The Annals of Applied Probability
2022, Vol. 32, No. 6, 4763–4802
https://doi.org/10.1214/22-AAP1802
© Institute of Mathematical Statistics, 2022

PROPERTIES OF EIGENVALUES AND EIGENVECTORS OF LARGE-DIMENSIONAL SAMPLE CORRELATION MATRICES

BY YANQING YIN[1,a] AND YANYUAN MA[2,b]

[1]*School of Mathematics and Statistics, Chongqing University,* [a]*yinyq799@nenu.edu.cn*
[2]*Department of Statistics, Pennsylvania State University,* [b]*yzm63@psu.edu*

This paper is to study the properties of eigenvalues and eigenvectors of high-dimensional sample correlation matrices. We first improve the result of Jiang (*Sankhyā* **66** (2004) 35–48), Xiao and Zhou (*J. Theoret. Probab.* **23** (2010) 1–20) and the Theorem 1 of El Karoui (*Ann. Appl. Probab.* **19** (2009) 2362–2405), both concerning the limiting spectral distribution and the extreme eigenvalues of sample correlation matrices, by allowing a more general fourth moment condition. Then, we establish a central limit theorem (CLT) for the linear statistics of the eigenvectors of large sample correlation matrices. We discover that the difference between the functional CLT of the sample covariance matrix and the sample correlation matrix is fundamentally influenced by the direction of a nonrandom projection vector. In the special case where the square root of the correlation matrix is identity, the difference will be determined by the sum of the fourth powers of the entries of the projection vector. These results also indicate that the eigenmatrix of sample correlation matrices is *not* asymptotically Haar if the underlying distribution is Gaussian. In other words, the normalization based on the sample variances affects the asymptotic properties of the eigenmatrix of the Wishart matrix. Furthermore, we establish a theorem concerning CLT for the linear statistics of the eigenvectors of large sample covariance matrices. This theorem improves the main results in Bai, Miao and Pan (*Ann. Probab.* **35** (2007) 1532–1572), which requires the assumption that the fourth moment of the underlying variable matches the one of Gaussian distribution, as well as Theorem 1.3 in Pan and Zhou (*Ann. Appl. Probab.* **18** (2008) 1232–1270), which relaxed the Gaussian like fourth moment requirement but assumes the maximum entries of the projection vector converge to 0 uniformly. We illustrate the usefulness of the theoretical results through an application in communications.

1. Introduction. Sample correlation matrix is a central object encountered in multivariate statistical analysis. Closely related to sample covariance matrix, sample correlation matrix is also frequently studied. Let $\mathbf{y} := (\mathbf{y}_1, \dots, \mathbf{y}_n)$ be a random sample of size n from a p-dimensional population with covariance matrix $\mathbf{\Sigma}_n$. We assume the following model for $\mathbf{y}$:

$$\mathbf{y}_j = \mathbf{\Gamma}_n \mathbf{x}_j, \quad j = 1, \dots, n,$$

where $\mathbf{\Gamma}_n$ satisfies that $\mathbf{\Gamma}_n \mathbf{\Gamma}_n^* = \mathbf{\Sigma}_n$, and $\mathbf{x}_j = (x_{1j}, \dots, x_{pj})^*$ consists of independently and identically distributed (i.i.d.) random variables with mean 0 and unit variance. Throughout this paper, we consider the case where there is no diagonal entry of $\mathbf{\Sigma}_n$ equals to 0. Let $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$. Then the sample covariance matrix is defined as

$$\widehat{\mathbf{S}}_n = \frac{1}{n} \sum_{j=1}^{n} \mathbf{y}_j \mathbf{y}_j^* = \frac{1}{n} \mathbf{Y}_n \mathbf{Y}_n^*.$$

Received August 2020; revised October 2021.
MSC2020 subject classifications. Primary 62H10; secondary 60F05.
Key words and phrases. Eigenvalue distribution, Haar distribution, high dimension correlation matrix, eigenvector, random matrix theory.

The population correlation matrix $\mathbf{R}_n$ and the sample correlation matrix $\widehat{\mathbf{R}}_n$ are then given respectively by

$$\mathbf{R}_n = [\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1/2}\boldsymbol{\Sigma}_n[\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1/2}, \qquad \widehat{\mathbf{R}}_n = [\mathrm{Diag}(\widehat{\mathbf{S}}_n)]^{-1/2}\widehat{\mathbf{S}}_n[\mathrm{Diag}(\widehat{\mathbf{S}}_n)]^{-1/2},$$

where for any square matrix $\mathbf{A}$, $\mathrm{Diag}(\mathbf{A})$ stands for the diagonal matrix formed by the diagonal entries of $\mathbf{A}$.

Denote

$$\mathbf{G}_n = (g_{kh})_{k,h=1}^{p} = [\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1/2}\boldsymbol{\Gamma}_n, \qquad \Xi_j = \mathbf{G}_n\mathbf{x}_j\mathbf{x}_j^*\mathbf{G}_n^*,$$

$$e_k'\mathbf{G}_n := \mathbf{g}_k^* \quad \text{and} \quad \boldsymbol{\Xi}_n = \frac{1}{n}\sum_{j=1}^{n}\Xi_j = \frac{1}{n}\mathbf{G}_n\mathbf{X}_n\mathbf{X}_n^*\mathbf{G}_n^*, \tag{1}$$

where e_k is the kth column of the identity matrix. It is easy to see that

$$\mathrm{E}(\Xi_j) = \mathrm{E}(\boldsymbol{\Xi}_n) = \mathbf{R}_n \quad \text{hence } \mathrm{Diag}\big(\mathrm{E}(\Xi_j)\big) = \mathrm{Diag}\big(\mathrm{E}(\boldsymbol{\Xi}_n)\big) = \mathbf{I}_p. \tag{2}$$

Also, by the fact that

$$\mathrm{Diag}(\boldsymbol{\Xi}_n) = [\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1/2}\,\mathrm{Diag}(\widehat{\mathbf{S}}_n)[\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1/2} = [\mathrm{Diag}(\boldsymbol{\Sigma}_n)]^{-1}\,\mathrm{Diag}(\widehat{\mathbf{S}}_n),$$

one can verify that $\widehat{\mathbf{R}}_n$ can be written as

$$\widehat{\mathbf{R}}_n = [\mathrm{Diag}(\boldsymbol{\Xi}_n)]^{-1/2}\boldsymbol{\Xi}_n[\mathrm{Diag}(\boldsymbol{\Xi}_n)]^{-1/2}.$$

The rapid development of computation power allows us to save and analyze very big data sets with very large dimension. Thus, the classical statistical theory, which assumes that the data dimension p is fixed while sample size n diverges, faces challenges when dealing with high-dimensional data set. This has prompted many new development in theories and application tools to handle high-dimensional data. Among these, high-dimensional statistical inference involving sample correlation matrix has gained much popularity recently. See for instance, Cai and Jiang (2011), Hero and Rajaratnam (2011), Zheng et al. (2019), Fan, Guo and Zheng (2019) for important developments in this area.

A very important tool in analyzing large dimension sample covariance and correlation matrices is the random matrix theory (RMT). Initiated from the pioneering investigation of the energy level distribution of a large number of particles in Quantum Mechanics, the original interests of RMT mainly focus on eigenvalue distributions of large-dimensional random matrices. Let $\mathbf{A}_n$ be a $p \times p$ symmetric matrix with eigenvalues $\lambda_1 \le \cdots \le \lambda_p$ and let $\mathbf{A}_n = \mathbf{U}_n\Lambda_n\mathbf{U}_n^*$ be its spectral decomposition. For any unit vector $\boldsymbol{\pi}_n \in \mathbb{C}^p$, let $\mathbf{q}_n = \mathbf{U}_n\pi_n := (q_1, \ldots, q_p)'$. We make the following definitions.

DEFINITION 1.1. The *empirical spectral distribution* (ESD) of $\mathbf{A}_n$ is defined as

$$F^{\mathbf{A}_n}(x) = \frac{1}{p}\sum_{j=1}^{p} I(\lambda_j \le x),$$

where $I(\cdot)$ is the indicator function.

If, as p, n tend to infinity, the limit of $F^{\mathbf{A}_n}(x)$ exists, the limit distribution is called the *limit spectral distribution* (LSD).

DEFINITION 1.2. Given π_n, the *vector empirical spectral distribution* (VESD) function based on eigenvalues and eigenvectors of matrix $\mathbf{A}_n$ is defined as

$$F_{v,\pi_n}^{A_n}(x) = \sum_{j=1}^{p} |q_j|^2 I(\lambda_j \le x).$$

When the underlying data is generated from a multivariate Gaussian distribution, the sample covariance matrix $\widehat{\mathbf{S}}_n$ is a Wishart matrix and its eigenmatrix will follow the Haar distribution, see Anderson (2003) or Corollary 2.2 of Dumitriu and Edelman (2002). In other words, denote $\widehat{\mathbf{S}}_n = \mathbf{U}_n \Lambda_n \mathbf{U}_n^*$ as the spectral decomposition of $\widehat{\mathbf{S}}_n$, then $\mathbf{U}_n$ should follow the uniform distribution over the group formed by all unitary matrices.

Assume that $\mathbf{U}_n$ follows a Haar distribution, then for any unit vector $\boldsymbol{\pi}_n \in \mathbb{C}^p$, $\mathbf{q}_n = \mathbf{U}_n \pi_n := (q_1, \dots, q_p)'$ will follow a uniform distribution over the unit sphere. In other words, the vector $\mathbf{q}_n$ has the same distribution as $\mathbf{z}/|\mathbf{z}|$ where $\mathbf{z} = (z_1, \dots, z_p)' \sim N_p(\mathbf{0}, \mathbf{I}_p)$. Here and throughout the paper, $|\cdot|$ denote the Euclidean norm of a vector. And the stochastic process

$$\mathbb{Q}_p(t) = \sqrt{\frac{p}{2}} \sum_{j=1}^{[pt]} \left(|q_j|^2 - \frac{1}{n}\right) \stackrel{d}{=} \sqrt{\frac{p}{2}} \frac{1}{|\mathbf{z}|^2} \sum_{j=1}^{[pt]} \left(|z_j|^2 - \frac{|\mathbf{z}|^2}{p}\right)$$

convergence to a Brownian Bridge $\mathbb{B}(t)$ as $p \to \infty$ (see page 334 in Bai and Silverstein (2010)).

For any $\mathbf{A}_n$, we shall make a time transform

$$\mathcal{Q}_p^{\mathbf{A}_n}(x) = \mathbb{Q}_p\big(F^{\mathbf{A}_n}(x)\big).$$

Then $\mathcal{Q}_p^{\widehat{\mathbf{S}}_n}(x)$ will approximate $\mathbb{B}(F^{\rho,H}(x))$, where $F^{\rho,H}(x)$ is the LSD of $\widehat{\mathbf{S}}_n$. Here, ρ is the limit of the dimension to sample size ratio p/n and H is the LSD of Σ_n.

Recall the definitions of ESD and VESD, it follows that

$$\mathcal{Q}_p^{\mathbf{A}_n}(x) = \sqrt{\frac{p}{2}} \big(F_{v,\pi_n}^{A_n}(x) - F^{A_n}(x)\big).$$

We thus convert the problem of studying $\mathbb{Q}_p(t)$, which often arises in studies of convergence into a Brownian bridge and related central limit theorems hence is of research interest, to the problem of investigating the difference between the ESD and VESD.

Throughout this paper, we denote $\|\cdot\|$ as the spectral norm of a matrix. Let C stand for a constant that may take different values depending on the context. For any real sequences a_n and b_n, we use $a_n = o(b_n)$ to denote the relationship $a_n/b_n \to 0$ as $b_n \to \infty$ and $a_n = O(b_n)$ to denote the relationship $a_n/b_n \le C$ as $b_n \to \infty$. Let $\mathbf{I}_p$ stand for the p-dimensional identity matrix and we will leave out the subscript if there is no ambiguity. Let $z = u + iv \in \mathbb{C}^+$ be a complex number. We also use $\mathbb{D}_{\mathbf{A}}$ to stand for $\mathrm{Diag}(\mathbf{A})$ for any square matrix $\mathbf{A}$.

The rest of the paper is organized as the follows. In Section 2, we summarize some previous results in random matrix theory and highlight our new contributions. Section 3 contains the main results. The proofs of the theorems are provided in Section 5.

2. Existing results and new contributions.

We first summarize some known results in random matrix theory concerning the distributional properties of the eigenvalues and eigenvectors of a sample covariance matrix, as well as those of a sample correlation matrix. We then summarize our new contributions.

2.1. *Background and existing results.*

We first introduce a main tool in RMT, the Stieltjes transform defined as follows.

DEFINITION 2.1. For any function G of bounded variation on the real line, its Stieltjes transform is defined by

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

Similar to the characteristic function, Stieltjes transform is a useful tool for investigating distribution functions and distribution function sequences. Below are some lemmas that establish the mathematical foundations, which can be found in Akhiezer and Glazman (1993). One can also see Theorems B.8–B.10 in Bai and Silverstein (2010).

LEMMA 2.1. *For any continuity points $a < b$ of G, we have*

$$G\{[a,b]\} = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_a^b \Im\big(s_G(x+i\varepsilon)\big)\, dx.$$

LEMMA 2.2. *Assume that $\{G_n\}$ is a sequence of functions of bounded variation and $G_n(-\infty) = 0$ for all n. Then,*

$$\lim_{n\to\infty} s_{G_n}(z) = s_G(z) \quad \forall z \in \mathbb{C}^+$$

if and only if there is a function of bounded variation G with $G(-\infty) = 0$ and Stieltjes transform $s_G(z)$ and such that $G_n \to G$ vaguely.

LEMMA 2.3. *Let G be a function of bounded variation and $x_0 \in \mathbb{R}$. Assume that $\lim_{z\in\mathbb{C}^+\to x_0} \Im(s_G(z))$ exists. Let the limit be $\Im(s_G(x_0))$. Then G is differentiable at x_0, and its derivative is $\pi^{-1}\Im(s_G(x_0))$.*

Lemma 2.1 is referred to as the inversion formula. It shows a one-to-one correspondence between G and its Stieltjes transform $s_G(z)$ when G is a finite signed measure. Lemma 2.2 and 2.3 show that proving the convergence properties of ESDs of a sequence of random matrices can be converted to proving the convergence of the corresponding Stieltjes transforms. In addition, the LSD can be easily established by its limit Stieltjes transform. Then one can obtain the density function of a signed measure via its Stieltjes transform.

The celebrated M–P law (Marčenko and Pastur (1967)), also in Wachter (1978) and Silverstein (1995), states that if $p/n \to \rho \in (0,\infty)$, $F^{\mathbf{\Sigma}_n} \xrightarrow{d} H$ and the sequence $(\mathbf{\Sigma}_n)_p$ is bounded in spectral norm, then almost surely, the ESD $F^{\widehat{\mathbf{S}}_n}$ of the sample covariance matrix $\widehat{\mathbf{S}}_n$ will tend weakly to a nonrandom p.d.f. $F^{\rho,H}$ as $n \to \infty$. And for each $z \in \mathbb{C}^+$, $s(z) = s_F(z) := s_{F^{\rho,H}}(z)$ is the unique solution to the equation

$$s(z) = \int \frac{1}{t(1-\rho-\rho z s(z)) - z}\, dH(t). \tag{3}$$

Let $\underline{\mathbf{S}}_n = \frac{1}{n}\mathbf{Y}_n^*\mathbf{Y}_n$. Note that the spectra of $\widehat{\mathbf{S}}_n$ and $\underline{\mathbf{S}}_n$ only differ by $|p-n|$ zero eigenvalues. It follows that

$$F^{\underline{\mathbf{S}}_n}(x) = (1-\rho_n)I_{[0,\infty)} + \rho_n F^{\widehat{\mathbf{S}}_n}(x),$$

from which we get

$$\begin{aligned} &\underline{F}^{\rho,H}(x) = (1-\rho)I_{[0,\infty)} + \rho F^{\rho,H}(x), \\ &s_{F^{\underline{\mathbf{S}}_n}}(z) = -\frac{1-\rho_n}{z} + \rho_n s_{F^{\widehat{\mathbf{S}}_n}}(z), \quad z \in \mathbb{C}^+, \end{aligned} \tag{4}$$

and as $n \to \infty$

$$\underline{s}(z) := \underline{s}_{F^{\rho,H}}(z) = -\frac{1-\rho}{z} + \rho s(z), \quad z \in \mathbb{C}^+. \tag{5}$$

Hence, (3) can be re-expressed as

$$\underline{s}(z) = -\left(z - \rho \int \frac{t}{1+t\underline{s}(z)}\, dH(t)\right)^{-1}.$$

Furthermore, it leads to

$$z = -\frac{1}{\underline{s}(z)} + \rho \int \frac{t}{1 + t\underline{s}(z)}\, dH(t).$$

The almost sure convergence of the extreme eigenvalues of a sample covariance matrix to the edge of the spectrum is considered in Geman (1980), Yin, Bai and Krishnaiah (1988), Bai and Yin (1993), Tikhomirov (2015). Then in Bai and Silverstein (1998), Bai and Silverstein (1999), the no eigenvalue outside and exact separation results are obtained. As for the second order limit theorem for the eigenvalues of a sample covariance matrix, the first work was done by Jonsson (1982), where the author studied the fluctuation for the sums of the powers of the sample eigenvalues of a Wishart matrix. Then Bai and Silverstein (2004) prove a general CLT for the linear spectral statistics (LSS) under the condition that the fourth moment of the underlying distribution match that of a Gaussian distribution and that the test functions are analytic. These conditions are later relaxed in Pan and Zhou (2008), Lytova and Pastur (2009), Shcherbina (2011), Zheng, Bai and Yao (2015), Najim and Yao (2016). It is worth mentioning that there are also fruitful results on another very important aspect of random matrix theory, the so-called Tracy–Widom law (known as T–W law). This law describes the fluctuation of the extreme eigenvalue around its mean after proper normalization. We refer the readers to Tracy and Widom (1994), Johnstone (2001), El Karoui (2007), Lee and Schnelli (2016), Johnstone and Ma (2012) and reference therein.

Compared to the eigenvalue studies, the literature regarding properties of the eigenvectors of a random matrix is very limited due to the difficulty of mathematical formulation since we are interested in the situation where the dimension increases with the sample size. The earliest work is Silverstein (1990), where the real sample covariance matrices are considered. Then this line of work is continued by Bai, Miao and Pan (2007), Pan and Zhou (2008), Ledoit and Péché (2011). In the former two papers, some isotropic conditions (see ($\mathcal{I}$C1) and ($\mathcal{I}$C2) in the beginning of Section 3) are assumed. It is shown that, under the isotropic condition ($\mathcal{I}$C1), the VESD of sample covariance matrix still converges to the M–P law. Recently, Yang (2020) considers the properties of eigenvector of sample covariance matrix when the population covariance matrix is anisotropic under some regular conditions. The convergence rate of VESD of the general sample covariance matrix is considered by Xia, Qin and Bai (2013) and more recently by Xi, Yang and Yin (2020).

Concerning the eigenstructure of a sample correlation matrix, Jiang (2004) was the first to prove that when $\mathbf{\Gamma} = \mathbf{I}_p$, if the dimension p and the sample size n are such that $p/n \to y \in (0, \infty)$, then the ESD of $\widehat{\mathbf{R}}_n$ converges to the Marčenko–Pastur distribution. The almost sure convergence of the smallest eigenvalue of a sample correlation matrix to the left edge of the spectrum was obtained in Xiao and Zhou (2010). The more general case where $\mathbf{G}_n \neq \mathbf{I}$ was considered in El Karoui (2009), where by applying the concentration of measure theory, the author proved that under a moment condition slightly stronger than that on the fourth moment, the ESD of $\widehat{\mathbf{R}}_n$ will also converge to the M–P law and the extreme eigenvalues will converge to the edges. Then the Tracy–Widom law for sample correlation matrix was obtained in Bao, Pan and Zhou (2012), Pillai and Yin (2012). The CLT for LSS of the sample correlation matrix under $\mathbf{G}_n = \mathbf{I}$ was investigated in Gao et al. (2017). As for more general cases, we refer the readers to Mestre and Vallet (2017), where the population is assumed to be Gaussian. Recently, the asymptotic result of eigenstructure of the sample correlation matrix for spiked model is investigated in Morales-Jimenez et al. (2021), whose sample covariance version is obtained in Paul (2007).

To the best of our knowledge, no work has been done to investigate the properties of eigenvectors of the sample correlation matrix even under the Gaussian case except for the specific spike model Morales-Jimenez et al. (2021).

2.2. *Our contributions.* The contributions of this paper are summarized as follows.

(1): We show that the ESD of the sample correlation matrix $\widehat{\mathbf{R}}_n$ convergences to the M–P law under a fourth moment assumption, while the extreme eigenvalues tend almost surely to the edges of the spectrum. Our results improve that in Jiang (2004), Xiao and Zhou (2010), which assumes $\mathbf{G} = \mathbf{I}$. Our results also improve Theorem 1 in El Karoui (2009), which relaxed the assumption $\mathbf{G} = \mathbf{I}$, but required an assumption stronger than the fourth moment condition of the underlying distribution.

(2): We establish a theorem concerning the CLT for the linear statistics of the eigenvectors of the large-dimensional sample covariance matrix. This theorem improves the main results of Bai, Miao and Pan (2007), which requires the fourth moment of the underlying variable to match that of the Gaussian distribution, as well as Theorem 1.3 in Pan and Zhou (2008), which relaxed the Gaussian like fourth moment condition but assumed the maximum entries of the projection vector converge to 0 uniformly.

(3): We prove an original CLT for the linear statistics of the eigenvectors of the large-dimensional sample correlation matrix. We show that the difference between the functional CLT of a sample covariance matrix and a sample correlation matrix is essentially influenced by the direction of a nonrandom projection vector π_n. To be specific, if $\|\pi_n\|_\infty \to 0$, where $\|\cdot\|_\infty$ stands for the ℓ_∞-norm, then there will be no difference between the forms of these two CLTs. The only difference is based on the change of the spectra of "population covariance matrix". However, this is not the case when the maximum entries of π_n do not converge to 0. In the special case where $\mathbf{G}_n = \mathbf{I}$, the additional terms will be determined by the sum of the fourth power of the entries of π_n. These results also indicate that unlike sample covariance matrix, the eigenmatrix of sample correlation matrices should *not* be asymptoticly Haar when the fourth moment equals to 3. This means normalization using the sample variances does affect the asymptotic properties of the eigenmatrix of the Wishart matrix even when the population covariance matrix equals to identity.

3. Main theorems. Before presenting the main theorem, we first list the necessary assumptions:

(C1) [*Assumption on Distribution I*]: Let $\{x_{ij}\}$ be a double array of infinite rows and columns. Assume that the entries of $\{x_{ij}\}$ are i.i.d. complex random variables with zero mean and unit variance. $\mathbf{X} = (x_{i,j})$, where $1 \le i \le p$, $1 \le j \le n$ is a $p \times n$ matrix of the upper-left corner of an infinite length double array;

(C2) [*Assumption on Distribution II*]: $\mathrm{E}\,|x_{1,1}|^4 < \infty$;

(C3) [*High-Dimensional Framework*]: $\rho_n = p/n \to \rho \in (0, \infty)$ as $n \to \infty$;

(C4) [*Population Correlation matrix*]: $\mathbf{G}_n$ is nonrandom with its spectral norm bounded in p. $H_n = F^{\mathbf{R}_n} \xrightarrow{D} H$, a proper distribution function.

(C5) [*Assumptions on Integrand*]: $\zeta_1, \dots, \zeta_k$ are k functions defined and analytic on an open region $\mathcal{D}$ of the complex plane which contains the real interval

$$\Bigl[\liminf_n \lambda_{\min}^{\mathbf{R}_n} I_{(0,1)}(\rho)(1-\sqrt{\rho})^2, \limsup_n \lambda_{\max}^{\mathbf{R}_n}(1+\sqrt{\rho})^2\Bigr]. \tag{6}$$

($\mathcal{I}$C1) [*Isotropic condition I*]: π_n is a vector over the p-dimensional unit sphere satisfying that $\pi_n^*(\mathbf{R}_n - z\mathbf{I})^{-1}\pi_n \to s_H(z)$, where $s_H(z)$ is the Stieltjes transform of H.

($\mathcal{I}$C2) [*Isotropic condition II*]: Define $\mathcal{R}_n(z) = (\mathbf{I} + \underline{s}_{F^{\rho_n,H_n}}(z)\mathbf{R}_n)^{-1}$, here and in the following, $\underline{s}_{F^{\rho_n,H_n}}(z)$ is the unique solution to the self-consistent equation

$$\underline{s}_{F^{\rho_n,H_n}}(z) = -\left(z - \rho_n \int \frac{t}{1 + t\underline{s}_{F^{\rho_n,H_n}}(z)}\, dH_n(t)\right)^{-1}.$$

Then

$$\sup_z \sqrt{n}\left|\pi_n^* \mathcal{R}_n(z)\pi_n - \int \frac{1}{\underline{s}_{F^{\rho_n,H_n}}(z)t+1}\,dH_n(t)\right| \to 0.$$

($\mathcal{I}$C3) [*Isotropic condition III*]:

$$\sup_{z,k} \sqrt{n}\left|e_k'\mathcal{R}_n(z)\pi_n + z^{-1}\pi_{n(k)} \int \frac{1}{\underline{s}_{F^{\rho_n,H_n}}(z)t+1}\,dH_n(t)\right| \to 0,$$

$$\sup_{z,k} \sqrt{n}\left|\pi_n^*\mathcal{R}_n(z)e_k + z^{-1}\bar{\pi}_{n(k)} \int \frac{1}{\underline{s}_{F^{\rho_n,H_n}}(z)t+1}\,dH_n(t)\right| \to 0,$$

$$\sup_{z,k} \sqrt{n}\left|e_k'\mathbf{G}_n^*\mathcal{R}_n(z)\pi_n + z^{-1}e_k'\mathbf{G}_n^*\pi_n \int \frac{1}{\underline{s}_{F^{\rho_n,H_n}}(z)t+1}\,dH_n(t)\right| \to 0,$$

$$\sup_{z,k} \sqrt{n}\left|\pi_n^*\mathcal{R}_n(z)\mathbf{G}_n e_k + z^{-1}\pi_n^*\mathbf{G}_n e_k \int \frac{1}{\underline{s}_{F^{\rho_n,H_n}}(z)t+1}\,dH_n(t)\right| \to 0.$$

REMARK 3.1. (C1)–(C4) are standard assumptions in spectral analysis in RMT. ($\mathcal{I}$C1)–($\mathcal{I}$C3) can be viewed as isotropic conditions, which are critical for obtaining the properties of eigenvector of sample correlation matrix under isotropic case and will be used in the second to last equations of (42)–(51) in Section 5.5.5. Note that all those conditions hold for all projection vector π_n if $\mathbf{R}_n = \mathbf{I}$ or more general, if $\sqrt{n}(\lambda_{\max}(\mathbf{R}_n) - \lambda_{\min}(\mathbf{R}_n)) \to 0$.

We first give a theorem concerning the LSD of a sample correlation matrix $\widehat{\mathbf{R}}_n$ as an improvement to Jiang (2004), Xiao and Zhou (2010) and Theorem 1 in El Karoui (2009). It is worth reminding the readers that in the following except Theorem 3.5, the parameters H_n, H, $s(z)$ and $\underline{s}(z)$ all refer to the correlation matrix $\mathbf{R}_n$ other than the covariance matrix $\mathbf{\Sigma}_n$. However, they are the same when the diagonal entries of $\mathbf{\Sigma}_n$ all equal to one.

THEOREM 3.1 (Theorem on eigenvalues I). *Under the Assumptions* (C1)–(C4), *we have*

$$F^{\widehat{\mathbf{R}}_n}(x) \to F^{\rho,H}(x), \quad a.s.$$

When $\mathbf{G}_n = \mathbf{I}$, *we have*

$$\lambda_{\max}(\widehat{\mathbf{R}}_n) \to (1+\sqrt{\rho})^2, \quad a.s.,$$
$$\lambda_{\min}(\widehat{\mathbf{R}}_n) \to I_{(0,1)}(\rho)(1-\sqrt{\rho})^2, \quad a.s.$$

In addition, if $|x_{1,1}| \le \eta_n\sqrt{n}$, *where* η_n *tends to* 0, *we have for any* $\varepsilon > 0$,

$$\mathrm{P}\big(\lambda_{\max}(\widehat{\mathbf{R}}_n) > (1+\sqrt{\rho})^2 + \varepsilon\big) = O\big(n^{-l}\big),$$
$$\mathrm{P}\big(\lambda_{\min}(\widehat{\mathbf{R}}_n) < I_{(0,1)}(\rho)(1-\sqrt{\rho})^2 - \varepsilon\big) = O\big(n^{-l}\big),$$

for all positive integer l.

The above theorem implies that the existence of the fourth moment of the underlying variable is sufficient to ensure the convergence of LSD of a sample correlation matrix to the M–P law. Meanwhile, after truncation, the extreme eigenvalues will tend to the edges of the LSD with high probability.

We next present a theorem concerning the no eigenvalue outside result for the sample correlation matrix.

THEOREM 3.2 (Theorem on eigenvalues II). *Under the Assumptions* (C1)–(C4), *assume further that the interval* $[a, b]$ *with* $a > 0$ *lies outside the support of* $F^{\rho,H}$ *and* F^{ρ_n,H_n} *for all* n *sufficiently large. Then we have*

$$\mathrm{P}\big(\textit{no eigenvalues of } \widehat{\mathbf{R}}_n \textit{ appears in } [a, b] \textit{ for all large } n\big) = 1.$$

Now we consider the properties of the eigenvectors of the sample correlation matrix. For any given projection vector π_n and $z \in \mathbb{C}^+$, define

$$s^{\mathbf{R}_n}_{\rho_n,\pi_n}(z) = -z^{-1}\pi_n^*\mathcal{R}_n(z)\pi_n.$$

It is then easy to see that $s^{\mathbf{R}_n}_{\rho_n,\pi_n}(z)$ is the Stieltjes transform of a distribution, which we shall denote by $F^{\mathbf{R}_n}_{\rho_n,\pi_n}$. Note that this class of distributions is also called the anisotropic M–P laws, see for instance Xi, Yang and Yin (2020).

First, we have the following theorem on the vector limit spectral distribution (VLSD).

THEOREM 3.3 (Theorem on VLSD). *Under the Assumptions* (C1)–(C4), *for any given* x, *we have*

$$F^{\widehat{\mathbf{R}}_n}_{\upsilon,\pi_n}(x) - F^{\mathbf{R}_n}_{\rho_n,\pi_n}(x) \to 0, \quad a.s.$$

In particular, if the condition ($\mathcal{I}C1$) *is satisfied, we have*

$$F^{\widehat{\mathbf{R}}_n}_{\upsilon,\pi_n}(x) \to F^{\rho,H}(x), \quad a.s.$$

Theorem 3.3 indicates that although ESD and VESD are not the same in general, they do share the same limit when $\mathbf{R}_n$ is close to identity.

Next, we present the functional CLT for VESD (denoted as CLT for vector linear spectral statistics (VLSS) in the following). Define

$$\mathbb{G}_n(x) = \sqrt{n}\big(F^{\widehat{\mathbf{R}}_n}_{\upsilon,\pi_n}(x) - F^{\mathbf{R}_n}_{\rho_n,\pi_n}(x)\big).$$

To make the statements of the next theorem more concise, we also give the following definitions. Denote

$$\mathcal{G}_{kl}(\nu_4, \alpha) = \left(\nu_4\sum_{i=1}^p |g_{ki}|^2|g_{li}|^2 + \alpha r_{k\ell}^2\right),$$

$$\begin{aligned}
&L_{0,0}(z_1, z_2)\\
&\quad= \frac{\nu_4\underline{s}(z_1)\underline{s}(z_2)}{z_1z_2}\lim_{n\to\infty}\sum_{k=1}^p e_k'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_ke_k'\mathbf{G}_n^*\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)\mathbf{G}_ne_k\\
&\qquad+\frac{\alpha(\underline{s}(z_2)-\underline{s}(z_1))}{z_1z_2(z_2-z_1)}\lim_{n\to\infty}\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_n\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)\mathbf{R}_n\mathcal{R}_n(z_1)\pi_n,
\end{aligned}$$

$$L_{1,1}(z_1, z_2) = \frac{1}{z_1z_2}\lim_{n\to\infty}\sum_{k,l=1}^p \mathcal{G}_{kl}(\nu_4, \alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*e_ke_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*e_l,$$

$$L_{2,2}(z_1, z_2) = \frac{1}{z_1z_2}\lim_{n\to\infty}\sum_{k,l=1}^p \mathcal{G}_{kl}(\nu_4, \alpha)e_k'\pi_n\pi_n^*\mathcal{R}_n(z_1)e_ke_l'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l,$$

$$L_{3,3}(z_1, z_2) = \frac{1}{z_1^2z_2^2}\lim_{n\to\infty}\sum_{k,l=1}^p \mathcal{G}_{kl}(\nu_4, \alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)e_ke_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l,$$

$$\begin{aligned}
&L_{0,1}(z_1,z_2)\\
&\quad=\frac{\underline{s}(z_1)}{z_1z_2}\lim_{n\to\infty}\sum_{k=1}^{p}e_k'\mathcal{R}_n(z_2)\pi_n\pi_n^*e_k\\
&\qquad\times\left(\nu_4\sum_{l=1}^{p}e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_l|g_{kl}|^2+\alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_ne_ke_k'\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\right),\\
&L_{0,2}(z_1,z_2)\\
&\quad=\frac{\underline{s}(z_1)}{z_1z_2}\lim_{n\to\infty}\sum_{k=1}^{p}e_k'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_k\\
&\qquad\times\left(\nu_4\sum_{l=1}^{p}e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_l|g_{kl}|^2+\alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_ne_ke_k'\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\right),\\
&L_{0,3}(z_1,z_2)\\
&\quad=-\frac{\underline{s}(z_1)}{z_1z_2^2}\lim_{n\to\infty}\sum_{k=1}^{p}e_k'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_k\\
&\qquad\times\left(\nu_4\sum_{l=1}^{p}e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_le_l'\mathbf{g}_k\mathbf{g}_k^*e_l+\alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_ne_ke_k^*\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\right),\\
&L_{1,2}(z_1,z_2)=\frac{1}{z_1z_2}\lim_{n\to\infty}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*e_ke_l'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l,\\
&L_{1,3}(z_1,z_2)=\frac{1}{z_1z_2^2}\lim_{n\to\infty}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*e_ke_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l,\\
&L_{2,3}(z_1,z_2)=\frac{1}{z_1z_2^2}\lim_{n\to\infty}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\pi_n\pi_n^*\mathcal{R}_n(z_1)e_ke_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l.
\end{aligned}$$

Then we have the following theorem.

THEOREM 3.4 (Theorem on CLT for VLSS for $\widehat{\mathbf{R}}_n$). *Under the Assumptions* (C1)–(C5), *set* $\nu_4=\mathrm{E}|x_{1,1}|^4-|\mathrm{E}x_{1,1}^2|^2-2$, $\alpha=\mathrm{E}(x_{1,1})^2+1$, *we have the following results*.

RT (Tightness): *The k-dimensional random vectors*

$$\Psi_n=(\psi_{1,n},\dots,\psi_{k,n})'=\left(\int\zeta_1(x)\,d\mathbb{G}_n(x),\dots,\int\zeta_k(x)\,d\mathbb{G}_n(x)\right)'$$

form a tight sequence.

RM (Limiting mean vector): *The random vectors Ψ_n converge weakly to a mean zero Gaussian vector $\Psi=(\psi_1,\dots,\psi_k)'$*.

RC (Limiting variance-covariance function): *Let* $\mathrm{E}(x_{1,1})^2=0$ *if $x_{1,1}$ is complex*, *then we have for* $1\le t,s\le k$,

$$\mathrm{Cov}(\psi_t,\psi_s)=-\frac{1}{2\pi^2}\int_{\mathcal{C}_1}\int_{\mathcal{C}_2}\zeta_t(z_1)\zeta_s(z_2)\varpi(z_1,z_2)\,dz_1\,dz_2, \tag{7}$$

where $\mathcal{C}_1, \mathcal{C}_2$ *are two nonoverlapping contours enclosing the support of* $F^{\rho,H}$ *and*

$$\begin{aligned}
&\varpi(z_1,z_2)\\
&\quad = L_{0,0}(z_1,z_2) + 4^{-1}L_{1,1}(z_1,z_2) + 4^{-1}L_{1,1}(z_1,z_2) + z_1z_2L_{3,3}(z_1,z_2)\\
&\qquad + 2^{-1}L_{0,1}(z_1,z_2) + 2^{-1}L_{0,2}(z_1,z_2) + 2^{-1}L_{0,1}(z_2,z_1) + 2^{-1}L_{0,2}(z_2,z_1)\\
&\qquad + z_2L_{0,3}(z_1,z_2) + z_1L_{0,3}(z_2,z_1) + 2^{-1}z_2L_{1,3}(z_1,z_2) + 2^{-1}z_1L_{1,3}(z_2,z_1)\\
&\qquad + 4^{-1}L_{1,2}(z_1,z_2) + 4^{-1}L_{1,2}(z_2,z_1) + 2^{-1}z_2L_{2,3}(z_1,z_2) + 2^{-1}z_1L_{2,3}(z_2,z_1).
\end{aligned}$$

The above theorem is general since we allow the anisotropic structure. Under the special case that the population correlation matrix is isotropic, also denote

$$\mathcal{L}_\pi = \lim_{n\to\infty}\sum_{k=1}^p |\pi_{(k)}|^4, \qquad \mathcal{L}_{\pi,\mathbf{G}}^{(1)} = \lim_{n\to\infty}\sum_{k=1}^p\left(\left|\sum_{l=1}^p \bar{g}_{l,k}\pi_{(l)}\right|^2\right)^2,$$

$$\mathcal{L}_{\pi,\mathbf{R}} = \lim_{n\to\infty}\sum_{k,l=1}^p r_{k\ell}^2|\pi_{(k)}|^2|\pi_{(l)}|^2, \qquad \mathcal{L}_{\pi,\mathbf{G}}^{(2)} = \lim_{n\to\infty}\sum_{k,l=1}^p\sum_{i=1}^p |g_{ki}|^2|g_{li}|^2|\pi_{(k)}|^2|\pi_{(l)}|^2,$$

$$\mathcal{L}_{\pi,\mathbf{G}}^{(3)} = \lim_{n\to\infty}\sum_{k,l=1}^p\left(|\pi_{(k)}|^2|g_{kl}|^2\left|\sum_{i=1}^p \bar{g}_{il}\pi_{(i)}\right|^2\right),$$

we have the following corollary.

COROLLARY 3.1. *Under the assumptions of Theorem* 3.4 *and assume the isotropic condition* ($\mathcal{I}$C1)–($\mathcal{I}$C3), *the results of Theorem* 3.4 *hold with*

$$\begin{aligned}
\varpi(z_1,z_2) &= L_{0,0}(z_1,z_2) + L_{1,1}(z_1,z_2) + z_1z_2L_{3,3}(z_1,z_2) + L_{0,1}(z_1,z_2) + L_{0,1}(z_2,z_1)\\
&\quad + z_2L_{0,3}(z_1,z_2) + z_1L_{0,3}(z_2,z_1) + z_2L_{1,3}(z_1,z_2) + z_1L_{1,3}(z_2,z_1),
\end{aligned}$$

where

$$\begin{aligned}
L_{0,0}(z_1,z_2) &= \left(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(1)}z_1z_2s^2(z_1)s^2(z_2)\underline{s}(z_1)\underline{s}(z_2) + \frac{\alpha(z_2\underline{s}(z_2)-z_1\underline{s}(z_1))^2}{\rho^2z_1z_2(z_2-z_1)(\underline{s}(z_2)-\underline{s}(z_1))}\right),\\
L_{1,1}(z_1,z_2) &= (\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)} + \alpha\mathcal{L}_{\pi,\mathbf{R}})s(z_1)s(z_2),\\
L_{3,3}(z_1,z_2) &= (\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)} + \alpha\mathcal{L}_{\pi,\mathbf{R}})s^2(z_1)s^2(z_2),\\
L_{0,1}(z_1,z_2) &= -\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(3)}z_1\underline{s}(z_1)s^2(z_1)s(z_2) - \frac{\alpha\mathcal{L}_\pi s(z_2)(1+z_1s(z_1))^2}{z_1\underline{s}(z_1)},\\
L_{0,3}(z_1,z_2) &= -\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(3)}z_1\underline{s}(z_1)s^2(z_1)s^2(z_2) - \frac{\alpha\mathcal{L}_\pi s^2(z_2)(1+z_1s(z_1))^2}{z_1\underline{s}(z_1)},\\
L_{1,3}(z_1,z_2) &= (\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)} + \alpha\mathcal{L}_{\pi,\mathbf{R}})s(z_1)s^2(z_2).
\end{aligned}$$

The next corollary, which gives the properties of the eigenvector of the "normalized with sample variances" Wishart matrix is immediate.

COROLLARY 3.2. *Under the assumptions of Theorem* 3.4, *if* $\mathbf{G}_n = \mathbf{I}_n$, *then we have*

$$\operatorname{Cov}(\psi_t,\psi_s) = -\frac{1}{2\pi^2}\int_{\mathcal{C}_1}\int_{\mathcal{C}_2}\zeta_t(z_1)\zeta_s(z_2)\varpi(z_1,z_2)\,dz_1\,dz_2,$$

where $\mathcal{C}_1, \mathcal{C}_2$ are two nonoverlapping contours enclosing the support of $F^{\rho,H}$ and

$$\begin{aligned}\varpi(z_1,z_2)\\ &= \left(\frac{(4\nu_4+3\alpha)\mathcal{L}_\pi \underline{s}(z_1)\underline{s}(z_2)}{z_1z_2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))^2} + \frac{\alpha(z_2\underline{s}(z_2)-z_1\underline{s}(z_1))^2}{\rho^2 z_1z_2(z_2-z_1)(\underline{s}(z_2)-\underline{s}(z_1))}\right).\end{aligned}$$

REMARK 3.2. Under the special case where $\mathbf{G}_n = \mathbf{I}_n$, comparing with Theorem 3.5 below we can see that the additional term caused by normalization using the sample variance is a term multiplied by $3\nu_4 + 3\alpha$. When the underlying variables are real, we have $\alpha = \mathrm{E}(x_{1,1})^2 + 1 = 2$. Thus the additional term will vanish if $\nu_4 = -2$. This is quite sensible since $\nu_4 = -2$ corresponds to the single case where the underlying variables follow the Bernoulli distribution with $\mathrm{P}(x_{1,1} = -1) = \mathrm{P}(x_{1,1} = -1) = 1/2$, thus the diagonal entries of $\widehat{\mathbf{S}}_n$ all equal to 1, which implies $\widehat{\mathbf{R}}_n = \widehat{\mathbf{S}}_n$.

It is seen from the theorem that in the special case where $\mathbf{G}_n = \mathbf{I}$, the additional terms in the functional CLT of the eigenvectors caused by normalization are all determined by the sum of the fourth power of the entries of π_n. This result gives strong evidence that the eigenmatrix of sample correlation matrices should *not* be asymptoticly Haar unless $4\nu_4 + 3\alpha = 0$. In other words, the normalization by the sample variances does affect the asymptotic properties of the eigenmatrix of the Wishart matrix. A surprising observation here is when $\nu_4 = -1.5$ under real case ($\mathrm{E}\,|x_{1,1}|^4 = 1.5$)]. This causes the term multiplied by π_n to vanish and the eigenmatrix of the sample correlation matrix to satisfy the necessary condition of asymptoticly Haar, see Section 10.2 in Bai and Silverstein (2010).

We also have the following improved CLT for VLSS for $\widehat{\mathbf{S}}_n$, which improves the results of Theorem 2 in Bai, Miao and Pan (2007) as well as Theorem 1.3 in Pan and Zhou (2008).

THEOREM 3.5 (Theorem on CLT for VLSS for $\widehat{\mathbf{S}}_n$). *Assume the Assumptions* (C1)–(C5), *where all $\mathbf{R}_n$ and $\widehat{\mathbf{R}}_n$ that appear in the assumptions and the definitions of $\underline{s}_{F^{\rho_n,H_n}}(z)$, $\mathbb{G}_n(x)$ are replaced with $\mathbf{\Sigma}_n$ and $\widehat{\mathbf{S}}_n$, respectively. Set $\nu_4 = \mathrm{E}\,|x_{1,1}|^4 - |\,\mathrm{E}\,x_{1,1}^2|^2 - 2$, $\alpha = \mathrm{E}(x_{1,1})^2 + 1$. Let $\mathcal{S}_n(z) = (\mathbf{I} + \underline{s}_{F^{\rho_n,H_n}}(z)\mathbf{\Sigma}_n)^{-1}$. We have*:

(1): *The same results as* RT, RM *in Theorem* 3.4 *hold.*
(2): *The same result as* RC *in Theorem* 3.4 *holds with*

$$\begin{aligned}\varpi(z_1,z_2)\\ &= \frac{\nu_4\underline{s}(z_1)\underline{s}(z_2)}{z_1z_2}\lim_{n\to\infty}\sum_{k=1}^p e_k'\mathbf{\Gamma}_n^*\mathcal{S}_n(z_1)\pi_n\pi_n^*\mathcal{S}_n(z_1)\mathbf{\Gamma}_n e_k e_k'\mathbf{\Gamma}_n^*\mathcal{S}_n(z_2)\pi_n\pi_n^*\mathcal{S}_n(z_2)\mathbf{\Gamma}_n e_k\\ &\quad + \frac{\alpha(\underline{s}(z_2)-\underline{s}(z_1))}{z_1z_2(z_2-z_1)}\lim_{n\to\infty}\pi_n^*\mathcal{S}_n(z_1)\mathbf{\Sigma}_n\mathcal{S}_n(z_2)\pi_n\pi_n^*\mathcal{S}_n(z_2)\mathbf{\Sigma}_n\mathcal{S}_n(z_1)\pi_n.\end{aligned}$$

If, in addition, ($\mathcal{I}$C1) *and* ($\mathcal{I}$C2) *are satisfied with $\mathbf{R}_n$ replaced by $\mathbf{\Sigma}_n$, we have*

$$\varpi(z_1,z_2) = \left(\frac{\nu_4\mathcal{L}_\pi\underline{s}(z_1)\underline{s}(z_2)}{z_1z_2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))^2} + \frac{\alpha(z_2\underline{s}(z_2)-z_1\underline{s}(z_1))^2}{\rho^2z_1z_2(z_2-z_1)(\underline{s}(z_2)-\underline{s}(z_1))}\right).$$

REMARK 3.3. Comparing with the results in Bai, Miao and Pan (2007) and Pan and Zhou (2008), we allow more general population covariance matrix structure. It is easy to see that the main difference comes from the first term of $\varpi(z_1, z_2)$. We try to give some insight of this term here. Consider the special case where $\mathbf{\Gamma}_n = \mathbf{I}$ and the underlying variables follow the real Bernoulli distribution with $\mathrm{P}(x_{1,1} = -1) = \mathrm{P}(x_{1,1} = -1) = 1/2$. Let $\pi_n =$

$e_1 = (1, 0, \dots,)'$. Then we know that the diagonal entries of $\widehat{\mathbf{S}}_n$ all equal to 1 and $\nu_4 = -2$, $\alpha = 2$. Also, one can obtain that the two terms of $\varpi(z_1, z_2)$ are exactly

$$\frac{-2\underline{s}(z_1)\underline{s}(z_2)}{z_1 z_2 (1+\underline{s}(z_1))^2 (1+\underline{s}(z_2))^2} \quad \text{and} \quad \frac{2(\underline{s}(z_2) - \underline{s}(z_1))}{z_1 z_2 (z_2 - z_1)(1+\underline{s}(z_1))^2 (1+\underline{s}(z_2))^2}.$$

When $z_1 = z_2 = z$ (here $\frac{\underline{s}(z_2)-\underline{s}(z_1)}{z_2 - z_1}$ is understood as $\frac{d\underline{s}(z)}{dz}$), we have

$$\begin{aligned}\varpi(z, z) &= \frac{-2}{z^2(1+\underline{s}(z))^4}\left(\underline{s}^2(z) - \frac{d\underline{s}(z)}{dz}\right) \\ &= \frac{-2}{z^2(1+\underline{s}(z))^4}\left(\underline{s}^2(z) - \frac{\underline{s}^2(z)}{1 - \frac{\rho \underline{s}^2(z)}{(1+\underline{s}^2(z))^2}}\right),\end{aligned}$$

which will tend to 0 as $\rho \to 0$. Recall that under this case, for given z, $\varpi(z, z)$ is the limiting "complex variance" (the expectation of the square of the difference between a complex random variable and its expectation) of the first diagonal entry of the random matrix $\sqrt{n}(\widehat{\mathbf{S}}_n - z\mathbf{I})^{-1}$, which should equal to 0 when $p = 1$. The above arguments indicate the consistency of our results under this special case since $\varpi(z, z)$ is a continued function of ρ.

4. An application in communications. In this section, we discuss the application of our main theorem in communications.

Consider a symbol synchronous direct sequence code division multiple access (DS-CDMA) system with n users and p processing gains. The discrete-time model for the received signal $\mathbf{r}$ in a symbol interval is

$$\mathbf{r} = \sum_{j=1}^{n} x_j \mathbf{s}_j + \boldsymbol{\epsilon},$$

where $\mathbf{s}_j = (s_{1,j}, s_{2,j}, \dots, s_{p,j})' \in \mathbb{C}^p$ is the signature sequence of user j and x_j is the transmitted symbol of user j. Further, $\boldsymbol{\epsilon} \sim N(0, \sigma^2\mathbf{I})$ is the background Gaussian noise. Assume the x_j's are independent with mean zero and variance P_j's, where each P_j is known as the received power of the corresponding user. The goal in wireless communications is to demodulate the transmitted x_j for each user and a relevant performance measure is the classical signal-to-interference ratio (SIR). Define $\mathbb{S}_j = (\mathbf{s}_1, \dots, \mathbf{s}_{j-1}, \mathbf{s}_{j+1}, \dots, \mathbf{s}_n)$ and $\mathbb{P}_j = \operatorname{diag}(P_1, \dots, P_{j-1}, P_{j+1}, \dots, P_n)$, then the SIR for user j is defined as

$$\beta_j = P_j \mathbf{s}_j^* (\mathbb{S}_j \mathbb{P}_j \mathbb{S}_j^* + \sigma^2 \mathbf{I})^{-1} \mathbf{s}_j, \quad 1 \le j \le n.$$

See Tse and Hanly (1999) for more details.

However, due to high computational cost (Pan and Zhou (2008)) of the classical SIR, some simpler but near performance measures have been considered. For example, Honig and Xiao (2001) considered the reduced-rank linear receiver for multistage Wiener (MSW), where the output SIR for user j is defined as

$$\beta_{jm} = P_j \mathbf{s}_j^* \mathbb{A}_{jm} (\mathbb{A}_{jm}^* \mathbb{Q}_j \mathbb{A}_{jm})^{-1} \mathbb{A}_{jm}^* \mathbf{s}_j, \tag{8}$$

where $m \le n$ and $\mathbb{A}_{jm} = [\mathbf{s}_j, \mathbb{Q}_j \mathbf{s}_j, \dots, \mathbb{Q}_j^{m-1} \mathbf{s}_j]$ with $\mathbb{Q}_j = (\mathbb{S}_j \mathbb{P}_j \mathbb{S}_j^* + \sigma^2 \mathbf{I})$. By modeling the signature sequences $\mathbf{s}_j$'s as random vectors, one may analyze the SIR using random matrix theory when the number of users and the processing gain approach infinity simultaneously (known as the large system). For example, assuming the entries $s_{i,j}$ are i.i.d. with zero mean and common variance $1/p$, Pan and Zhou (2008) proved that the asymptotic distribution of the SIR under MSW is Gaussian after proper normalization.

Indeed, in the random signature sequences model, for the user j, the signature sequence covariance matrix $\mathbb{H}_j := \mathbb{S}_j \mathbb{S}_j^*$ is a random sample covariance matrix which is independent with its signature sequence $\mathbf{s}_j$. By the analysis in page 1246 in Pan and Zhou (2008), especially equations (3.1)–(3.4), to derive the limit distribution of SIR, it is sufficient to consider the joint distribution of

$$\left\{\sqrt{p}\left(\mathbf{s}_j^* \mathbb{H}_j \mathbf{s}_j - \frac{1}{p} \operatorname{tr} \mathbb{H}_j\right), \ldots, \sqrt{p}\left(\mathbf{s}_j^* \mathbb{H}_j^{2m-1} \mathbf{s}_j - \frac{1}{p} \operatorname{tr} \mathbb{H}_j^{2m-1}\right)\right\}.$$

Note that for any given u,

$$\sqrt{p}\left(\mathbf{s}_j^* \mathbb{H}_j^u \mathbf{s}_j - \frac{1}{p} \operatorname{tr} \mathbb{H}_j^u\right) = \sqrt{p} |\mathbf{s}_j|^2 \left(\frac{\mathbf{s}_j^* \mathbb{H}_j^u \mathbf{s}_j}{|\mathbf{s}_j|^2} - \frac{1}{p} \operatorname{tr} \mathbb{H}_j^u\right) + \sqrt{p} \frac{1}{p} \operatorname{tr} \mathbb{H}_j^u (|\mathbf{s}_j|^2 - 1),$$

where the fluctuation of the term $\sqrt{p}(\frac{\mathbf{s}_j^* \mathbb{H}_j^u \mathbf{s}_j}{|\mathbf{s}_j|^2} - \frac{1}{p} \operatorname{tr} \mathbb{H}_j^u)$ can be determined by applying the functional CLT for eigenvector of sample covariance matrix.

However, since the difference components in $\mathbf{s}_j$ represent different antennas, it is more reasonable to assume the variabilities differ for these components. This implies we should first normalize the signature sequence $\mathbf{s}_j$'s using sample variances before the analysis described above. To be more specific, define

$$t_{i,j} = \frac{s_{i,j}}{\sqrt{v_i^{(j)}}}, \quad 1 \le i \le p, 1 \le j \le n,$$

where $v_i^{(j)} = (n-1)^{-1} \sum_{k \neq j} s_{i,k}^2$. Let

$$\mathbf{t}_j = (t_{1,j}, t_{2,j}, \ldots, t_{p,j})', \qquad \mathbb{T}_j = (\mathbf{t}_1, \ldots, \mathbf{t}_{j-1}, \mathbf{t}_{j+1} \cdots, \mathbf{t}_n),$$

$\mathbb{B}_{jm} = [\mathbf{s}_j, \mathbb{V}_j \mathbf{s}_j, \ldots, \mathbb{V}_j^{m-1} \mathbf{s}_j]$ with $\mathbb{V}_j = (\mathbb{T}_j \mathbb{P}_j \mathbb{T}_j^* + \sigma^2 \mathbf{I})$, the output SIR for user j is then

$$\widetilde{\beta}_{jm} = P_j \mathbf{t}_j^* \mathbb{B}_{jm} (\mathbb{B}_{jm}^* \mathbb{V}_j \mathbb{B}_{jm})^{-1} \mathbb{B}_{jm}^* \mathbf{t}_j. \tag{9}$$

We have the following result when $P_1 = P_2 = \cdots = P_n = 1$, which is a direct application of Corollary 3.2.

THEOREM 4.1. *Assume that the* $\{w_{i,j}, i, j = 1, \ldots\}$ *are i.i.d. complex random variables with zero mean, unit variance and finite fourth moment* μ_4. *The entries of the signature sequences are* $s_{i,j} = \frac{\sigma_i w_{i,j}}{\sqrt{p}}$, *where* $\sigma_i > 0$ *and* $\sup_i \sigma_i < \infty$ *for* $i = 1, \ldots, p$, $j = 1, \ldots, n$. *In addition*, $p/n \to \infty$ *as* $n \to \infty$. *Then for any integer* m, $\widetilde{\beta}_{jm}$ *has the same asymptotical distribution as* β_{jm} *under the situation when* $\sigma_1 = \cdots = \sigma_p = 1$.

Theorem 4.1 indicates that we can normalize the signature sequences using the sample variances before computing the output SIR. After normalization, the fluctuation of the output SIR under the diagonal covariance matrix case is the same as the one under the case where the covariance matrix equals to the identity. The asymptotic distribution of β_{1m} under the situation when $\sigma_1 = \cdots = \sigma_p = 1$ was given in Theorem 1.1 in Pan and Zhou (2008) and Theorem 4.1 ensures that this result is still valid even if the unit variance assumption does not hold.

5. Proof of the main theorems. We organize the proof of these theorems as the following. In Section 5.1, we list the notation and lemmas that will be used in our proofs. Sections 5.2–5.6 contain the detailed proofs of the theorems.

5.1. *Notations and primary lemmas.* We begin by showing necessary notation and primary lemmas. Let

$$\mathbf{r}_j = \mathbf{G}_n \mathbf{x}_j/\sqrt{n}, \qquad \mathbf{A}(z) = \sum_{i=1}^{n} \mathbf{r}_i \mathbf{r}_i^T - z\mathbf{I}, \qquad \mathbf{A}_j(z) = \mathbf{A}(z) - \mathbf{r}_j \mathbf{r}_j^T,$$

and

$$\mathbf{A}_{kj}(z) = \mathbf{A}_j(z) - \mathbf{r}_k \mathbf{r}_k^T \quad \text{for } 1 \le j, k \le n.$$

Denote

$$\breve{\mathbf{A}}_j(z) = \sum_{i<j} \mathbf{r}_i \mathbf{r}_i^T + \sum_{i>j} \breve{\mathbf{r}}_i \breve{\mathbf{r}}_i^T - z\mathbf{I},$$

where $\breve{\mathbf{r}}_{j+1}, \dots, \breve{\mathbf{r}}_n$ are independent copies of $\mathbf{r}_{j+1}, \dots, \mathbf{r}_n$. Let E_j denote the conditional expectation given the samples $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_j$, also let $\mathrm{E}_{(-j)}$ denote the conditional expectation given $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{j-1}, \breve{\mathbf{x}}_{j+1}, \breve{\mathbf{x}}_{j+2}, \dots, \breve{\mathbf{x}}_n$. Moreover, define

$$b_j(z) = \frac{1}{1 + \frac{1}{n}\operatorname{tr}(\mathbf{A}_j^{-1}(z)\mathbf{R}_n)}, \qquad b_{ij}(z) = \frac{1}{1 + \frac{1}{n}\operatorname{tr}(\mathbf{A}_{ij}^{-1}(z)\mathbf{R}_n)},$$

$$\breve{b}_j(z) = \frac{1}{1 + \frac{1}{n}\operatorname{tr}(\breve{\mathbf{A}}_j^{-1}(z)\mathbf{R}_n)}, \qquad b(z) = \frac{1}{1 + \frac{1}{n}\mathrm{E}\operatorname{tr}(\mathbf{A}^{-1}(z)\mathbf{R}_n)},$$

$$\beta_{i(j)}(z) = \frac{1}{1 + \mathbf{r}_i^* \mathbf{A}_{ij}^{-1} \mathbf{r}_i}, \qquad \beta_j(z) = \frac{1}{1 + \mathbf{r}_j^* \mathbf{A}_j^{-1}(z) \mathbf{r}_j},$$

$$\widehat{\gamma}_j(z) = \mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\mathbf{r}_j - \frac{1}{n}\operatorname{tr}(\mathbf{A}_j^{-1}(z)\mathbf{R}_n),$$

$$\widehat{\gamma}_{i(j)}(z) = \mathbf{r}_i^T \mathbf{A}_{ij}^{-1}(z)\mathbf{r}_i - \frac{1}{n}\operatorname{tr}(\mathbf{A}_{ij}^{-1}(z)\mathbf{R}_n),$$

$$\mathbb{D}_{\mathbf{\Xi}_n}^{(-j)} = \operatorname{diag}\left(\frac{1}{n}\sum_{i \ne j} \mathbf{\Xi}_i\right) - \frac{n-1}{n}\mathbf{I}_p,$$

$$\delta_j(z) = \mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\pi_n \pi_n^* \mathbf{A}_j^{-1}(z)\mathbf{r}_j - \frac{1}{n}\pi_n^* \mathbf{A}_j^{-1}(z)\mathbf{R}_n \mathbf{A}_j^{-1}(z)\pi_n.$$

It is easy to see that

$$\max(|b_j(z)|, |b_{ij}(z)|, |b(z)|, |\beta_{i(j)}(z)|, |\beta_j(z)|) \le |z|/\operatorname{Im}(z), \tag{10}$$

where $\operatorname{Im}(z)$ stands for the imaginary part of z. With the notation above, $\mathbf{A}^{-1}(z)$ can be decomposed as

$$\mathbf{A}^{-1}(z) = \mathbf{A}_j^{-1}(z) - \beta_j(z)\mathbf{A}_j^{-1}(z)\mathbf{r}_j \mathbf{r}_j^T \mathbf{A}_j^{-1}(z), \tag{11}$$

and $\beta_j(z)$ can be further written as

$$\beta_j(z) = b_j(z) - \beta_j(z) b_j(z)\widehat{\gamma}_j(z) = b_j(z) - b_j^2(z)\widehat{\gamma}_j(z) + \beta_j(z) b_j^2(z)\widehat{\gamma}_j^2(z), \tag{12}$$

and

$$\beta_{i(jk)}(z) = b_{ijk}(z) - \beta_{i(jk)}(z) b_{ijk}(z)\widehat{\gamma}_{i(jk)}(z). \tag{13}$$

The matrix $\mathbf{A}_j^{-1}(z)$ can be further decomposed as

$$\mathbf{A}_j^{-1}(z) = \mathbf{B}_{j1}(z) + \mathbf{B}_{j2}(z) + \mathbf{B}_{j3}(z) + \mathbf{B}_{j4}(z), \tag{14}$$

where

$$\mathbf{B}_{j1}(z) = -\left(z\mathbf{I} - \frac{n-1}{n}b(z)\cdot\mathbf{R}_n\right)^{-1} := \mathbb{R}_n(z),$$

$$\mathbf{B}_{j2}(z) = b(z)\sum_{i\neq j}\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^T - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_{ij}^{-1}(z),$$

$$\mathbf{B}_{j3}(z) = \sum_{i\neq j}(\beta_{i(j)}(z) - b(z))\mathbb{R}_n(z)\mathbf{r}_i\mathbf{r}_i^T\mathbf{A}_{ij}^{-1}(z) \quad \text{and}$$

$$\mathbf{B}_{j4}(z) = n^{-1}b(z)\mathbb{R}_n(z)\sum_{i\neq j}(\mathbf{A}_{ij}^{-1}(z) - \mathbf{A}_j^{-1}(z)) = -n^{-1}b(z)\mathbb{R}_n(z)\sum_{i\neq j}\mathbf{A}_{ij}^{-1}\mathbf{r}_i\mathbf{r}_i^T\mathbf{A}_{ij}^{-1}\beta_{i(j)}.$$

We are now able to present the following lemmas. The first lemma is used to bound the moments of some random quadratic forms, which play important role in random matrix theory.

LEMMA 5.1 (Lemma B.26 in Bai and Silverstein (2010)). *Let* $\mathbf{A} = (a_{jk})$ *be an* $n\times n$ *nonrandom matrix and* $\mathbf{x} = (x_1,\dots,x_n)'$ *be a random vector of independent entries. Assume that* $\mathrm{E}x_j = 0$, $\mathrm{E}|x_j|^2 = 1$ *and* $\mathrm{E}|x_j|^l \le \nu_l$. *Then for* $p\ge 1$,

$$\mathrm{E}\left|\mathbf{x}^*\mathbf{A}\mathbf{x} - \operatorname{tr}\mathbf{A}\right|^p \le C_p\left[(\nu_4\operatorname{tr}\mathbf{A}\mathbf{A}^*)^{p/2} + \nu_{2p}\operatorname{tr}(\mathbf{A}\mathbf{A}^*)^{p/2}\right],$$

where C_p *is a constant depending on* p *only.*

The second lemma is used to obtain the covariance of two random quadratic forms. It will be applied to find the limiting variance-covariance functions when establishing the CLTs.

LEMMA 5.2 (See (1.15) in Bai and Silverstein (2004)). *Let* $\mathbf{x} = (x_1,\dots,x_n)'$ *be a random vector of independent entries with* 0 *means and unit variances, then we have*

$$E(\mathbf{x}^*\mathbf{A}\mathbf{x} - \operatorname{tr}\mathbf{A})(\mathbf{x}^*\mathbf{B}\mathbf{x} - \operatorname{tr}\mathbf{B}) = (E|x_1|^4 - |Ex_1^2|^2 - 2)\sum_{i=1}^n a_{ii}b_{ii} + |Ex_1^2|^2\operatorname{tr}\mathbf{A}\mathbf{B}^T + \operatorname{tr}\mathbf{A}\mathbf{B}.$$

The next lemma is the well-known CLT for martingale.

LEMMA 5.3 (Theorem 35.12 of Billingsley (1995)). *Suppose that for each* n, $Y_{n1}, Y_{n2}, \dots, Y_{nr_n}$ *is a real martingale difference sequence with respect to the increasing* σ*-field* $\{\mathcal{F}_{nj}\}$ *with second moments. If, as* $n\to\infty$,

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

where σ^2 *is a positive constant, and, for each* $\varepsilon > 0$,

$$\sum_{j=1}^{r_n}\mathrm{E}(Y_{nj}^2 I_{(|Y_{nj}|\ge\varepsilon)}) \to 0,$$

then

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

The following lemma is the key tool for establishing the CLTs for linear statistics of eigenvectors of sample correlation matrix and sample covariance matrix when $\nu_4 \neq 0$. We believe this lemma will have independent interest. It indicates that when n is large, $\mathbf{A}_j^{-1}(z)$ is "close to" $\mathbb{R}_n(z)$ in a certain sense and we shall replace $\mathbf{A}_j^{-1}(z)$ with $\mathbb{R}_n(z)$ without changing the limit in probability under some conditions.

LEMMA 5.4. *For any nonrandom p-dimensional complex vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *with uniformly bounded Euclidean norms*, *under Assumptions* (C1)–(C4) *and* $|x_{1,1}| \le \eta_n n^{1/2}$ *with* $\eta_n \to 0$, *we have*

$$\mathbf{a}^*(\mathbf{B}_{j2}(z) + \mathbf{B}_{j3}(z) + \mathbf{B}_{j4}(z))\mathbf{b} \to 0$$

in probability.

PROOF. Recall that $\mathbf{x}_1, \dots, \mathbf{x}_n$ are i.i.d.. For convenience of notation, we will prove this lemma by removing $\mathbf{x}_j$ from the samples without loss of generality. We have

$$\begin{aligned} &\mathrm{E}\left|\sum_{i=1}^n \mathbf{a}^* b(z)\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2 \\ &\quad = \sum_{i=1}^n \mathrm{E}\left|\mathbf{a}^* b(z)\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2 \\ &\qquad + |b(z)|^2 \mathrm{E}\sum_{i\neq j} \mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_i^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_j^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}. \end{aligned} \tag{15}$$

It is easy to see that, from Lemma 5.1,

$$\begin{aligned} &\sum_{i=1}^n \mathrm{E}\left|\mathbf{a}^* b(z)\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2 \\ &\quad = \sum_{i=1}^n \mathrm{E}\left|\mathbf{r}_i^*\mathbf{A}_i^{-1}(z)\mathbf{b}\mathbf{a}^* b(z)\mathbb{R}_n(z)\mathbf{r}_i - \frac{1}{n}\mathbf{a}^* b(z)\mathbb{R}_n(z)\mathbf{R}_n\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2 = O(n^{-1}). \end{aligned} \tag{16}$$

Also, we have

$$\begin{aligned} &\left|\mathrm{E}\sum_{i\neq j} \mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_i^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_j^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right| \\ &\quad = \left|\mathrm{E}\sum_{i\neq j} \mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)(\mathbf{A}_i^{-1}(z) - \mathbf{A}_{ij}^{-1}(z))\right. \\ &\qquad \left.\times \mathbf{b}\mathbf{b}^*(\mathbf{A}_j^{-1}(\bar{z}) - \mathbf{A}_{ij}^{-1}(\bar{z}))\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right| \\ &\quad = \left|\mathrm{E}\sum_{i\neq j} \beta_{j(i)}(z)\beta_{i(j)}(\bar{z})\mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\right. \\ &\qquad \left.\times \mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right| \\ &\quad \le \sum_{i\neq j} \mathrm{E}^{1/2}|\beta_{j(i)}(z)\beta_{i(j)}(\bar{z})|^2 \end{aligned} \tag{17}$$

$$\times \mathrm{E}^{1/2}\left|\mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\right.$$
$$\left.\times \mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right|^2$$
$$\le C\sum_{i\ne j}\mathrm{E}^{1/2}\left|\mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\right.$$
$$\left.\times \mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right|^2.$$

Note that

$$\begin{aligned}
&\mathrm{E}\left|\mathbf{a}^*\mathbb{R}_n(z)\left(\mathbf{r}_i\mathbf{r}_i^* - \frac{1}{n}\mathbf{R}_n\right)\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\right.\\
&\qquad\left.\times \mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\right|^2\\
&\quad\le C\left\{\mathrm{E}\left|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\right.\right.\\
&\qquad\left.\times \mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(\bar{z})\mathbf{a}\mathbf{a}^*\mathbb{R}_n(z)\mathbf{r}_i\right|^2\\
&\qquad+\frac{1}{n^2}\mathrm{E}\left|\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\right.\\
&\qquad\left.\left.\times \mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\right|^2\right\}\\
&\quad\le C\left\{\mathrm{E}\left|\left(\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\right.\right.\right.\\
&\qquad\left.-\frac{1}{n}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\right)\\
&\qquad\left.\times \mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i\right|^2\\
&\qquad+\frac{1}{n^2}\mathrm{E}\left|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\right.\\
&\qquad\times \mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i - \frac{1}{n}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\\
&\qquad\left.\times \mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\right|^2\\
&\qquad+\frac{1}{n^2}\mathrm{E}\left|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\right.\\
&\qquad\times \mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\\
&\qquad-\frac{1}{n}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\\
&\qquad\left.\times \mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\right|^2
\end{aligned}\tag{18}$$

$$+\frac{1}{n^4}\mathrm{E}\bigg|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\mathbb{R}_n(\bar{z})$$
$$\times\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\bigg|^2$$
$$+\frac{1}{n^4}\mathrm{E}\bigg|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)$$
$$\times\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\bigg|^2\bigg\}.$$

We then deal with the five terms above.

First, by applying Lemma 5.1, we obtain that

$$\begin{aligned}
&\mathrm{E}\bigg|\bigg(\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i-\frac{1}{n}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\bigg)\\
&\qquad\times\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i\bigg|^2\\
&\le\mathrm{E}^{1/2}\bigg|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i-\frac{1}{n}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\bigg|^4\\
&\qquad\times\mathrm{E}^{1/2}\bigg|\mathbf{r}_j^*\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_j-\frac{1}{n}\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i\bigg|^4\\
&\le Cn^{-1}\mathrm{E}^{1/4}|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_j|^4\\
&\qquad\times\mathrm{E}^{1/4}|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j|^4\\
&\qquad\times\big(n^{-1}\mathrm{E}^{1/4}|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_i|^4\\
&\qquad\times\mathrm{E}^{1/4}|\mathbf{r}_i^*\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i|^4\big)\\
&=o(n^{-4}).
\end{aligned}\tag{19}$$

Second, also by applying Lemma 5.1 and taking the similar procedure we get

$$\begin{aligned}
&\frac{1}{n^2}\mathrm{E}\bigg|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{r}_i\\
&\qquad-\frac{1}{n}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\bigg|^2\\
&=o(n^{-4}).
\end{aligned}\tag{20}$$

For the same reason, we have

$$\begin{aligned}
&\frac{1}{n^2}\mathrm{E}\bigg|\mathbf{r}_i^*\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{r}_i\\
&\qquad-\frac{1}{n}\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\bigg(\mathbf{r}_j\mathbf{r}_j^*-\frac{1}{n}\mathbf{R}_n\bigg)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\bigg|^2\\
&=o(n^{-4}).
\end{aligned}\tag{21}$$

Furthermore, taking the same but more simple procedures, we have

$$\frac{1}{n^4}\mathrm{E}\left|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\right|^2 = o(n^{-4}), \tag{22}$$

and

$$\frac{1}{n^4}\mathrm{E}\left|\mathbf{r}_j^*\mathbf{A}_{ij}^{-1}(z)\mathbf{b}\mathbf{b}^*\mathbf{A}_{ij}^{-1}(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(\bar{z})\left(\mathbf{r}_j\mathbf{r}_j^* - \frac{1}{n}\mathbf{R}_n\right)\mathbb{R}_n(z)\mathbf{a}\mathbf{a}^*\mathbb{R}_n(\bar{z})\mathbf{R}_n\mathbf{A}_{ij}^{-1}(z)\mathbf{r}_j\right|^2 = o(n^{-4}). \tag{23}$$

Substituting (19)–(23) to (18) then to (17), combining with (15)–(16), we then get

$$\mathbf{a}^*\mathbf{B}_{j_2}(z)\mathbf{b} \xrightarrow{p} 0.$$

Then, we have

$$\begin{aligned}\mathrm{E}\left|\mathbf{a}^*\mathbf{B}_{j3}(z)\mathbf{b}\right| &\le C\sum_{i=1}^{n}\mathrm{E}^{1/2}\left|\beta_i(z)-b(z)\right|^2\mathrm{E}^{1/2}\left|\mathbf{a}^*\mathbb{R}_n(z)\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2\\ &\le \frac{C}{n}\sum_{i=1}^{n}\mathrm{E}^{1/2}\left|\beta_i(z)-b(z)\right|^2 = O(n^{-1/2})\end{aligned}$$

and

$$\mathrm{E}\left|\mathbf{a}^*\mathbf{B}_{j4}(z)\mathbf{b}\right| \le \frac{C}{n}\sum_{i=1}^{n}\mathrm{E}^{1/2}\left|\mathbf{a}^*\mathbb{R}_n(z)\mathbf{A}_i^{-1}(z)\mathbf{r}_i\mathbf{r}_i^*\mathbf{A}_i^{-1}(z)\mathbf{b}\right|^2 = O(n^{-1}).$$

Combining the argument above, we finally get that

$$\mathrm{E}\left|\mathbf{a}^*\left(\mathbf{B}_{j2}(z)+\mathbf{B}_{j3}(z)+\mathbf{B}_{j4}(z)\right)\mathbf{b}\right| = o(1). \tag{24}$$

The proof of the lemma is complete. □

The remaining lemmas all concern the properties of random diagonal matrices $\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}$, $\mathbb{D}_{\Xi_j} - \mathbf{I}$ and $\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}$. They are helpful in dealing with those terms that were caused by normalization using sample variances when we consider the properties of sample correlation matrices. In particular, we successfully relax the moment condition of some existing results for sample correlation matrices by virtue of Lemma 5.7.

LEMMA 5.5. *Assume that* $\mathrm{E}\,x_{1,1} = 0$, $\mathrm{E}\,|x_{1,1}|^2 = 1$, $\mathrm{E}\,|x_{1,1}|^4 < \infty$ *and* $|x_{1,1}| \le \eta_n\sqrt{n}$. *Then for any given* j *and* l, *we have* $\mathrm{E}\,\|\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\|^l \to 0$, $\mathrm{E}\,\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\|^l \to 0$, *as* $n \to \infty$.

PROOF. Note that $\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}$ is a diagonal matrix and its kth diagonal entry is $d_k = \sum_{j=1}^n(\frac{1}{n}\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - \frac{1}{n})$. Thus we have $\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\|^l = \max_{k=1}^p|\sum_{j=1}^n(\frac{1}{n}\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - \frac{1}{n})|^l$, where $\mathbf{g}_k^*$ is the kth row of the matrix $\mathbf{G}_n$. Denote $\triangle_{k,j} = (\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - 1) = (\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1)$, then from Lemma 5.1, we have that for $l \le 2$, $\mathrm{E}\,|\triangle_{i,j}|^l \le C$ and for $l > 2$, $\mathrm{E}\,|\triangle_{i,j}|^l \le \eta_n n^{l-2}$. Thus we obtain that for any even l,

$$\begin{aligned}E|d_k|^l &= \mathrm{E}\left(\frac{1}{n}\sum_{j=1}^n(\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - 1)\right)^l = \frac{1}{n^l}\sum_{j_1,j_2,\dots,j_l=1}^{n}\mathrm{E}\prod_{t=1}^{l}(\mathbf{g}_k^*\mathbf{x}_{j_t}\mathbf{x}_{j_t}^*\mathbf{g}_k - 1)\\ &\le n^{-l}\big(C_{l,1}n^{l/2}(\mathrm{E}\,|\triangle_{k,j}|^2)^{l/2} + C_{l,2}n^{l/2-1}(\mathrm{E}\,|\triangle_{k,j}|^2)^{l/2-3}(\mathrm{E}\,|\triangle_{k,j}|^3)^2\\ &\quad + \cdots + n\,\mathrm{E}\,|\triangle_{k,j}|^l)\big) = o(n^{-1}),\end{aligned} \tag{25}$$

where for any $j = 1, \dots, C_{l,j}$ stands for a constant only depending on l. In the last step above we used the fact that the number of terms is bounded by a constant depend on l.

For l odd, the above results also hold due to the fact $\mathrm{E}|d_k|^{l-1} \le (\mathrm{E}|d_{i,i}|^l)^{(l-1)/l}$. Thus we arrive at $\mathrm{E}\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\|^l \to 0$ as $n \to \infty$ since the above result is uniform for all $1 \le k \le p$.

For the same reason, we have $\mathrm{E}\|\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\|^l \to 0$.

We hence complete the proof of this lemma. $\square$

LEMMA 5.6. *Assume that* $\mathrm{E}\, x_{1,1} = 0$, $\mathrm{E}|x_{1,1}|^2 = 1$ *and* $|x_{1,1}| \le \eta_n n^{1/4}$ *with* $\eta_n \to 0$. *For any given j and l, as $n \to \infty$, we have*:

(1): $\mathrm{E}|\widehat{\gamma}_j(z)|^l = O(n^{-l/2})$,
(2): $\mathrm{E}|\mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j|^l = O(n^{\max\{-l/2-1,-l\}})$,
(3): $\mathrm{E}|\mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\pi_n\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j|^l = o(n^{\max\{-l/2-1,-l\}})$,
(4): $\mathrm{E}\|(\mathbb{D}_{\boldsymbol{\Xi}_j} - \mathbf{I})\|^l = o(n^{l/2-1})$ *when* $l > 2$ *and* $\mathrm{E}\|(\mathbb{D}_{\boldsymbol{\Xi}_j} - \mathbf{I})\|^l = O(1)$ *when* $l \le 2$.

PROOF. By Lemma 5.1 and Lemma 5.5, we have that

$$
\begin{aligned}
&\mathrm{E}\big|\mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\pi_n\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j\big|^l \\
&\quad = \mathrm{E}\bigg|\bigg(\mathbf{r}_j^* \mathbf{A}_j^{-1}(z)\pi_n\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j - \frac{1}{n}\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{R}_n\mathbf{A}_j^{-1}(z)\pi_n\bigg) \\
&\qquad + \frac{1}{n}\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{R}_n\mathbf{A}_j^{-1}(z)\pi_n\bigg|^l \\
&\quad \le Cn^{-l}\,\mathrm{E}\big(\big|\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{R}_n\mathbf{A}_j^{-1}(\bar{z})\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\pi_n\big|^{l/2} \\
&\qquad + \big(\eta_n n^{1/4}\big)^{\max\{(2l-4),0\}}\big|\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{R}_n\mathbf{A}_j^{-1}(\bar{z})\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\pi_n\big|^{l/2}\big) \\
&\quad \le Cn^{-l}n^{\max\{l/2-1,0\}}\,\mathrm{E}\big\|\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\big\|^l = o\big(n^{\max\{-l/2-1,-l\}}\big).
\end{aligned} \tag{26}
$$

This proves (3). The proofs of (1), (2) and (4) are the same thus omitted. $\square$

LEMMA 5.7. *Under Assumptions* (C1)–(C3), *we have*

$$\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\| = o(1), \quad a.s.$$

and

$$\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\| = O_p\big(n^{-1/2}\big).$$

PROOF. Recall that $\boldsymbol{\Xi} = \frac{1}{n}\mathbf{G}\mathbf{X}\mathbf{X}^*\mathbf{G}^*$. Note that the kth diagonal entry of $\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}$ is $d_k = \sum_{j=1}^n (\frac{1}{n}\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - \frac{1}{n})$. Denote $\triangle_{k,j} = (\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - 1) = (\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1)$, we know that for any given k, $\{\triangle_{k,1}, \dots, \triangle_{k,n}\}$ is a sequence of i.i.d. random variables. It is easy to see that $\mathrm{E}|\triangle_{k,1}|^2 \le C < \infty$ and $\mathrm{E}|\triangle_{i,j}|^l \le (\eta_n n)^{l-2}$ for $l > 2$. Thus the second conclusion is an easy consequence of (25) by choosing $l = 2$.

We now turn to the first conclusion. To obtain the a.s. convergence under the fourth moment condition, we need to do more than what we did in the proof of Lemma 5.5. Under Assumptions (C1)–(C3), following a standard technical, we will truncate the variable at $\eta_n\sqrt{n}$, where η_n is a sequence converging to 0. This part is postponed to the Appendix.

After the truncation and recentralization step, choose l to be an even number in the region $[\log(n\eta_n^4/(\nu_4+3)), \log(n\eta_n^4/(\nu_4+3))]$. Now, since l tends to infinity, we need to reconsider the bounds on high order moments of $\triangle_{k,j_t}$. First, it is easy to verify from 5.2 that $\mathrm{E}|\triangle_{k,j_t}|^2 \le$

ν_4+2. Then, noting that the matrix $\mathbf{g}_k\mathbf{g}_k^*$ is of rank 1, by the proof of Lemma 9.1 in Bai and Silverstein (2010), we know that for large n,

$$\mathrm{E}\,|\triangle_{k,j_t}|^{\ell} \le (\nu_4+3)80^l n^{\ell-2}\eta_n^{2(\ell-2)} \le 6400(\nu_4+3)(80\eta_n)^{\ell-2}(\eta_n n)^{\ell-2} \le (\eta_n n)^{\ell-2}$$

when $2<\ell\le\sqrt{2}\log(n\eta_n^4/(\nu_4+3))$. Recall that

$$E|d_k|^l = \mathrm{E}\left(\frac{1}{n}\sum_{j=1}^{n}(\mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k-1)\right)^l = \frac{1}{n^l}\sum_{j_1,j_2,\dots,j_l=1}^{n}\mathrm{E}\prod_{t=1}^{l}\triangle_{k,j_t}. \tag{27}$$

It is easy to see by the independence of $\mathbf{x}_1,\dots,\mathbf{x}_n$ that if there is a single j in the sequence $(j_1,\dots,j_l)$ then the expectation of the term $\prod_{t=1}^{l}\triangle_{k,j_t}$ is zero. Thus we only need to consider the terms that have contribution to the moment, that is, those terms where there is no single j in the sequence $(j_1,\dots,j_l)$. To achieve this, let us denote m as the number of dimensional j's in the sequence $(j_1,\dots,j_l)$, denote the noncoincident j's as $\mathring{\mathbb{j}}_1,\dots,\mathring{\mathbb{j}}_m$ (index are set according to the order of appearance) respectively, also denote q_ζ for $1\le\zeta\le m$ as the number of the multiplicity of j_ζ. Immediate results are

$$m\le l/2,\qquad \sum_{\zeta=1}^{m}q_\zeta = l.$$

Now, to find the bound of (27), we need the following steps.

(1): Set the arrangement of different $\mathring{\mathbb{j}}_1,\dots,\mathring{\mathbb{j}}_m$ and bound the number of different arrangements.

Since the first appearance of $\mathring{\mathbb{j}}_1$ must take place at the first place of sequence $j_1,\dots,j_l$, the number of different arrangements of the first appearance of $\mathring{\mathbb{j}}_1,\dots,\mathring{\mathbb{j}}_m$ is $\binom{l-1}{m-1}$. Then the number of different arrangements of the second appearances of $\mathring{\mathbb{j}}_1,\dots,\mathring{\mathbb{j}}_m$ can be bounded by $\binom{l-m}{m}$. All other arrangements can obviously be bounded by m^{l-2m}.

(2): Bound the expectation by the given arrangements.

We have

$$\mathrm{E}\prod_{\zeta=1}^{m}(\triangle_{k,\mathring{\mathbb{j}}_\zeta})^{q_\zeta} \le \prod_{\zeta=1}^{m}\mathrm{E}\,|\triangle_{k,\mathring{\mathbb{j}}_\zeta}|^{q_\zeta} \le (\nu_4+2)^m\prod_{\zeta=1}^{m}(\eta_n n)^{q_\zeta-2} = (\nu_4+2)^m(\eta_n n)^{l-2m}. \tag{28}$$

Now, combining the arguments above, we finally arrive at when n large and $\frac{4\sqrt{\eta_n}l}{\log n}\to 0$,

$$\begin{aligned}
E|d_k|^l &\le \frac{1}{n^l}\sum_{m=1}^{l/2}n^m\binom{l-1}{m-1}\binom{l-m}{m}m^{l-2m}(\nu_4+2)^m(\eta_n n)^{l-2m}\\
&\le \eta_n^l\sum_{m=1}^{l/2}n^m l^m l^m m^{l-2m}(\nu_4+2)^m(\eta_n n)^{-2m} \le \eta_n^l\sum_{m=1}^{l/2}l^{2m}m^{l-2m}n^{-m/2}\\
&\le \eta_n^l\sum_{m=1}^{l/2}(\ln^{-1/8})^{2m}m^{l-2m}n^{-m/4} \le \eta_n^l\sum_{m=1}^{l/2}(m+1)^l(n^{1/4})^{-m}\\
&\le \eta_n^l\sum_{m=1}^{l/2}n^{1/4}\left(\frac{l}{\log n^{1/4}}\right)^l \le (\eta_n)^{l/2}.
\end{aligned} \tag{29}$$

Here, in the second to last inequality we use the fact that

$$a^{-c}(c+1)^b \le a\left(\frac{b}{\log a}\right)^b$$

for $a > 1$, $b > 0$, $c > 0$. Thus we have for any $\varepsilon > 0$, and any t, as n large enough,

$$\mathbf{P}\big(\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\| > \varepsilon\big) \le (\varepsilon\sqrt{\eta_n})^l \le n^{-t}. \tag{30}$$

The proof of this lemma is thus complete. □

Next we will proceed with our proofs of the theorems.

5.2. *Proof of Theorem* 3.1. Under (C1)–(C4), we have from Lemma 5.7 that $\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\| \to 0$ a.s. This implies that $\|\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \mathbf{I}\| \to 0$ almost surely. We thus have

$$\begin{aligned}
&\|\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2}\boldsymbol{\Xi}_n\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \boldsymbol{\Xi}_n\| \\
&\quad = \|\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2}\boldsymbol{\Xi}_n\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \boldsymbol{\Xi}_n\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} + \boldsymbol{\Xi}_n\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \boldsymbol{\Xi}_n\| \\
&\quad \le \|(\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \mathbf{I})\boldsymbol{\Xi}_n\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2}\| + \|\boldsymbol{\Xi}_n(\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - \mathbf{I})\| \to 0, \quad \text{a.s.}
\end{aligned} \tag{31}$$

Here we used the fact that $\|n^{-1}\mathbf{X}_n\mathbf{X}_n^T\| \to (1+\sqrt{\rho})^2$, a.s. as $n \to \infty$, which implies $\|\boldsymbol{\Xi}_n\| \le C\|\mathbf{R}_n\|(1+\sqrt{\rho})^2$, $a.s$. Then the first result of this theorem follows from the well-known Weyl inequality.

We now treat the second and the last results. When $\mathbf{G}_n = \mathbf{I}$, note that

$$\begin{aligned}
&\mathrm{P}\big(\lambda_{\max}(\widehat{\mathbf{R}}_n) > (1+\sqrt{\rho})^2 + \varepsilon\big) \\
&\quad \le \mathrm{P}\big(\lambda_{\max}(n^{-1}\mathbf{X}_n\mathbf{X}_n^T) > (1+\sqrt{\rho})^2 + \varepsilon\big) + \mathrm{P}\big(\lambda_{\max}(\mathbb{D}_{\boldsymbol{\Xi}_n}) > 1+\varepsilon\big),
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{P}\big(\lambda_{\min}(\widehat{\mathbf{R}}_n) < I_{(0,1)}(\rho)(1-\sqrt{\rho})^2 - \varepsilon\big) \\
&\quad \le \mathrm{P}\big(\lambda_{\min}(n^{-1}\mathbf{X}_n\mathbf{X}_n^T) < I_{(0,1)}(\rho)(1-\sqrt{\rho})^2 - \varepsilon\big) + \mathrm{P}\big(\lambda_{\min}(\mathbb{D}_{\boldsymbol{\Xi}_n}) < 1-\varepsilon\big),
\end{aligned}$$

the last two results follows from Theorem 9.13 in Bai and Silverstein (2010) and (30).

5.3. *Proof of Theorem* 3.2. Theorem 1 in Bai and Silverstein (1998) considers the corresponding result for sample covariance matrices. Adapting this result to $\boldsymbol{\Xi}_n$, combining with (31) and Weyl's inequality, we get the desired result.

5.4. *Proof of Theorem* 3.3. First, by Lemma 5.7, it is easy to see that

$$\begin{aligned}
&\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n - \pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n \\
&\quad = \pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}(\boldsymbol{\Xi}_n - \widehat{\mathbf{R}}_n)(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n = o_{\text{a.s.}}(1).
\end{aligned} \tag{32}$$

In Bai, Miao and Pan (2007), it is proved that

$$\pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n - \big(-z^{-1}\pi_n^*(1+\underline{s}(z)\mathbf{R}_n)^{-1}\pi_n\big) = o_{\text{a.s.}}(1).$$

Then the first conclusion of this theorem is the consequence of the fact $\underline{s}_{F^{\rho_n,H_n}}(z) - \underline{s}(z) = o(1)$. And the second conclusion is from the fact that when the condition ($\mathcal{I}$C1) is satisfied, we have

$$-z^{-1}\pi_n^*(1+\underline{s}(z)\mathbf{R}_n)^{-1}\pi_n \to \int \frac{1}{-z(\underline{s}(z)t+1)}\,dH(t) = s(z).$$

5.5. *Proofs of Theorem 3.4, Corollary 3.1 and Corollary 3.2.* This subsection is devoted to dealing with the proofs of Theorem 3.4, Corollary 3.1 and Corollary 3.2. Notice that these two corollaries are both special cases of Theorem 3.4. Their proofs will be involved in the proof of Theorem 3.4 and the differences only appear in Section 5.5.5.

Note that for any cumulative distribution function G and any function g that is analytic on an open set containing the support of G, we have by the Cauchy integral formula that

$$\int g(x)\,dG(x) = -\frac{1}{2\pi \mathrm{i}}\int_{\mathcal{C}} g(z)s_G(z)\,dz,$$

where $\mathcal{C}$ is a contour properly chosen as given in Section 5.5.1. Following the idea of Bai and Silverstein (2004), the proof of Theorem 3.4 is based on analyzing the sequence of random processes defined by

$$\mathbb{M}_n(z) = \sqrt{n}\big(s_{F_{v,\pi_n}^{\hat{\mathbf{R}}_n}} - s_{\rho_n,\pi_n}^{\mathbf{R}_n}\big) = \sqrt{n}\big(\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n - s_{\rho_n,\pi_n}^{\mathbf{R}_n}\big).$$

The sketch of the proof is as follows.

- Step 1: To handle the high order moments, we first truncate the underlying variable at a proper order of n.
- Step 2: In Section 5.5.2, we decompose $\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n$ into several parts.
- Step 3: We find in Section 5.5.3 the main terms that contribute to the limiting property of the random process

$$\mathbb{M}_{n,1}(z) = \sqrt{n}\big(\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n - \mathrm{E}\,\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n\big).$$

- Step 4: In Sections 5.5.4–5.5.6, we obtain the CLT for $\mathbb{M}_{n,1}(z)$ by applying Lemma 5.3 and prove the tightness.
- Step 5: It is proved in Section 5.5.7 that

$$\mathbb{M}_{n,2}(z) = \sqrt{n}\big(\mathrm{E}\,\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n - s_{\rho_n,\pi_n}^{\mathbf{R}_n}\big)$$

  convergences to 0.

Combining the above Steps, we complete the proof of Theorem 3.4, Corollary 3.1 and Corollary 3.2.

The truncation step is essentially the same as Section 7 in Bai, Miao and Pan (2007) by making slight adjustments and thus omitted to reduce the length of the paper. After truncation, recentralization and rescale, we will proceed with our proof by assuming $\mathrm{E}\,x_{1,1} = 0$, $\mathrm{E}\,|x_{1,1}|^2 = 1$, $|x_{1,1}| \le \eta_n n^{1/4}$ and $\mathrm{E}\,|x_{1,1}|^4 < \infty$. We now proceed the remaining steps one by one.

5.5.1. *The choosing of contour $\mathcal{C}$.* We define the contour $\mathcal{C}$ as

$$\mathcal{C} = \mathcal{C}_\ell \cup \mathcal{C}_u \cup \mathcal{C}_b \cup \mathcal{C}_r,$$

where

$$\mathcal{C}_u = \{x + iv_0 : x \in [x_\ell, x_r]\}, \qquad \mathcal{C}_\ell = \{x_\ell + iv : |v| \le v_0\},$$
$$\mathcal{C}_b = \{x - iv_0 : x \in [x_\ell, x_r]\}, \qquad \mathcal{C}_r = \{x_r + iv : |v| \le v_0\},$$

where x_r is a number which is greater than the right endpoint of interval (6) and x_ℓ is a negative number if the left endpoint of interval ((6)) is zero, otherwise x_ℓ is a positive number smaller than the left endpoint of interval (6), and $v_0 > 0$ is to be determined. Let $\mathcal{C}_n = \mathcal{C} \cap \{z : |\Im z| > n^{-2}\}$. Define $\mathcal{B}_n = \{(\lambda_{\min}(\hat{\mathbf{R}}_n) < x_\ell) \cup (\lambda_{\max}(\hat{\mathbf{R}}_n) > x_r)\}$. By Theorem 3.1, we have that $P(\mathcal{B}_n) = o(n^{-t})$ for any given $t > 0$.

Let

$$\widehat{M}_n(z) = \begin{cases} M_n(z) & \text{if } z \in \mathcal{C}_n, \\ M_n(x_\ell + in^{-2}) & \text{if } \Re z = x_\ell, \Im z \in [0, n^{-2}], \\ M_n(x_\ell - in^{-2}) & \text{if } \Re z = x_\ell, \Im z \in [-n^{-2}, 0), \\ M_n(x_r + in^{-2}) & \text{if } \Re z = x_r, \Im z \in [0, n^{-2}], \\ M_n(x_r - in^{-2}) & \text{if } \Re z = x_r, \Im z \in [-n^{-2}, 0). \end{cases}$$

Observe that on event $\mathcal{B}_n$, when $\Re z$ equals either x_ℓ or x_r, we have $|M_n(z)| \le 1/\varepsilon$, hence

$$\left| p \oint_{\mathcal{C}} g(z)\big(M_n(z) - \widehat{M}_n(z)\big)\, dz \right| = \left| p \oint_{\mathcal{C}\setminus\mathcal{C}_n} g(z)\big(M_n(z) - \widehat{M}_n(z)\big)\, dz \right| \le K \frac{p}{n^2} \cdot 1/\varepsilon = o(1).$$

Therefore, in order to establish the limit theorem for $p \oint_{\mathcal{C}} g(z) M_n(z)\, dz$, it suffices to study $p \oint_{\mathcal{C}} g(z) \widehat{M}_n(z)\, dz$. Furthermore, since $\Im(z)$ can be chosen to be arbitrarily small, the contribution from the segments C_ℓ and C_r can be made small as well. This allows us to focus only on $z \in \mathcal{C}_u \cup \mathcal{C}_b$ in the following.

5.5.2. *Decomposition of* $\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n$. We here want to decompose the main object $\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n$ into several parts.

Using the formula for any matrix $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}$, and Lemma 5.7 we have

$$\begin{aligned}
&\pi_n^*(\widehat{\mathbf{R}}_n - z\mathbf{I})^{-1}\pi_n \\
&\quad = \pi_n^*\big(\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2}\boldsymbol{\Xi}\mathbb{D}_{\boldsymbol{\Xi}_n}^{-1/2} - z\mathbf{I}\big)^{-1}\pi_n = \pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbb{D}_{\boldsymbol{\Xi}_n})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n \\
&\quad = \pi_n^*\big(\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2} + z\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbb{D}_{\boldsymbol{\Xi}_n})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\big)\pi_n \\
&\quad = \pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n + z\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n \\
&\qquad + z^2\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbb{D}_{\boldsymbol{\Xi}_n})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n \\
&\quad = \pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n + z\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n \\
&\qquad + z^2\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2}\pi_n \\
&\qquad + o\big(\|\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}\|^2\big).
\end{aligned} \tag{33}$$

Then, noting that

$$\mathbb{D}_{\boldsymbol{\Xi}_n}^{1/2} = \mathbf{I} + 2^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I}) - 8^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})^2 + o\big((\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})^2\big)$$

when $e_j'(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})e_j = o(1)$ for any $j = 1, \dots, p$, we obtain that

$$\begin{aligned}
&\pi_n^*(\widehat{\mathbf{R}} - z\mathbf{I})^{-1}\pi_n \\
&\quad = \pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n + 2^{-1}\pi_n^*(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n \\
&\qquad + 2^{-1}\pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})\pi_n + z\pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n \\
&\qquad + 4^{-1}\pi_n^*(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})\pi_n \\
&\qquad - 8^{-1}\pi_n^*(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})^2\pi_n - 8^{-1}\pi_n^*(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})^2(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n \\
&\qquad + 2^{-1}z\pi_n^*(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n} - \mathbf{I})(\boldsymbol{\Xi}_n - z\mathbf{I})^{-1}\pi_n
\end{aligned}$$

$$
\begin{aligned}
&+2^{-1}z\pi_n^*(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\Xi_n}-\mathbf{I})(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\Xi_n}-\mathbf{I})\pi_n\\
&+z^2\pi_n^*(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\Xi_n}-\mathbf{I})(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\Xi_n}-\mathbf{I})(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}\pi_n\\
&+o(\|\mathbb{D}_{\Xi_n}-\mathbf{I}\|^2).\\
:=\ &\Phi_0+2^{-1}\Phi_1+2^{-1}\Phi_2+z\Phi_3+4^{-1}\Phi_4-8^{-1}\Phi_5-8^{-1}\Phi_6\\
&+2^{-1}z\Phi_7+2^{-1}z\Phi_8+z^2\Phi_9+o(\|\mathbb{D}_{\Xi_n}-\mathbf{I}\|^2).
\end{aligned}
$$

It turns out that the first term Φ_0 is exactly the random process determined by sample covariance matrix $\boldsymbol{\Xi}_n$, which had been fully considered in the investigation of the properties of the eigenvector of the sample covariance matrix. While the other additional terms in the sample correlation case are all caused by the normalization using sample variances. Fortunately, under the existence of the fourth moment of the underlying distribution, it is proved that most of the additional terms have no contributions to the asymptotic properties.

5.5.3. *Simplify the terms to their main parts.* Note that according to Lemma 5.7, there are only four terms, say $\Phi_0,\dots,\Phi_3$, that have contribution to the limit of $\mathbb{M}_n(z)$. In this subsection, we will simplify the terms $\Phi_0,\dots,\Phi_3$ by splitting them into several terms and determine the corresponding main terms and remove all other terms that convergence to 0 in probability. We will proceed one by one.

5.5.3.1. *Simplification of* Φ_0. The limit properties of $\Phi_0=\pi_n^*(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}\pi_n$ has been investigated in Bai, Miao and Pan (2007) and the main term is proved to be

$$\phi_{0,j}(z)=-\sqrt{n}z\underline{s}(z)\,\mathrm{E}_j\,\delta_j(z).$$

5.5.3.2. *Simplification of* Φ_1 *and* Φ_2. Now we consider the terms Φ_1 and Φ_2. By (11) we have

$$
\begin{aligned}
&\sqrt{n}(\Phi_1-\mathrm{E}\,\Phi_1)\\
&\quad=\sqrt{n}\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\pi_n^*(\mathbb{D}_{\Xi_n}-\mathbf{I})\mathbf{A}^{-1}(z)\pi_n\\
&\quad=\sqrt{n}\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\big(\pi_n^*(\mathbb{D}_{\Xi_n}-\mathbf{I})\mathbf{A}^{-1}(z)\pi_n-\pi_n^*\mathbb{D}_{\Xi_n}^{(-j)}\mathbf{A}_j^{-1}(z)\pi_n\big)\\
&\quad=\frac{1}{\sqrt{n}}\sum_{j=1}^n\mathrm{E}_j\,\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z)\pi_n \qquad (34)\\
&\qquad-\sqrt{n}\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j(z)\pi_n^*\mathbb{D}_{\Xi_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\\
&\qquad-\frac{1}{\sqrt{n}}\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j(z)\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\\
&\quad:=\sum_{j=1}^n\phi_{1,j}(z)+\Delta_{1,1}+\Delta_{1,2}.
\end{aligned}
$$

Applying Lemma 5.5 and 5.6, we obtain

$$
\begin{aligned}
&\mathrm{E}\left|-\sqrt{n}\sum_{j=1}^{n}(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j(z)\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad\le n\sum_{j=1}^{n}\mathrm{E}\left|\beta_j(z)\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad\le Cn\sum_{j=1}^{n}\mathrm{E}\left|\pi_n^*\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad=o(1),
\end{aligned}
\tag{35}
$$

and

$$
\begin{aligned}
&\mathrm{E}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j(z)\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad\le Cn^{-1}\sum_{j=1}^{n}\mathrm{E}\left|\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad\le Cn^{-1}\sum_{j=1}^{n}\sqrt{\mathrm{E}\left\|(\mathbb{D}_{\Xi_j}-\mathbf{I})\right\|^4}\sqrt{\mathrm{E}\left|\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^4}\\
&\quad=O(n^{-1}),
\end{aligned}
\tag{36}
$$

which implies that both $\Delta_{1,1}$ and $\Delta_{1,2}$ converge to 0 in probability.

Following the same procedures, we can obtain that

$$
\sqrt{n}(\Phi_2-\mathrm{E}\,\Phi_2)=\sum_{j=1}^{n}\phi_{2,j}(z)+o_p(1),
$$

where $\phi_{2,j}(z)=\frac{1}{\sqrt{n}}\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_j}-\mathbf{I})\pi_n$.

5.5.3.3. *Simplification of* Φ_3. Now we turn to Φ_3. By equation (11), we have

$$
\begin{aligned}
&\sqrt{n}\big(\pi_n^*(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}\pi_n\\
&\qquad-\mathrm{E}\,\pi_n^*(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})(\boldsymbol{\Xi}_n-z\mathbf{I})^{-1}\pi_n\big)\\
&\quad=\sqrt{n}\sum_{j=1}^{n}(\mathrm{E}_j-\mathrm{E}_{j-1})\big(\pi_n^*\mathbf{A}^{-1}(z)(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})\mathbf{A}^{-1}(z)\pi_n\\
&\qquad-\pi_n^*\mathbf{A}_j^{-1}(z)\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\pi_n\big)\\
&\quad=\sqrt{n}\sum_{j=1}^{n}(\mathrm{E}_j-\mathrm{E}_{j-1})\big(\pi_n^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})\mathbf{A}_j^{-1}(z)\pi_n\\
&\qquad-\pi_n^*\mathbf{A}_j^{-1}(z)\mathbb{D}_{\boldsymbol{\Xi}_n}^{(-j)}\mathbf{A}_j^{-1}(z)\pi_n\\
&\qquad-\beta_j(z)\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})\mathbf{A}_j^{-1}(z)\pi_n\\
&\qquad-\beta_j(z)\pi_n^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\boldsymbol{\Xi}_n}-\mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n
\end{aligned}
\tag{37}
$$

$$
\begin{aligned}
&\quad + \beta_j(z)^2 \pi_n^* \mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\big)\\
&= \sqrt{n}\sum_{j=1}^n (\mathrm{E}_j - \mathrm{E}_{j-1})\big(\pi_n^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_j} - \mathbf{I})\mathbf{A}_j^{-1}(z)\pi_n\\
&\quad - \beta_j(z)\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\pi_n\\
&\quad - \beta_j(z)\pi_n^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\\
&\quad + \beta_j(z)^2 \pi_n^* \mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\big)\\
&:= \sum_{j=1}^n \phi_{3,j}(z) + \Delta_{3,1} + \Delta_{3,2} + \Delta_{3,3}.
\end{aligned}
$$

Following the same procedure as we did in the last paragraph, we can prove that both $\Delta_{3,1}$ and $\Delta_{3,2}$ are $o_p(1)$.

Then by Lemmas 5.6 and 5.5, we have

$$
\begin{aligned}
&\mathrm{E}\left|\sqrt{n}\sum_{j=1}^n (\mathrm{E}_j - \mathrm{E}_{j-1})\beta_j(z)^2\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\right|^2\\
&\quad \le Cn\sum_{j=1}^n \sqrt{\mathrm{E}\big|\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}_j^{-1}(z)\mathbf{r}_j\big|^4}\sqrt{\mathrm{E}\big|\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z)\mathbf{r}_j\big|^4}\\
&\quad = o(1).
\end{aligned}
\tag{38}
$$

Thus we have $\Delta_{3,3}$ convergence in probability to 0.

Now, we consider the limiting properties of the process

$$
\Upsilon(z) := \sum_{j=1}^n \big(\phi_{0,j}(z) + \phi_{1,j}(z) + \phi_{2,j}(z) + \phi_{3,j}(z)\big).
$$

To achieve this, we will draw support from the CLT for martingale, say Lemma 5.3.

5.5.4. *Verify the Lindeberg condition.* For any $z_1, \dots, z_r \in \mathbf{C}_+$, $\alpha_1, \dots, \alpha_r \in \mathbf{C}$ and any $\varepsilon > 0$, we have

$$
\begin{aligned}
&\sum_{j=1}^n E\left(\left|\sum_{\ell=1}^r \alpha_l\phi_{1,j}(z_\ell)\right|^2 I\left(\left|\sum_{\ell=1}^r \alpha_l\phi_{1,j}(z_\ell)\right| \ge \varepsilon\right)\right)\\
&\qquad \le \frac{1}{\varepsilon^2}\sum_{j=1}^n E\left|\sum_{\ell=1}^r \alpha_l\phi_{1,j}(z_\ell)\right|^4\\
&\qquad = \frac{1}{\varepsilon^2}\sum_{j=1}^n E\left|\sum_{\ell=1}^r \alpha_\ell E_j\frac{1}{\sqrt{n}}\pi_n^*(\mathbb{D}_{\Xi_j} - \mathbf{I})\mathbf{A}_j^{-1}(z_l)\pi_n\right|^4 = o(1).
\end{aligned}
\tag{39}
$$

For the same reason, we have

$$
\begin{aligned}
&\sum_{j=1}^n E\left(\left|\sum_{\ell=1}^r \alpha_l\phi_{2,j}(z_\ell)\right|^2 I\left(\left|\sum_{\ell=1}^r \alpha_l\phi_{2,j}(z_\ell)\right| \ge \varepsilon\right)\right)\\
&\qquad \le \frac{1}{\varepsilon^2}\sum_{j=1}^n E\left|\sum_{\ell=1}^r \alpha_\ell E_j\frac{1}{\sqrt{n}}\pi_n^*\mathbf{A}_j^{-1}(z_l)(\mathbb{D}_{\Xi_j} - \mathbf{I})\pi_n\right|^4 = o(1),
\end{aligned}
\tag{40}
$$

and

$$\sum_{j=1}^{n} E\left(\left|\sum_{\ell=1}^{r}\alpha_l\phi_{3,j}(z_\ell)\right|^2 I\left(\left|\sum_{\ell=1}^{r}\alpha_l\phi_{3,j}(z_\ell)\right|\geq\varepsilon\right)\right)$$

$$\leq\frac{1}{\varepsilon^2}\sum_{j=1}^{n}E\left|\sum_{\ell=1}^{r}\alpha_\ell E_j\frac{1}{\sqrt{n}}\pi_n^*\mathbf{A}_j^{-1}(z_l)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_l)\pi_n\right|^4=o(1). \tag{41}$$

Thus the second condition in Lemma 5.3 is satisfied.

5.5.5. *Convergence in finite dimensions.* We now focus on determining the variance-covariance function. In the following, we use "$\stackrel{\mathcal{I}}{=}$" to denote equal under the isotropic conditions (corresponds to Corollary 3.1) and use "$\stackrel{\mathcal{N}}{=}$" to denote equal under the null case $\mathbf{R}_n=\mathbf{G}_n=\mathbf{I}_p$ (corresponds to Corollary 3.2). We will frequently use the relationship

$$\frac{1}{z(1+\underline{s}(z))}=-s(z),$$

and

$$\mathcal{L}_\pi=\mathcal{L}_{\pi,\mathbf{G}}=\mathcal{L}_{\pi,\mathbf{G}}^{(1)}=\mathcal{L}_{\pi,\mathbf{R}}^{(2)}=\mathcal{L}_{\pi,\mathbf{G}}^{(3)}=\lim_{n\to\infty}\sum_{k=1}^{n}|\pi_{(k)}|^4,$$

under the null case.

We have from Lemma 5.2, Lemma 5.4, and (4.7)–(4.25) in Bai, Miao and Pan (2007) that

$$\begin{aligned}
&\sum_{j=1}^{n}\mathrm{E}_{j-1}\,\phi_{0,j}(z_1)\phi_{0,j}(z_2)\\
&=\frac{z_1z_2\underline{s}(z_1)\underline{s}(z_2)}{n}\sum_{j=1}^{n}\mathrm{E}_{j-1}\big(\mathrm{E}_j\,\delta_j(z_1)\,\mathrm{E}_j\,\delta_j(z_2)\big)\\
&=\frac{z_1z_2\underline{s}(z_1)\underline{s}(z_2)}{n}\\
&\quad\times\sum_{j=1}^{n}\Bigg(\nu_4\sum_{k=1}^{p}e_k'\mathbf{G}_n^*\mathbb{R}_n(z_1)\pi_n\pi_n^*\mathbb{R}_n(z_1)\mathbf{G}_ne_ke_k'\mathbf{G}_n^*\mathbb{R}_n(z_2)\pi_n\pi_n^*\mathbb{R}_n(z_2)\mathbf{G}_ne_k\\
&\quad+\alpha\,\mathrm{E}_{j-1}\operatorname{tr}\mathrm{E}_j\,\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{R}_n\,\mathrm{E}_j\,\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)\mathbf{R}_n\Bigg)+o_p(1)\\
&\to\frac{\nu_4\underline{s}(z_1)\underline{s}(z_2)}{z_1z_2}\\
&\quad\times\lim_{n\to\infty}\sum_{k=1}^{p}e_k'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_ke_k'\mathbf{G}_n^*\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)\mathbf{G}_ne_k\\
&\quad+\frac{\alpha(\underline{s}(z_2)-\underline{s}(z_1))}{z_1z_2(z_2-z_1)}\lim_{n\to\infty}\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_n\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)\mathbf{R}_n\mathcal{R}_n(z_1)\pi_n\\
&\stackrel{\mathcal{I}}{=}\lim_{n\to\infty}\left(\nu_4z_1z_2s^2(z_1)s^2(z_2)\underline{s}(z_1)\underline{s}(z_2)\mathcal{L}_{\pi,\mathbf{G}}^{(1)}+\frac{\alpha(z_2\underline{s}(z_2)-z_1\underline{s}(z_1))^2}{\rho^2z_1z_2(z_2-z_1)(\underline{s}(z_2)-\underline{s}(z_1))}\right)\\
&\stackrel{\mathcal{N}}{=}\frac{\nu_4\underline{s}(z_1)\underline{s}(z_2)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))^2}+\frac{\alpha(z_2\underline{s}(z_2)-z_1\underline{s}(z_1))^2}{\rho^2z_1z_2(z_2-z_1)(\underline{s}(z_2)-\underline{s}(z_1))}.
\end{aligned}\tag{42}$$

Also, we can obtain that

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\,\phi_{1,j}(z_1)\phi_{1,j}(z_2) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}_{j-1}\bigl(\mathrm{E}_j\,\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_1)\pi_n\,\mathrm{E}_j\,\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n\bigr) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathrm{E}_j\, e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^* e_k\,\mathrm{E}_j\, e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^* e_l \\
&\qquad \times \mathrm{E}\bigl[\bigl(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j-1\bigr)\bigl(\mathbf{x}_j^*\mathbf{g}_l\mathbf{g}_l^*\mathbf{x}_j-1\bigr)\bigr] \\
&\quad = \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)\,\mathrm{E}_j\, e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^* e_k\,\mathrm{E}_j\, e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^* e_l \\
&\quad = \sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\mathbb{R}_n(z_1)\pi_n\pi_n^* e_k e_l'\mathbb{R}_n(z_2)\pi_n\pi_n^* e_l + o_p(1) \\
&\quad \to z_1^{-1}z_2^{-1}\lim_{n\to\infty}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^* e_k e_l'\mathcal{R}_n(z_2)\pi_n\pi_n^* e_l \\
&\quad \overset{\mathcal{I}}{=} \lim_{n\to\infty}\biggl(\int \frac{z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)\biggl(\int \frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)\bigl(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)}+\alpha\mathcal{L}_{\pi,\mathbf{R}}\bigr) \\
&\quad \overset{\mathcal{N}}{=} \frac{(\nu_4+\alpha)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))(1+\underline{s}(z_2))},
\end{aligned}
\tag{43}
$$

and

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\,\phi_{2,j}(z_1)\phi_{2,j}(z_2) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}_{j-1}\bigl(\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\pi_n\,\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\pi_n\bigr) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)\,\mathrm{E}_j\, e_k'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)e_k\,\mathrm{E}_j\, e_l'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l \\
&\quad = \sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\pi_n\pi_n^*\mathbb{R}_n(z_1)e_k e_l'\pi_n\pi_n^*\mathbb{R}_n(z_2)e_l + o_p(1) \\
&\quad \to z_1^{-1}z_2^{-1}\lim_{n\to\infty}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\pi_n\pi_n^*\mathcal{R}_n(z_1)e_k e_l'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l \\
&\quad \overset{\mathcal{I}}{=} \lim_{n\to\infty}\biggl(\int \frac{z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)\biggl(\int \frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)\bigl(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)}+\alpha\mathcal{L}_{\pi,\mathbf{R}}\bigr) \\
&\quad \overset{\mathcal{N}}{=} \frac{(\nu_4+\alpha)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))(1+\underline{s}(z_2))}.
\end{aligned}
\tag{44}
$$

What is more, we get

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\,\phi_{3,j}(z_1)\phi_{3,j}(z_2)\\
&\quad= \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}_{j-1}\big(\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_1)\pi_n\\
&\qquad\times \mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n\big)\\
&\quad= \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathrm{E}_j\,e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)e_k\,\mathrm{E}_j\,e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l\\
&\qquad\times \mathrm{E}\big[\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j-1\big)\big(\mathbf{x}_j^*\mathbf{g}_l\mathbf{g}_l^*\mathbf{x}_j-1\big)\big]\\
&\quad= \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)\,\mathrm{E}_j\,e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)e_k\\
&\qquad\times \mathrm{E}_j\,e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l\\
&\quad= \sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathbb{R}_n(z_1)\pi_n\pi_n^*\mathbb{R}_n(z_1)e_ke_l'\mathbb{R}_n(z_2)\pi_n\pi_n^*\mathbb{R}_n(z_2)e_l+o_p(1)\\
&\quad\to z_1^{-2}z_2^{-2}\lim_{n\to\infty}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)e_ke_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l\\
&\quad\overset{\mathcal{I}}{=} z_1^{-2}z_2^{-2}\lim_{n\to\infty}\left(\int\frac{z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\right)^2\left(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\right)^2\big(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)}+\alpha\mathcal{L}_{\pi,\mathbf{R}}\big)\\
&\quad\overset{\mathcal{N}}{=} \frac{(\nu_4+\alpha)\mathcal{L}_\pi}{z_1^2z_2^2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))^2}.
\end{aligned}
\tag{45}
$$

Similarly, we have

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\,\phi_{1,j}(z_1)\phi_{2,j}(z_2)\\
&\quad= \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)\,\mathrm{E}_j\,e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*e_k\,\mathrm{E}_j\,e_l'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l\\
&\quad= \sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathbb{R}_n(z_1)\pi_n\pi_n^*e_ke_l'\pi_n\pi_n^*\mathbb{R}_n(z_2)e_l+o_p(1)\\
&\quad\to z_1^{-1}z_2^{-1}\lim_{n\to\infty}\sum_{k,l=1}^{p}\mathcal{G}_{kl}(\nu_4,\alpha)e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^*e_ke_l'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l\\
&\quad\overset{\mathcal{I}}{=} \lim_{n\to\infty}\left(\int\frac{z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\right)\left(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\right)\big(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)}+\alpha\mathcal{L}_{\pi,\mathbf{R}}\big)\\
&\quad\overset{\mathcal{N}}{=} \frac{(\nu_4+\alpha)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))(1+\underline{s}(z_2))},
\end{aligned}
\tag{46}
$$

and

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\, \phi_{1,j}(z_1)\phi_{3,j}(z_2) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)\, \mathrm{E}_j\, e_k'\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^* e_k\, \mathrm{E}_j\, e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l \\
&\quad = \sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\mathbb{R}_n(z_1)\pi_n\pi_n^* e_k e_l'\mathbb{R}_n(z_2)\pi_n\pi_n^*\mathbb{R}_n(z_2)e_l + o_p(1) \\
&\quad \to z_1^{-1}z_2^{-2}\lim_{n\to\infty}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\mathcal{R}_n(z_1)\pi_n\pi_n^* e_k e_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l \\
&\quad \overset{\mathcal{I}}{=} \lim_{n\to\infty}\biggl(\int \frac{-z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)\biggl(\int \frac{z_2^{-2}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)^2 \bigl(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)} + \alpha\mathcal{L}_{\pi,\mathbf{R}}\bigr) \\
&\quad \overset{\mathcal{N}}{=} \frac{-(\nu_4+\alpha)\mathcal{L}_\pi}{z_1 z_2^2(1+\underline{s}(z_1))(1+\underline{s}(z_2))^2}.
\end{aligned} \tag{47}
$$

Furthermore, one obtains

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\, \phi_{2,j}(z_1)\phi_{3,j}(z_2) \\
&\quad = \frac{1}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha)\, \mathrm{E}_j\, e_k'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1) e_k\, \mathrm{E}_j\, e_l'\mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_2)e_l \\
&\quad = \sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\pi_n\pi_n^*\mathbb{R}_n(z_1) e_k e_l'\mathbb{R}_n(z_2)\pi_n\pi_n^*\mathbb{R}_n(z_2)e_l + o_p(1) \\
&\quad \to z_1^{-1}z_2^{-2}\lim_{n\to\infty}\sum_{k,l=1}^{p} \mathcal{G}_{kl}(\nu_4,\alpha) e_k'\pi_n\pi_n^*\mathcal{R}_n(z_1) e_k e_l'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_l \\
&\quad \overset{\mathcal{I}}{=} \lim_{n\to\infty}\biggl(\int \frac{-z_1^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)\biggl(\int \frac{z_2^{-2}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)^2 \bigl(\nu_4\mathcal{L}_{\pi,\mathbf{G}}^{(2)} + \alpha\mathcal{L}_{\pi,\mathbf{R}}\bigr) \\
&\quad \overset{\mathcal{N}}{=} \frac{-(\nu_4+\alpha)\mathcal{L}_\pi}{z_1 z_2^2(1+\underline{s}(z_1))(1+\underline{s}(z_2))^2},
\end{aligned} \tag{48}
$$

and

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\, \phi_{0,j}(z_1)\phi_{1,j}(z_2) \\
&\quad = -z_1\underline{s}(z_1)\sum_{j=1}^{n} \mathrm{E}_{j-1}\bigl(\mathrm{E}_j\,\delta_j(z_1)\,\mathrm{E}_j\,\pi_n^*(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n\bigr) \\
&\quad = -z_1\underline{s}(z_1)\sum_{j=1}^{n}\sum_{k=1}^{p} \mathrm{E}_{j-1}\biggl(\mathrm{E}_j\biggl(\mathbf{r}_j^*\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{r}_j \\
&\qquad - \frac{1}{n}\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{R}_n\mathbf{A}_j^{-1}(z_1)\pi_n\biggr)
\end{aligned}
$$

$$
\begin{aligned}
&\quad \times \mathrm{E}_j\, e_k' \mathbf{A}_j^{-1}(z_2)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big)\Big)\\
&= -z_1\underline{s}(z_1)\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}_{j-1}\,\mathrm{E}_{(-j)}\bigg(\bigg(\mathbf{r}_j^*\mathbf{A}_j^{-1}(z_1)\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{r}_j\\
&\quad - \frac{1}{n}\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{R}_n\mathbf{A}_j^{-1}(z_1)\pi_n\bigg)\\
&\quad \times e_k'\breve{\mathbf{A}}_j^{-1}(z_2)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big)\bigg)\\
&= -z_1\underline{s}(z_1)\sum_{k=1}^{p} e_k'\mathbb{R}_n(z_2)\pi_n\pi_n^* e_k\Bigg(\nu_4\sum_{l=1}^{p} e_l'\mathbf{G}_n^*\mathbb{R}_n(z_1)\pi_n\pi_n^*\mathbb{R}_n(z_1)\mathbf{G}_n e_l|g_{kl}|^2\\
&\quad + \alpha\pi_n^*\mathbb{R}_n(z_1)\mathbf{R}_n e_k e_k'\mathbf{R}_n\mathbb{R}_n(z_1)\pi_n\Bigg) + o_p(1)\\
&\to z_1^{-1}z_2^{-1}\underline{s}(z_1)\lim_{n\to\infty}\sum_{k=1}^{p} e_k'\mathcal{R}_n(z_2)\pi_n\pi_n^* e_k\\
&\quad \times \Bigg(\nu_4\sum_{l=1}^{p} e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_n e_l|g_{kl}|^2\\
&\quad + \alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_n e_k e_k'\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\Bigg)\\
&\overset{\mathcal{I}}{=} \lim_{n\to\infty}\bigg\{\nu_4 z_1^{-1}\underline{s}(z_1)\bigg(\int\frac{dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\bigg)^2\bigg(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\bigg)\mathcal{L}_{\pi,\mathbf{G}}^{(3)}\\
&\quad + \alpha z_1^{-1}\underline{s}(z_1)\bigg(\int\frac{t\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\bigg)^2\bigg(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\bigg)\mathcal{L}_\pi\bigg\}\\
&\overset{\mathcal{N}}{=} \frac{(\nu_4+\alpha)\underline{s}(z_1)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))}.
\end{aligned}\tag{49}
$$

Similarly, we have

$$
\begin{aligned}
&\sum_{j=1}^{n}\mathrm{E}_{j-1}\,\phi_{0,j}(z_1)\phi_{2,j}(z_2)\\
&\quad = -z_1\underline{s}(z_1)\sum_{j=1}^{n}\mathrm{E}_{j-1}\big(\mathrm{E}_j\,\delta_j(z_1)\,\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\pi_n\big)\\
&\quad \to z_1^{-1}z_2^{-1}\underline{s}(z_1)\lim_{n\to\infty}\sum_{k=1}^{p} e_k'\pi_n\pi_n^*\mathcal{R}_n(z_2)e_k\\
&\qquad \times \Bigg(\nu_4\sum_{l=1}^{p} e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_n e_l e_l'\mathbf{g}_k\mathbf{g}_k^* e_l\\
&\qquad + \alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_n e_k e_k'\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\Bigg)
\end{aligned}\tag{50}
$$

$$\begin{aligned}&\stackrel{\mathcal{I}}{=}\lim_{n\to\infty}\biggl\{\nu_4 z_1^{-1}\underline{s}(z_1)\biggl(\int\frac{dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)^2\biggl(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)\mathcal{L}_{\pi,\mathbf{G}}^{(3)}\\&\quad+\alpha z_1^{-1}\underline{s}(z_1)\biggl(\int\frac{t\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)^2\biggl(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)\mathcal{L}_\pi\biggr\}\\&\stackrel{\mathcal{N}}{=}\frac{(\nu_4+\alpha)\underline{s}(z_1)\mathcal{L}_\pi}{z_1z_2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))},\end{aligned}$$

and

$$\begin{aligned}&\sum_{j=1}^n \mathrm{E}_{j-1}\,\phi_{0,j}(z_1)\phi_{3,j}(z_2)\\&\quad=-z_1\underline{s}(z_1)\sum_{j=1}^n\mathrm{E}_{j-1}\bigl(\mathrm{E}_j\,\delta_j(z_1)\,\mathrm{E}_j\,\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n\bigr)\\&\quad\to -z_1^{-1}z_2^{-2}\underline{s}(z_1)\lim_{n\to\infty}\sum_{k=1}^p e_k'\mathcal{R}_n(z_2)\pi_n\pi_n^*\mathcal{R}_n(z_2)e_k\\&\qquad\times\biggl(\nu_4\sum_{l=1}^p e_l'\mathbf{G}_n^*\mathcal{R}_n(z_1)\pi_n\pi_n^*\mathcal{R}_n(z_1)\mathbf{G}_ne_le_l'\mathbf{g}_k\mathbf{g}_k^*e_l\\&\qquad+\alpha\pi_n^*\mathcal{R}_n(z_1)\mathbf{R}_ne_ke_k'\mathbf{R}_n^*\mathcal{R}_n(z_1)\pi_n\biggr)\\&\quad\stackrel{\mathcal{I}}{=}-\lim_{n\to\infty}\biggl\{\nu_4 z_1^{-1}\underline{s}(z_1)\biggl(\int\frac{dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)^2\biggl(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)^2\mathcal{L}_{\pi,\mathbf{G}}^{(3)}\\&\qquad+\alpha z_1^{-1}\underline{s}(z_1)\biggl(\int\frac{t\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_1)t+1}\biggr)^2\biggl(\int\frac{z_2^{-1}\,dH_n(t)}{\underline{s}_{F^{\rho_n,H_n}}(z_2)t+1}\biggr)^2\mathcal{L}_\pi\biggr\}\\&\quad\stackrel{\mathcal{N}}{=}\frac{(\nu_4+\alpha)\underline{s}(z_1)\mathcal{L}_\pi}{z_1z_2^2(1+\underline{s}(z_1))^2(1+\underline{s}(z_2))^2}.\end{aligned}\tag{51}$$

Combining the argument above, we arrive at

$$\begin{aligned}&\lim_{n\to\infty}\sum_{j=1}^n\mathrm{E}_{j-1}\bigl(\phi_{0,j}(z_1)+2^{-1}\phi_{1,j}(z_1)+2^{-1}\phi_{2,j}(z_1)+z_1\phi_{3,j}(z_1)\bigr)\\&\quad\times\bigl(\phi_{0,j}(z_2)+2^{-1}\phi_{1,j}(z_2)+2^{-1}\phi_{2,j}(z_2)+z_2\phi_{3,j}(z_2)\bigr)\\&=L_{0,0}(z_1,z_2)+4^{-1}L_{1,1}(z_1,z_2)+4^{-1}L_{1,1}(z_1,z_2)+z_1z_2L_{3,3}(z_1,z_2)\\&\quad+2^{-1}L_{0,1}(z_1,z_2)+2^{-1}L_{0,2}(z_1,z_2)+2^{-1}L_{0,1}(z_2,z_1)+2^{-1}L_{0,2}(z_2,z_1)\\&\quad+z_2L_{0,3}(z_1,z_2)+z_1L_{0,3}(z_2,z_1)+2^{-1}z_2L_{1,3}(z_1,z_2)+2^{-1}z_1L_{1,3}(z_2,z_1)\\&\quad+4^{-1}L_{1,2}(z_1,z_2)+4^{-1}L_{1,2}(z_2,z_1)+2^{-1}z_2L_{2,3}(z_1,z_2)+2^{-1}z_1L_{2,3}(z_2,z_1).\end{aligned}$$

Under the isotropic conditions, we shall get the corresponding variance-covariance function of Corollary 3.1 by the second to last equations in (42)–(51). Using the relationship

$\frac{1}{z(1+\underline{s}(z))} = -s(z)$ under the null case, we obtain that

$$
\begin{aligned}
& L_{0,0}(z_1, z_2) = \nu_4 \mathcal{L}_\pi \frac{\underline{s}(z_1)\underline{s}(z_2)}{z_1 z_2 (1+\underline{s}(z_1))^2 (1+\underline{s}(z_2))^2} + \frac{\alpha (z_2 \underline{s}(z_2) - z_1 \underline{s}(z_1))^2}{\rho^2 z_1 z_2 (z_2 - z_1)(\underline{s}(z_2) - \underline{s}(z_1))}, \\
& L_{1,1}(z_1, z_2) = L_{2,2}(z_1, z_2) = L_{1,2}(z_1, z_2) = (\nu_4 + \alpha) s(z_1) s(z_2) \mathcal{L}_\pi, \\
& L_{3,3}(z_1, z_2) = (\nu_4 + \alpha) s^2(z_1) s^2(z_2) \mathcal{L}_\pi, \\
& L_{1,3}(z_1, z_2) = L_{2,3}(z_1, z_2) = (\nu_4 + \alpha) s(z_1) s^2(z_2) \mathcal{L}_\pi, \\
& L_{0,1}(z_1, z_2) = L_{0,2}(z_1, z_2) = (\nu_4 + \alpha)\big(1 + z_1 s(z_1)\big) s(z_1) s(z_2) \mathcal{L}_\pi, \\
& L_{0,3}(z_1, z_2) = (\nu_4 + \alpha)\big(1 + z_1 s(z_1)\big) s(z_1) s^2(z_2) \mathcal{L}_\pi.
\end{aligned} \tag{52}
$$

Those give us

$$
\begin{aligned}
& L_{1,1}(z_1, z_2) + z_1 z_2 L_{3,3}(z_1, z_2) + L_{0,1}(z_1, z_2) + L_{0,1}(z_2, z_1) \\
& \qquad + z_2 L_{0,3}(z_1, z_2) + z_1 L_{0,3}(z_2, z_1) + z_2 L_{1,3}(z_1, z_2) + z_1 L_{1,3}(z_2, z_1) \\
& \quad = (\nu_4 + \alpha)\mathcal{L}_\pi \big(s(z_1)s(z_2) + z_1 z_2 s^2(z_1) s^2(z_2) + \big(1 + z_1 s(z_1)\big) s(z_1) s(z_2) \\
& \qquad + \big(1 + z_2 s(z_2)\big) s(z_2) s(z_1) + z_2 \big(1 + z_1 s(z_1)\big) s(z_1) s^2(z_2) \\
& \qquad + z_1 \big(1 + z_2 s(z_2)\big) s^2(z_1) s(z_2) + z_2 s(z_1) s^2(z_2) + z_1 s(z_2) s^2(z_1)\big) \\
& \quad = (\nu_4 + \alpha)\mathcal{L}_\pi \big(\big(s(z_1)s(z_2) + z_1 z_2 s^2(z_1) s^2(z_2) + z_2 s(z_1) s^2(z_2) + z_1 s(z_2) s^2(z_1)\big) \\
& \qquad + \big(1 + z_1 s(z_1)\big) s(z_1) s(z_2) + \big(1 + z_2 s(z_2)\big) s(z_2) s(z_1) \\
& \qquad + z_2 \big(1 + z_1 s(z_1)\big) s(z_1) s^2(z_2) + z_1 \big(1 + z_2 s(z_2)\big) s^2(z_1) s(z_2)\big) \\
& \quad = 3(\nu_4 + \alpha)\mathcal{L}_\pi s(z_1) s(z_2) \big(1 + z_1 s(z_1)\big)\big(1 + z_2 s(z_2)\big) \\
& \quad = 3(\nu_4 + \alpha)\mathcal{L}_\pi \frac{\underline{s}(z_1)\underline{s}(z_2)}{z_1 z_2 (1+\underline{s}(z_1))^2 (1+\underline{s}(z_2))^2},
\end{aligned}
$$

which indicates the variance-covariance function under the null case.

5.5.6. *Proof of tightness.* We need to prove that for $t = 0, 1, 2, 3$,

$$
\sup_{n, z_1, z_2 \in \mathcal{C}_n} \frac{\mathrm{E}\,|\sum_{j=1}^n (\phi_{t,j}(z_1) - \phi_{t,j}(z_2))|^2}{|z_1 - z_2|^2} \le C. \tag{53}
$$

First, we have from Bai, Miao and Pan (2007) that

$$
\sup_{n, z_1, z_2 \in \mathcal{C}_n} \frac{\mathrm{E}\,|\sum_{j=1}^n (\phi_{0,j}(z_1) - \phi_{0,j}(z_2))|^2}{|z_1 - z_2|^2} \le C. \tag{54}
$$

Then, it is apparent that

$$
\frac{\sum_{j=1}^n (\phi_{1,j}(z_1) - \phi_{1,j}(z_2))}{z_1 - z_2} = \frac{-1}{\sqrt{n}} \sum_{j=1}^n \pi_n^* \mathbf{A}_j^{-1}(z_1) \mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j} - \mathbf{I})\pi_n. \tag{55}
$$

This implies that

$$
\begin{aligned}
&\frac{\mathrm{E}\left|\sum_{j=1}^{n}(\phi_{1,j}(z_1)-\phi_{1,j}(z_2))\right|^2}{|z_1-z_2|^2}\\
&\quad\le \frac{1}{n}\mathrm{E}\left|\sum_{j=1}^{n}\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\pi_n\right|^2\\
&\quad= \frac{1}{n}\mathrm{E}\left|\sum_{j=1}^{n}\sum_{k=1}^{n}e_k'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j-1\big)\right|^2\\
&\quad= \frac{1}{n}\sum_{j=1}^{n}\mathrm{E}\left|\sum_{k=1}^{n}e_k'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j-1\big)\right|^2\\
&\quad\le \frac{1}{n}\sum_{j=1}^{n}\mathrm{E}\sum_{k=1}^{n}\left|e_k'\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)e_k\right|^2\\
&\quad\le \frac{1}{n}\sum_{j=1}^{n}\mathrm{E}\,\mathrm{tr}\,\pi_n\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\mathbf{A}_j^{-1}(\bar z_2)\mathbf{A}_j^{-1}(\bar z_1)\pi_n^*\pi_n\le C.
\end{aligned}
\tag{56}
$$

Taking the same procedure, and noting that

$$
\begin{aligned}
&\frac{\sum_{j=1}^{n}(\phi_{3,j}(z_1)-\phi_{3,j}(z_2))}{z_1-z_2}\\
&\quad=\frac{n^{-1/2}\sum_{j=1}^{n}(\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_1)\pi_n-\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n)}{z_1-z_2}\\
&\quad=\frac{n^{-1/2}\sum_{j=1}^{n}(\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_1)\pi_n-\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n)}{z_1-z_2}\\
&\qquad+\frac{n^{-1/2}\sum_{j=1}^{n}(\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n-\pi_n^*\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n)}{z_1-z_2}\\
&\quad=\frac{-1}{\sqrt{n}}\sum_{j=1}^{n}\pi_n^*\mathbf{A}_j^{-1}(z_1)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)\pi_n\\
&\qquad-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\pi_n^*\mathbf{A}_j^{-1}(z_1)\mathbf{A}_j^{-1}(z_2)(\mathbb{D}_{\Xi_j}-\mathbf{I})\mathbf{A}_j^{-1}(z_2)\pi_n,
\end{aligned}
\tag{57}
$$

we can obtain the boundedness for $t=2,3$. The proof of tightness is then complete.

5.5.7. *Convergence to* 0 *of expectation.* This part is to investigate the limit of expectation. First, by (2.5)–(2.8) in Bai, Miao and Pan (2007), we have

$$
\begin{aligned}
&\sqrt{n}\big(\pi_n^*\,\mathrm{E}\,\Phi_0(z)\pi_n-s^{\mathbf{R}_n}_{\rho_n,\pi_n}(z)\big)\\
&\quad=\sqrt{n}\big(\big(-z^{-1}\pi_n^*\big(1+\mathrm{E}\,\underline{s}_n(z)\mathbf{R}_n\big)^{-1}\pi_n\big)-\big(-z^{-1}\pi_n^*\big(1+\underline{s}_{F^{\rho_n,H_n}}(z)\mathbf{R}_n\big)^{-1}\pi_n\big)\big)+o(1),
\end{aligned}
$$

where $\underline{s}_n(z)$ is defined in (4). Combining with (5.6) in Bai, Miao and Pan (2007), we conclude that

$$
\sqrt{n}\big(\pi_n^*\,\mathrm{E}\,\Phi_0(z)\pi_n-s^{\mathbf{R}_n}_{\rho_n,\pi_n}(z)\big)\to 0.
$$

Now, consider the expectation of $\sqrt{n}\Phi_t$ for $t = 1, 2, 3$. When $t = 1$, we have

$$\begin{aligned}
\sqrt{n}\,\mathrm{E}\,\Phi_1 &= \sqrt{n}\,\mathrm{E}\,\pi_n^*(\mathbb{D}_{\Xi_n} - \mathbf{I})\mathbf{A}^{-1}(z)\pi_n \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\,e_k'\mathbf{A}^{-1}(z)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\,e_k'\mathbf{A}_j^{-1}(z)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \\
&\quad - \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\,\beta_j(z)e_k'\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \\
&= -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\,b_j(z)e_k'\mathbf{A}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \qquad (58)\\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\,b_j(z)\beta_j(z)\widehat{\gamma}_j(z)e_k'\mathbf{A}_j^{-1}(z) \\
&\qquad \times \mathbf{r}_j\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\pi_n^* e_k\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \\
&= -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{k=1}^{p} b_j(z)\,\mathrm{E}\big(\mathbf{x}_j^*\mathbf{g}_k\mathbf{g}_k^*\mathbf{x}_j - 1\big) \\
&\quad \times \bigg(\mathbf{r}_j^*\mathbf{A}_j^{-1}(z)\pi_n\pi_n^* e_k e_k'\mathbf{A}_j^{-1}(z)\mathbf{r}_j - \frac{1}{n}\pi_n^* e_k e_k'\mathbf{A}_j^{-1}(z)\mathbf{R}_n\mathbf{A}_j^{-1}(z)\pi_n\bigg) \\
&= O\big(n^{-1/2}\big).
\end{aligned}$$

The estimates for $t = 2, 3$ are the same thus omitted.

5.6. *Proof of Theorem* 3.5. The proof of Theorem 3.5 is essentially a part of the proofs of Theorem 3.3 and Theorem 3.4 after a substitution of $\widehat{\mathbf{S}}_n$ for $\widehat{\mathbf{R}}_n$ and a substitution of $\boldsymbol{\Sigma}_n$ for $\mathbf{R}_n$ in all of the notation that we need. To be specific, we shall first prove that

$$F_{\upsilon,\pi_n}^{\widehat{\mathbf{S}}_n}(x) - F_{\rho_n,\pi_n}^{\boldsymbol{\Sigma}_n}(x) \to 0, \quad \text{a.s.}$$

Then, define

$$\mathbb{G}_n(x) = \sqrt{n}\big(F_{\upsilon,\pi_n}^{\widehat{\mathbf{S}}_n}(x) - F_{\rho_n,\pi_n}^{\boldsymbol{\Sigma}_n}(x)\big), \qquad \mathbb{M}_n(z) = \sqrt{n}\big(\pi_n^*(\widehat{\mathbf{S}}_n - z\mathbf{I})^{-1}\pi_n - s_{\rho_n,\pi_n}^{\boldsymbol{\Sigma}_n}\big).$$

Note that in the sample covariance matrix case, we do not need to decompose the object $\pi_n^*(\widehat{\mathbf{S}}_n - z\mathbf{I})^{-1}\pi_n$. That is, consider the term $\Phi_0 = \pi_n^*(\Xi_n - z\mathbf{I})^{-1}\pi_n$ after replacing $\mathbf{G}_n$ with $\boldsymbol{\Gamma}_n$ in the definition of Ξ_n. Recall the arguments concerning Φ_0 in Sections 5.5.3–5.5.6, we complete the proof of this theorem.

It is worth noting that comparing to Theorem 2 in Bai, Miao and Pan (2007), we need to consider the case where $\nu_4 \neq 0$. This is achieved by (42), where Lemma 5.4 is applied.

5.7. *Proof of Theorem* 4.1. By the argument in Section 3 of Pan and Zhou (2008), combining with Corollary 3.2 of this paper, to prove this theorem, it is enough to show that:

(1) $\max_i \frac{t_{i,1}}{\|\mathbf{t}_1\|} \to 0$ in probability.
(2) The asymptotical distribution of $\sqrt{p}(\|\mathbf{t}_1\|^2 - 1)$ is $N(0, \mathrm{E}\,|x_{1,1}|^4 - 1)$.

In fact, for $1 \le j \le n$, let $\pi^{(j)} = (\frac{t_{1,1}}{\|\mathbf{t}_1\|}, \frac{t_{2,1}}{\|\mathbf{t}_1\|}, \ldots, \frac{t_{p,1}}{\|\mathbf{t}_1\|})'$. If (1) is true, we have $\sum_{i=1}^{p} (\pi_i^{(j)})^4 \to 0$ in probability. Then, noting the fact that the population correlation matrix in this case is $\mathbf{I}$, by Corollary 3.2, we know that the fluctuation of the VLSS $\sqrt{p}(\frac{\mathbf{s}_j^* \mathbb{H}_j^u \mathbf{s}_j}{|\mathbf{s}_j|^2} - \frac{1}{p} \operatorname{tr} \mathbb{H}_j^u)$ will be the same as the one of sample covariance matrix under the case where the population covariance matrix equals the identity.

Recall the definition of $t_{i,1} = \frac{s_{i,1}}{\sqrt{v_i^{(1)}}}$, $1 \le i \le p$, with $v_i^{(1)} = \frac{1}{n-1} \sum_{k \neq 1} s_{i,k}^2$. We see that $(t_{1,1}, \ldots, t_{p,1})$ are independent. Write

$$\|\mathbf{t}_1\|^2 - 1 = \sum_{i=1}^{p} t_{i,1}^2 - 1 = \sum_{i=1}^{p} \frac{s_{i,1}^2}{v_i^{(1)}} - 1 = \frac{1}{p} \sum_{i=1}^{p} (x_{i,1}^2 - 1) + \frac{1}{p} \sum_{i=1}^{p} x_{i,1}^2 \left(\frac{\sigma_1^2}{v_i^{(1)}} - 1 \right).$$

First, (1) is an easy consequence of the facts that $\|\mathbf{t}_1\| \to 1$ in probability. We now prove (2). We have by the central limit theorem that $\frac{1}{\sqrt{p}} \sum_{i=1}^{p} (x_{i,1}^2 - 1) \xrightarrow{d} N(0, \mathrm{E}\,|x_{1,1}|^4 - 1)$. Also, we have by lemma 5.7 that $\max_i \frac{\sigma_1^2}{v_i^{(1)}} - 1 = O_p(p^{-1/2})$, which implies that

$$\frac{1}{\sqrt{p}} \sum_{i=1}^{p} x_{i,1}^2 \left(\frac{\sigma_1^2}{v_i^{(1)}} - 1 \right) \to 0$$

in probability. These prove (2).

The proof of this theorem is thus complete.

APPENDIX: TRUNCATION AND CENTRALIZATION IN LEMMA 5.7

Given the condition $\mathrm{E}\,|x_{11}|^4 < \infty$, we have that for any $\eta > 0$,

$$\sum_{k=1}^{\infty} \eta^{-2} 2^{2k} \mathrm{P}(|x_{11}| \ge \eta 2^{k/2}) < \infty.$$

Then, we can select a slowly decreasing sequence of constants $\tau_n \to 0$ such that

$$\sum_{k=1}^{\infty} \eta_{2^k}^{-2} 2^{2k} \mathrm{P}(|x_{11}| \ge \eta_{2^k} 2^{k/2}) < \infty.$$

Denote $\tilde{x}_{jk} = x_{jk} I(|x_{jk}| \le \eta_n \sqrt{n})$. Also denote $\widetilde{\mathbf{X}}_n = (\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n) = (\tilde{x}_{jk})_{p \times n}$, $\widetilde{\mathbf{S}}_n = \frac{1}{n} \widetilde{\mathbf{X}}_n \widetilde{\mathbf{X}}_n^*$ and $\widetilde{\mathbf{\Xi}}_n = \frac{1}{n} \mathbf{G}_n \widetilde{\mathbf{X}}_n \widetilde{\mathbf{X}}_n^* \mathbf{G}_n^*$.

We have

$$\begin{aligned}
\mathrm{P}(\mathbf{\Xi}_n \neq \widetilde{\mathbf{\Xi}}_n, \text{i.o.}) &= \mathrm{P}(\widehat{\mathbf{S}}_n \neq \widetilde{\mathbf{S}}_n, \text{i.o.}) \\
&= \lim_{N \to \infty} \sum_{l=N}^{\infty} \mathrm{P}\left(\bigcup_{2^l < n \le 2^{l+1}} \bigcup_{k=1}^{n} \bigcup_{j=1}^{p} \{x_{jk} \neq \tilde{x}_{jk}\} \right) \\
&\le \lim_{N \to \infty} \sum_{l=N}^{\infty} \mathrm{P}\left(\bigcup_{2^l < n \le 2^{l+1}} \bigcup_{k=1}^{2^{l+1}} \bigcup_{j=1}^{2\rho 2^{l+1}} \{|x_{jk}| \ge \eta_{2^l} 2^{l/2}\} \right) \\
&= \lim_{N \to \infty} \sum_{l=N}^{\infty} \mathrm{P}\left(\bigcup_{k=1}^{2^{l+1}} \bigcup_{j=1}^{2\rho 2^{l+1}} \{|x_{jk}| \ge \eta_{2^l} 2^{l/2}\} \right) \\
&\le 8\rho \lim_{N \to \infty} \sum_{l=N}^{\infty} 2^{2l} \mathrm{P}(|x_{11}| \ge \eta_{2^l} 2^{l/2}) \to 0.
\end{aligned}$$

This implies that we can replace the variables in the data matrix $\mathbf{X}_n$ with the truncated version without changing any properties of $\widehat{\mathbf{S}}_n$ and $\boldsymbol{\Xi}_n$ almost surely.

We next need the recentralization steps to make the variables have zero mean again. To this end, denote $\bar{\boldsymbol{\Xi}}_n = \frac{1}{n}\mathbf{G}_n\bar{\mathbf{X}}_n\bar{\mathbf{X}}_n^*\mathbf{G}_n^*$ where $\bar{\mathbf{X}}_n = (\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_n) = (\bar{x}_{i,j})_{p\times n}$ with $\bar{x}_{i,j} = \tilde{x}_{jk} - \mathrm{E}\,\tilde{x}_{jk}$. Note that $\mathrm{E}\,\tilde{x}_{jk} \le C(\eta_n\sqrt{n})^{-3}$, we have

$$
\begin{aligned}
&\|\widetilde{\bar{\boldsymbol{\Xi}}}_n - \bar{\boldsymbol{\Xi}}_n\| \\
&= \max_k \left| \frac{1}{n}\sum_{j=1}^n \mathbf{g}_k^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{g}_k - \frac{1}{n}\sum_{j=1}^n \mathbf{g}_k^*\bar{\mathbf{x}}_j\bar{\mathbf{x}}_j^*\mathbf{g}_k \right| \\
&= \max_k \left| \frac{1}{n}\sum_{j=1}^n \mathbf{g}_k^*(\mathbf{x}_j\mathbf{x}_j^* - \bar{\mathbf{x}}_j\bar{\mathbf{x}}_j^*)\mathbf{g}_k \right| \le \max_k \frac{1}{n}\sum_{j=1}^n |\mathbf{g}_k^*(\mathbf{x}_j\mathbf{x}_j^* - \bar{\mathbf{x}}_j\bar{\mathbf{x}}_j^*)\mathbf{g}_k| \\
&\le \max_k \left(\frac{1}{n}\sum_{j=1}^n |\mathbf{g}_k^*(\mathrm{E}\,\widetilde{\mathbf{x}}_j)\widetilde{\mathbf{x}}_j^*\mathbf{g}_k| + \frac{1}{n}\sum_{j=1}^n |\mathbf{g}_k^*\widetilde{\mathbf{x}}_j(\mathrm{E}\,\widetilde{\mathbf{x}}_j^*)\mathbf{g}_k| + \frac{1}{n}\sum_{j=1}^n |\mathbf{g}_k^*\,\mathrm{E}\,\widetilde{\mathbf{x}}_j\,\mathrm{E}\,\widetilde{\mathbf{x}}_j^*\mathbf{g}_k| \right) \\
&= O\big((\eta_n\sqrt{n})^{-2}\big) \to 0.
\end{aligned} \tag{59}
$$

Acknowledgments. The authors thank the Editor, the Associate Editor, the anonymous reviewers for their constructive feedback and helpful comments, which have substantially improved the quality of this paper.

Funding. Yanqing Yin's work is partially supported by a project under Grant NSFC11801234.

Yanyuan Ma's work is partially supported by grants from the National Institute of Health.

REFERENCES

AKHIEZER, N. I. and GLAZMAN, I. M. (1993). *Theory of Linear Operators in Hilbert Space*. Dover, New York. MR1255973

ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR1990662

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 https://doi.org/10.1214/009117906000001079

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 https://doi.org/10.1214/aop/1022855421

BAI, Z. D. and SILVERSTEIN, J. W. (1999). Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.* **27** 1536–1555. MR1733159 https://doi.org/10.1214/aop/1022677458

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 https://doi.org/10.1214/aop/1078415845

BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2567175 https://doi.org/10.1007/978-1-4419-0661-8

BAI, Z. D. and YIN, Y. Q. (1993). Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.* **21** 1275–1294. MR1235416

BAO, Z., PAN, G. and ZHOU, W. (2012). Tracy–Widom law for the extreme eigenvalues of sample correlation matrices. *Electron. J. Probab.* **17** no. 88, 32 pp. MR2988403 https://doi.org/10.1214/EJP.v17-1962

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR1324786

CAI, T. T. and JIANG, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Statist.* **39** 1496–1525. MR2850210 https://doi.org/10.1214/11-AOS879

DUMITRIU, I. and EDELMAN, A. (2002). Matrix models for beta ensembles. *J. Math. Phys.* **43** 5830–5847. MR1936554 https://doi.org/10.1063/1.1507823

EL KAROUI, N. (2007). Tracy–Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.* **35** 663–714. MR2308592 https://doi.org/10.1214/009117906000000917

EL KAROUI, N. (2009). Concentration of measure and spectra of random matrices: Applications to correlation matrices, elliptical distributions and beyond. *Ann. Appl. Probab.* **19** 2362–2405. MR2588248 https://doi.org/10.1214/08-AAP548

FAN, J., GUO, J. and ZHENG, S. (2019). Estimating number of factors by adjusted eigenvalues thresholding. Preprint. Available at arXiv:1909.10710.

GAO, J., HAN, X., PAN, G. and YANG, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 677–693. MR3641402 https://doi.org/10.1111/rssb.12189

GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261. MR0566592

HERO, A. and RAJARATNAM, B. (2011). Large-scale correlation screening. *J. Amer. Statist. Assoc.* **106** 1540–1552. MR2896855 https://doi.org/10.1198/jasa.2011.tm11015

HONIG, M. L. and XIAO, W. (2001). Performance of reduced-rank linear interference suppression. *IEEE Trans. Inf. Theory* **47** 1928–1946. MR1842528 https://doi.org/10.1109/18.930928

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

JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961 https://doi.org/10.1214/aos/1009210544

JOHNSTONE, I. M. and MA, Z. (2012). Fast approach to the Tracy–Widom law at the edge of GOE and GUE. *Ann. Appl. Probab.* **22** 1962–1988. MR3025686 https://doi.org/10.1214/11-AAP819

JONSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38. MR0650926 https://doi.org/10.1016/0047-259X(82)90080-X

LEDOIT, O. and PÉCHÉ, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Related Fields* **151** 233–264. MR2834718 https://doi.org/10.1007/s00440-010-0298-3

LEE, J. O. and SCHNELLI, K. (2016). Tracy–Widom distribution for the largest eigenvalue of real sample covariance matrices with general population. *Ann. Appl. Probab.* **26** 3786–3839. MR3582818 https://doi.org/10.1214/16-AAP1193

LYTOVA, A. and PASTUR, L. (2009). Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* **37** 1778–1840. MR2561434 https://doi.org/10.1214/09-AOP452

MARČENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR, Sb.* **1** 457–483. Available at http://iopscience.iop.org/0025-5734/1/4/A01/pdf/0025-5734_1_4_A01.pdf.

MESTRE, X. and VALLET, P. (2017). Correlation tests and linear spectral statistics of the sample correlation matrix. *IEEE Trans. Inf. Theory* **63** 4585–4618. MR3666978 https://doi.org/10.1109/TIT.2017.2689780

MORALES-JIMENEZ, D., JOHNSTONE, I. M., MCKAY, M. R. and YANG, J. (2021). Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. *Statist. Sinica* **31** 571–601. MR4286186 https://doi.org/10.5705/ss.20

NAJIM, J. and YAO, J. (2016). Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *Ann. Appl. Probab.* **26** 1837–1887. MR3513608 https://doi.org/10.1214/15-AAP1135

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 https://doi.org/10.1214/07-AAP477

PAUL, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865

PILLAI, N. S. and YIN, J. (2012). Edge universality of correlation matrices. *Ann. Statist.* **40** 1737–1763. MR3015042 https://doi.org/10.1214/12-AOS1022

SHCHERBINA, M. (2011). Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. *J. Math. Phys. Anal. Geom.* **69** 176–192.

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

SILVERSTEIN, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.* **55** 331–339. MR1370408 https://doi.org/10.1006/jmva.1995.1083

TIKHOMIROV, K. (2015). The limit of the smallest singular value of random matrices with i.i.d. entries. *Adv. Math.* **284** 1–20. MR3391069 https://doi.org/10.1016/j.aim.2015.07.020

TRACY, C. A. and WIDOM, H. (1994). Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159** 151–174. MR1257246

TSE, D. N. C. and HANLY, S. V. (1999). Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Trans. Inf. Theory* **45** 641–657. MR1677023 https://doi.org/10.1109/18.749008

WACHTER, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18. MR0467894 https://doi.org/10.1214/aop/1176995607

XI, H., YANG, F. and YIN, J. (2020). Convergence of eigenvector empirical spectral distribution of sample covariance matrices. *Ann. Statist.* **48** 953–982. MR4102683 https://doi.org/10.1214/19-AOS1832

XIA, N., QIN, Y. and BAI, Z. (2013). Convergence rates of eigenvector empirical spectral distribution of large dimensional sample covariance matrix. *Ann. Statist.* **41** 2572–2607. MR3161438 https://doi.org/10.1214/13-AOS1154

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 https://doi.org/10.1007/s10959-009-0270-2

YANG, F. (2020). Linear spectral statistics of eigenvectors of anisotropic sample covariance matrices. Preprint. Available at arXiv:2005.00999.

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 https://doi.org/10.1007/BF00353874

ZHENG, S., BAI, Z. and YAO, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Statist.* **43** 546–591. MR3316190 https://doi.org/10.1214/14-AOS1292

ZHENG, S., CHENG, G., GUO, J. and ZHU, H. (2019). Test for high-dimensional correlation matrices. *Ann. Statist.* **47** 2887–2921. MR3988776 https://doi.org/10.1214/18-AOS1768