xi_{21}^2(z))]\\
&\quad=O(\varepsilon_n).
\end{aligned}$$

Therefore, combining the above argument with (3.38), we have

$$\sup_{i\neq j}\bigl|\sqrt{n}E(\mathbf{e}_j^T\mathbf{A}_1^{-1}(z)\mathbf{e}_i)\bigr|\to 0. \tag{A.8}$$

Next, applying (A.3) two times gives

$$\begin{aligned}
&E(\mathbf{e}_i^T\mathbf{A}_1^{-1}(z_1)\bar{\mathbf{s}}_1)\\
&\quad=\frac{1}{n}\sum_{m=2}^{n}E(\mathbf{e}_i^T\mathbf{A}_{m1}^{-1}(z_1)\mathbf{s}_m\beta_{m1}(z_1))\\
&\quad=\frac{b_{12}^2(z_1)(n-1)}{n}[-E(\mathbf{e}_i^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{s}_2\xi_{21}(z_1))\\
&\qquad\qquad+E(\mathbf{e}_i^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{s}_2\beta_{21}(z_1)\xi_{21}^2(z_1))].
\end{aligned}$$

Obviously, we conclude from (A.6), (3.11) and Hölder's inequality that

$$\left|\frac{n-1}{n}E(\mathbf{e}_i^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{s}_2\beta_{21}(z_1)\xi_{21}^2(z_1))\right|=O(n^{-1/2}\varepsilon_n),$$

while (3.6), (3.8) and (A.8) yield

$$\begin{aligned}
&\max_i\left|\frac{n-1}{n}E(\mathbf{e}_i^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{s}_2\xi_{21}(z_1))\right|\\
&\quad=\max_i\left|\frac{EX_{11}^3(n-1)}{n^2}\sum_{j=1}^{p}E[\mathbf{e}_i^T\mathbf{A}_{21}^{-1}(z_1)\mathbf{e}_j(\mathbf{A}_{21}^{-1}(z_1))_{jj}]\right|
\end{aligned}$$

$$\leq \frac{|EX_{11}^3|}{n} \max_i \sum_{j\neq i}^{p} |E[\mathbf{e}_i^T \mathbf{A}_1^{-1}(z_1)\mathbf{e}_j (\mathbf{A}_1^{-1}(z_1))_{jj}]| + \frac{\mathfrak{M}}{n}$$

$$\leq \mathfrak{M}\left|EX_{11}^3 E\frac{1}{n}\operatorname{tr}\mathbf{A}_1^{-1}(z_1)\right| \max_{i\neq j} |E(\mathbf{e}_i^T \mathbf{A}_1^{-1}(z_1)\mathbf{e}_j)| + \frac{\mathfrak{M}}{n}$$

$$= o(n^{-1/2}).$$

Here we also use the estimate, via (A.7),

$$E\big|\big(\mathbf{e}_i^T \mathbf{A}_1^{-1}(z_1)\mathbf{e}_j - E(\mathbf{e}_i^T \mathbf{A}_1^{-1}(z_1)\mathbf{e}_j)\big)\big((\mathbf{A}^{-1}(z_1))_{jj} - E(\mathbf{A}^{-1}(z_1))_{jj}\big)\big| = O(n^{-1}).$$

Thus, the proof of (A.2) is complete. □

A.2. Truncation of the underlying random variables. To guarantee the results holding under the fourth moment, it is necessary to truncate and centralize the underlying r.v.'s at an appropriate rate. As in [5], (1.8), one may select a positive sequence ε_n so that

$$\varepsilon_n \to 0 \quad \text{and} \quad \varepsilon_n^{-4} EX_{11}^4 I\big(|X_{11}| \geq \varepsilon_n\sqrt{n}\big) \to 0. \tag{A.9}$$

Set $\widehat{X}_{ij} = X_{ij} I(|X_{ij}| \leq \varepsilon_n\sqrt{n}) - EX_{ij} I(|X_{ij}| \leq \varepsilon_n\sqrt{n})$ and $\widetilde{\mathbf{X}}_n = \mathbf{X}_n - \widehat{\mathbf{X}}_n = (\tilde{X}_{ij})$ with $\widehat{\mathbf{X}}_n = (\hat{X}_{ij})$. Let $\sigma_n = \sqrt{E|\hat{X}_{11}|^2}$, $\check{\mathbf{S}}_n = (n\sigma_n^2)^{-1}\widehat{\mathbf{X}}_n\widehat{\mathbf{X}}_n^T$ and $\check{\mathbf{A}}^{-1}(z) = (\check{\mathbf{S}}_n - zI)^{-1}$. Moreover, introduce $\bar{\check{\mathbf{s}}} = \frac{1}{n}\sum_{j=1}^n \check{\mathbf{s}}_j$, where $\check{\mathbf{s}}_j$ is the jth column of the matrix $(\sigma_n)^{-1}\widehat{\mathbf{X}}_n$.

LEMMA 7. *Assume that $X_{ij}, i = 1, \ldots, p, j = 1, \ldots, n$ are i.i.d. with $EX_{11} = 0$, $E|X_{11}|^2 = 1$ and $E|X_{11}|^4 < \infty$, for $z \in \mathcal{C}_n^+$, we have then*

$$\sqrt{n}\big(\bar{\mathbf{s}}^T \mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \bar{\check{\mathbf{s}}}^T \check{\mathbf{A}}^{-1}(z)\bar{\check{\mathbf{s}}}\big) \xrightarrow{i.p.} 0, \tag{A.10}$$

where the convergence in probability holds uniformly for $z \in \mathcal{C}_n^+$. Moreover,

$$\sqrt{n}(\bar{\mathbf{s}}^T \bar{\mathbf{s}} - \bar{\check{\mathbf{s}}}^T \bar{\check{\mathbf{s}}}) \xrightarrow{i.p.} 0. \tag{A.11}$$

PROOF. Write

$$\sqrt{n}\big(\bar{\mathbf{s}}^T \mathbf{A}^{-1}(z)\bar{\mathbf{s}} - \bar{\check{\mathbf{s}}}^T \check{\mathbf{A}}^{-1}(z)\bar{\check{\mathbf{s}}}\big) = u_{n1} + u_{n2} + u_{n3},$$

where

$$u_{n1} = \sqrt{n}[(\bar{\mathbf{s}} - \bar{\check{\mathbf{s}}})^T \mathbf{A}^{-1}(z)\bar{\mathbf{s}}], u_{n2} = \sqrt{n}[\bar{\check{\mathbf{s}}}^T (\mathbf{A}^{-1}(z) - \check{\mathbf{A}}^{-1}(z))\bar{\mathbf{s}}]$$

and

$$u_{n3} = \sqrt{n}[\bar{\check{\mathbf{s}}}^T \check{\mathbf{A}}^{-1}(z)(\bar{\mathbf{s}} - \bar{\check{\mathbf{s}}})].$$

Consider u_{n1} on the $\mathcal{C}_u$ first. It is observed that

$$\begin{aligned}|u_{n1}| &\le \sqrt{n}\|(\bar{\mathbf{s}}-\bar{\tilde{\mathbf{s}}})^T\|\|\mathbf{A}^{-1}(z)\|\|\bar{\mathbf{s}}\| \le \frac{\sqrt{n}}{v_0}\|(\bar{\mathbf{s}}-\bar{\tilde{\mathbf{s}}})^T\|\|\bar{\mathbf{s}}\| \\ &\le \frac{\sqrt{n}}{v_0}\left|1-\frac{1}{\sigma_n}\right|\|\bar{\mathbf{s}}\|^2+\frac{\sqrt{n}}{v_0}\frac{1}{\sigma_n}\|\bar{\tilde{\mathbf{s}}}\|\|\bar{\mathbf{s}}\|,\end{aligned} \tag{A.12}$$

since $\bar{\mathbf{s}}-\bar{\tilde{\mathbf{s}}}=(1-\frac{1}{\sigma_n})\bar{\mathbf{s}}+\frac{1}{\sigma_n}\bar{\tilde{\mathbf{s}}}$ with $\bar{\tilde{\mathbf{s}}}=\sum_{j=1}^n \tilde{\mathbf{s}}_j/n$ and $\tilde{\mathbf{s}}_j$ being the jth column of $\widetilde{\mathbf{X}}_n$. Moreover, it follows from (A.9) that

$$1-\sigma_n^2 \le 2EX_{11}^2 I\big(|X_{11}|\ge \varepsilon_n\sqrt{n}\big) \le 2\varepsilon_n^{-2}n^{-1}EX_{11}^4 I\big(|X_{11}|\ge\varepsilon_n\sqrt{n}\big)=o(\varepsilon_n^2 n^{-1}),$$

which implies that

$$\sqrt{n}(1-1/\sigma_n)=\sqrt{n}(\sigma_n^2-1)/[\sigma_n(1+\sigma_n)]=o(n^{-1/2}). \tag{A.13}$$

On the other hand,

$$E\|\bar{\tilde{\mathbf{s}}}\|^2=E\left[\sum_{i=1}^p\left|\frac{1}{n}\sum_{j=1}^n \tilde{X}_{ij}\right|^2\right]=\frac{1}{n^2}\sum_{i=1}^p\sum_{j=1}^n E\tilde{X}_{ij}^2 \le \frac{\mathfrak{M}}{n\varepsilon_n^2}EX_{11}^4 I\big(|X_{11}|\ge\varepsilon_n\sqrt{n}\big),$$

which, via (A.9), gives that

$$\sqrt{n}\|\bar{\tilde{\mathbf{s}}}\| \xrightarrow{\text{i.p.}} 0. \tag{A.14}$$

In addition, $\|\bar{\mathbf{s}}\|^2$ is uniformly integrable because (3.17) remains true for $k=2$ without truncation by a careful check on its argument. This, together with (A.12)–(A.14), ensures that u_{n1} converges in probability to zero uniformly on $\mathcal{C}_u$.

Analyze u_{n2} next. Since $\mathbf{X}_n-\sigma_n^{-1}\widehat{\mathbf{X}}_n=(1-\sigma_n^{-1})\mathbf{X}_n+\sigma_n^{-1}\widetilde{\mathbf{X}}_n$, we have

$$\begin{aligned}|u_{n2}| &\le \sqrt{n}\|\bar{\check{\mathbf{s}}}^T\|\|\mathbf{A}^{-1}(z)-\check{\mathbf{A}}^{-1}(z)\|\|\bar{\mathbf{s}}\| \le \frac{\sqrt{n}}{v_0^2}\|\bar{\check{\mathbf{s}}}^T\|\|\mathbf{A}(z)-\check{\mathbf{A}}(z)\|\|\bar{\mathbf{s}}\| \\ &\le \frac{1}{v_0^2\sqrt{n}}\|\bar{\check{\mathbf{s}}}^T\|\|\bar{\mathbf{s}}\|[\|\mathbf{X}_n-\sigma_n^{-1}\widehat{\mathbf{X}}_n\|\|\mathbf{X}_n^T\|+\|\sigma_n^{-1}\widehat{\mathbf{X}}_n\|\|\mathbf{X}_n^T-\sigma_n^{-1}\widehat{\mathbf{X}}_n^T\|] \\ &\le \frac{1}{v_0^2\sqrt{n}}\|\bar{\check{\mathbf{s}}}^T\|\|\bar{\mathbf{s}}\|[(1-\sigma_n^{-1})\|\mathbf{X}_n\|\|\mathbf{X}_n^T\|+\sigma_n^{-1}\|\widetilde{\mathbf{X}}_n\|\|\mathbf{X}_n^T\| \\ &\qquad +\|\sigma_n^{-1}\widehat{\mathbf{X}}_n\|(1-\sigma_n^{-1})\|\mathbf{X}_n^T\|+\|\sigma_n^{-1}\widehat{\mathbf{X}}_n\|\sigma_n^{-1}\|\widetilde{\mathbf{X}}_n^T\|].\end{aligned}$$

As before, $\|\bar{\check{\mathbf{s}}}\|$ and $\|\bar{\mathbf{s}}\|$ are uniformly integrable. Moreover, the spectral norms $\|\mathbf{X}_n^T\|/\sqrt{n}$ and $\|\sigma_n^{-1}\widehat{\mathbf{X}}_n\|/\sqrt{n}$ both converge to $(1+\sqrt{c})^2$ with probability one by [25]. In addition, $\|\widetilde{\mathbf{X}}_n^T\|/\sqrt{nE\tilde{X}_{11}^2}$ converges to $(1+\sqrt{c})^2$ with probability one. From (A.9) we have

$$nE\tilde{X}_{11}^2 \le 2\varepsilon_n^{-2}EX_{11}^4 I\big(|X_{11}|\ge\varepsilon_n\sqrt{n}\big)=O(\varepsilon_n^2),$$

which, together with (A.13), yields that u_{n2} converges in probability to zero uniformly on $\mathcal{C}_u$.

Clearly, the argument for u_{n1} works for u_{n3} as well. Moreover, note that $\|\mathbf{A}^{-1}(z)\|$ is bounded for $z \in \mathcal{C}_l, u_l < 0$. As for $z \in \mathcal{C}_l, u_l > 0$ or $z \in \mathcal{C}_r$, by [25] we have

$$\lim_{n\to\infty} \min(u_r - \lambda_{\max}(\mathbf{A}), \lambda_{\min}(\mathbf{A}) - u_l) > 0, \qquad \text{a.s.}$$

and

$$\lim_{n\to\infty} \min(u_r - \lambda_{\max}(\check{\mathbf{A}}), \lambda_{\min}(\check{\mathbf{A}}) - u_l) > 0, \qquad \text{a.s.}$$

Therefore, the above argument for u_{nj}, $j = 1, 2, 3$ for $z \in \mathcal{C}_u$ of course applies to the cases (1) $z \in \mathcal{C}_l, u_l < 0$; (2) $z \in \mathcal{C}_l, u_l > 0$; (3) $z \in \mathcal{C}_r$. Thus, (A.10) holds.

Finally, the above argument for (A.10) certainly works for (A.11). Thus, the proof is complete. □

Acknowledgments. The authors would like to thank the editor, an associate editor and a referee for their constructive comments which have helped to improve the paper a great deal.

REFERENCES

[1] ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York. MR0771294

[2] BAI, Z. and SARANADASA, H. (1996). Effect of high dimension: By an example of a two sample problem. *Statist. Sinica* **6** 311–329. MR1399305

[3] BAI, Z. D., MIAO, B. Q. and PAN, G. M. (2007). On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.* **35** 1532–1572. MR2330979

[4] 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

[5] 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

[6] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. MR0233396

[7] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York. MR1324786

[8] BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692

[9] DIACONIS, P. and EVANS, S. N. (2001). Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** 2615–2633. MR1828463

[10] HOTELLING, H. (1931). The generalization of Student's ratio. *Ann. Math. Statist.* **2** 360–378.

[11] JIANG, T. (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā* **66** 35–48. MR2082906

[12] JOHANSSON, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** 151–204. MR1487983

[13] JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961

[14] JONSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38. MR0650926

[15] LEHMANN, E. L. and ROMANO, J. P. (2005). *Testing Statistical Hypotheses*, 3rd ed. Springer, New York. MR2135927

[16] MARČENKO, V. A. and PASTUR, L. A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sb.* **1** 457–483.

[17] PAN, G. M. and ZHOU, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* **18** 1232–1270. MR2418244

[18] SILVERSTEIN, J. W. (1989). On the eigenvectors of large-dimensional sample covariance matrices. *J. Multivariate Anal.* **30** 1–16. MR1003705

[19] SILVERSTEIN, J. W. (1990). Weak convergence of random functions defined by the eigenvectors of sample covariance matrices. *Ann. Probab.* **18** 1174–1194. MR1062064

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

[21] SINAI, Y. and SOSHNIKOV, A. (1998). Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat.* (*N.S.*) **29** 1–24. MR1620151

[22] WACHTER, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18. MR0467894

[23] XIAO, H. and ZHOU, W. (2010). Almost sure limit of the smallest eigenvalue of some sample correlation matrices. *J. Theoret. Probab.* **23** 1–20. MR2591901

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

[25] YIN, Y. Q., BAI, Z. D. and KRISHNAIAH, P. R. (1988). On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** 509–521. MR0950344

DIVISION OF MATHEMATICAL SCIENCES
SCHOOL OF PHYSICAL
AND MATHEMATICAL SCIENCES
NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE 637371
E-MAIL: gmpan@ntu.edu.sg

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 117546
E-MAIL: stazw@nus.edu.sg