TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 377, Number 6, June 2024, Pages 3805–3843
https://doi.org/10.1090/tran/9070
Article electronically published on April 11, 2024

SEPARABLE SAMPLE COVARIANCE MATRICES UNDER ELLIPTICAL POPULATIONS WITH APPLICATIONS

HUIQIN LI, GUANGMING PAN, YANQING YIN, AND WANG ZHOU

ABSTRACT. This paper is to investigate the spectral properties of separable covariance matrices under elliptical populations. The separable covariance matrix model can handle both cross-row and cross-column correlations thus gain more popularity recently. Under the high-dimensional setting where the dimension p and the sample size n tend to infinity proportionally, we find the limit of the empirical spectral distribution and establish the central limit theorems (CLT) for linear spectral statistics of such kinds of sample covariance matrices. Some applications of our established CLT are also given.

1. Introduction and main results

Spectral properties of sample covariance matrices are important in engineering and statistics, see [16], [1]. When the sample size n tends to infinity while the dimension p of population is fixed, the sample covariance matrix is asymptotical a good estimator for the population covariance matrix. However, as has been well known, it is not the case when the dimension p is comparable or even larger compared with the sample size n. Thus, it is reasonable to investigate the relationship between the population covariance matrix and its sample version under the high-dimensional framework.

We begin by introducing some definitions. Let $\mathbf{A}$ be any $n\times n$ Hermitian matrix whose eigenvalues are denoted by $\lambda_j, j=1,2,\cdots,n$. Then the Empirical Spectral Distribution (ESD) of $\mathbf{A}$ is defined as

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

where $I(\cdot)$ is the indicator function. The Stieltjes transform of $F^{\mathbf{A}}(x)$ is given by

$$m_{F^{\mathbf{A}}}(z)=\int_{-\infty}^{+\infty}(x-z)^{-1}dF^{\mathbf{A}}(x), \quad z=u+iv\in\mathbb{C}^{+}.$$

Received by the editors September 16, 2021, and, in revised form, October 5, 2022, October 5, 2022, and April 25, 2023.

2020 *Mathematics Subject Classification.* Primary 60F05; Secondary 60H10.

Key words and phrases. Sample covariance matrix, high-dimensional, elliptical distribution, central limit theorem, random matrix theory.

Yanqing Yin is the corresponding author.

The first and third authors were partially supported by NSFC grant 12271065. The second author was partially supported by MOE Tier grant 2018-T2-2-112 and by a MOE Tier 1 grant RG133/18 at Nanyang Technological University. The fourth author was partially supported by a grant A-8000440-00-00 at the National University of Singapore.

As a classical matrix model, the sample covariance matrix has been studied in random matrix theory for a long time, which is defined as follows. Suppose that $\mathbf{y_1}, \mathbf{y_2}, \cdots, \mathbf{y_n}$ are n observations drawn from a centered p dimensional population. Then the sample covariance matrix is $n^{-1}\sum_{i=1}^n \mathbf{y_i}\mathbf{y_i}^*$. In this paper, we will consider more general matrices called separable sample covariance matrices. To be specific, we consider the matrix model

$$\mathbf{S}_n = n^{-1}\mathbf{A}_n\mathbf{X}_n\mathbf{B}_n\mathbf{B}_n^*\mathbf{X}_n^*\mathbf{A}_n^*,$$

where $\mathbf{X}_n = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$ is a $p\times n$ random matrix whose columns are n independent observations from a p dimensional population $\mathbb{X}$ with zero mean vector and unit population covariance matrix while $\mathbf{A}_n$ is a $p \times p$ matrix and $\mathbf{B}_n$ an $n \times n$ matrix. The separable covariance model describes a process where the time correlation does not depend on the spatial location and the spatial correlation does not depend on time, i.e. there is no space-time interaction and this model has applications in many fields such as wireless communications [17] and spatio-temporal statistics [11, 13]. Furthermore, it includes many covariance type matrices that have been well studied in random matrix theory as special cases.

The existing literature on this model mainly considers the case where the population follows the independent component structure (ICS). That is, the components of $\mathbb{X}$ are assumed to be independent. [18] firstly obtained the limiting spectral distribution (LSD) of $\mathbf{S}_n$ under ICS. It was shown that, unlike the usual sample covariance case where $\mathbf{B}_n = \mathbf{I}_n$ (here and throughout this paper, $\mathbf{I}_n$ stands for the n-dimensional identity matrix), the LSD of $\mathbf{S}_n$ is more complex and proved to be the solution of a system of equations. Then [14] proved the no eigenvalue outside result when $\mathbf{B}_n$ is diagonal with non-negative entries. The central limit theorems (CLT) for linear spectral statistics (LSS) were studied in [4] and [12].

However, this assumption excludes lots of important populations such as the elliptical populations. A p-dimensional random vector $\mathbf{x}$ follows an elliptical distribution if and only if it has a stochastic representation:

$$\mathbf{x} = \rho\mathbf{A}\mathbf{u} + \mu, \tag{1.1}$$

where $\mathrm{E}\mathbf{x} = \mu$, the non-random p by p matrix $\mathbf{A}$ satisfies $\mathrm{rank}(\mathbf{A}) = p$, $\rho \geq 0$ is a random variable being the radius of $\mathbf{x}$, and $\mathbf{u}$ is the p-dimensional random direction independent of ρ and being uniformly distributed on the unit sphere S^{p-1} in $\mathbb{C}^p(\mathbb{R}^p)$, denoted by $\mathbf{u} \sim U_{\mathbb{C}(\mathbb{R})}(S^{p-1})$ in the sequel. In view of its abilities in describing heavy tails and tail dependence among components of a population, the family of elliptical distributions has found its applications in lots of scientific areas, including but not limited to statistics and economics. Many commonly used distributions, such as multivariate F distribution, multivariate Pearson type II distributions, power exponential distributions, and multivariate Kotz-type distributions, belong to the family of elliptical distributions. See, for instance [6].

Now, let us introduce the model that will be investigated in this paper.

Definition 1.1. The matrix $\mathbf{S}_n = \frac{1}{n}\mathbf{A}_n\mathbf{X}_n\mathbf{B}_n\mathbf{B}_n^*\mathbf{X}_n^*\mathbf{A}_n^*$ is defined as the separable elliptical sample covariance matrix if the following conditions are satisfied:

(a) The columns of $\mathbf{X}_n$ follow the elliptical distribution. That is, $\mathbf{x}_j = \rho_j\mathbf{u_j}, 1 \leq j \leq n$, where ρ_j's are independent and identically distributed (i.i.d.) random variables with $\mathrm{E}\,\rho_1^2 = p$ and $\mathrm{E}\left(\rho_1^4\right) = p^2 + \tau p + o(p)$ and $\mathbf{u}_j \overset{i.i.d.}{\sim}$

$U_{\mathbb{C}(\mathbb{R})}(S^{p-1})$. Furthermore,

$$\sup_p \mathrm{E}\,|(\rho^2 - p)/\sqrt{p}|^{2+\varepsilon} < \infty$$

for some positive constant ε;

(b) $\mathbf{A}_n$ is a $p \times p$ matrix and $\mathbf{B}_n$ is an $n \times n$ matrix, both are non-random and have uniformly bounded spectral norms;

(c) As $n \to \infty$, the ESDs of $\mathbf{\Sigma}_n = \mathbf{A}_n \mathbf{A}_n^*$, $\mathbf{\Phi}_n = \mathbf{B}_n^* \mathbf{B}_n$, denoted by H_{Σ_n} and H_{Φ_n}, converge weakly to two proper distributions H_Σ and H_Φ respectively;

(d) $c_n = p/n \to c \in (0, \infty)$ as $n \to \infty$.

Remark 1.2. We give some explanations on condition (a) of our model. This condition indicates that the column of $\mathbf{X}_n$ follows an elliptical distribution with squared radius variable ρ^2. To make sure that the considered model is identifiable, the expectation of ρ^2 is assumed to be the dimensional parameter p. Also, the normalized squared radius variable $\rho^2/\sqrt{p}$ is assumed to have a finite moment with order slightly higher than 2 and to have a limit variance τ. When $\tau = 2$, each column of $\mathbf{X}_n$ is a standard p dimensional Gaussian vector. We also remark that the deterministic matrix A in (1.1) is absorbed into A_n in our model.

For this model, when $\mathbf{B}_n = \mathbf{I}_n$, the LSD and CLT for LSS were considered in [9] and [8]. To the best of our knowledge, there is no literature on the LSD and CLT for LSS when $\mathbf{B}_n \neq \mathbf{I}_n$. The following is a theorem concerning the LSD of $\mathbf{S}_n$.

Theorem 1.3. *Suppose the assumptions* (a)–(d) *hold. With probability* 1*, as* $n \to \infty$*, the ESD of* $\mathbf{S}_n$ *converges weakly to a non-random probability distribution function* F^{c,H_Σ,H_Φ}*. To be specific, we have*

(1) *If* $H_\Sigma = 1_{[0,\infty)}$ *or* $H_\Phi = 1_{[0,\infty)}$*, then* $F^{c,H_\Sigma,H_\Phi} = 1_{[0,\infty)}$;

(2) *If* $H_\Sigma \neq 1_{[0,\infty)}$ *and* $H_\Phi \neq 1_{[0,\infty)}$*, for each* $z \in \mathbb{C}^+$,

$$\begin{cases} m(z) = -z^{-1}(1-c^{-1}) - z^{-1}c^{-1}\int \frac{1}{1+q_1(z)y}dH_\Phi(y), \\ m(z) = -z^{-1}\int \frac{1}{1+q_2(z)x}dH_\Sigma(x), \\ m(z) = -z^{-1} - c^{-1}q_1(z)q_2(z) \end{cases} \tag{1.2}$$

is viewed as a system of equations for the complex vector $(m(z), q_1(z), q_2(z))$*, then* (1.2) *has a unique solution in the set*

$$U = \{(m(z), q_1(z), q_2(z)) : \Im m(z) > 0, \Im(zq_1(z)) > 0, \Im q_2(z) > 0\}.$$

Also, the Stieltjes transform of F^{c,H_Σ,H_Φ}*, denoted by* $m_F(z)$*, together with the other two functions* $g_1(z)$ *and* $g_2(z)$*, both of which are analytic on* $\mathbb{C}^+$*, is given by this solution.*

Let $F^{c_n,H_{\Sigma_n},H_{\Phi_n}}$ obtained from F^{c,H_Σ,H_Φ} with (c, H_Σ, H_Φ) replaced by $(c_n, H_{\Sigma_n}, H_{\Phi_n})$. Let $\mathbb{S}_n = \frac{1}{n}\mathbf{B}_n^*\mathbf{X}_n^*\mathbf{A}_n^*\mathbf{A}_n\mathbf{X}_n\mathbf{B}_n$. Obviously, one has

$$F^{\mathbb{S}_n}(x) = c_n F^{\mathbf{S}_n}(x) + (1-c_n)1_{[0,\infty)}(x).$$

Denote by $\mathbb{F}^{c,H_\Sigma,H_\Phi}$ the limiting spectral distribution of $F^{\mathbb{S}_n}$ and denote by $\underline{m}(z)$ the Stieltjes transform of $\mathbb{F}^{c,H_\Sigma,H_\Phi}$. Then one finds

$$\mathbb{F}^{c,H_\Sigma,H_\Phi}(x) = cF^{c,H_\Sigma,H_\Phi}(x) + (1-c)1_{[0,\infty)}(x).$$

Likewise, $\mathbb{F}^{c_n,H_{\Sigma_n},H_{\Phi_n}}$ is obtained from $\mathbb{F}^{c,H_\Sigma,H_\Phi}$ with (c, H_Σ, H_Φ) replaced by $(c_n, H_{\Sigma_n}, H_{\Phi_n})$.

Next, we will give a result about the CLT for LSS of $\mathbf{S}_n$. For that purpose, define

$$G_n(x) = F^{\mathbf{S}_n}(x) - F^{c_n, H_{\Sigma_n}, H_{\Phi_n}}(x).$$

Let $\mathbf{e}_j$ be the j-th column of an identity matrix. Denote

$$\mathfrak{G}_{\triangle}(z) = \left(g_1(z)\mathbf{B}_n\mathbf{B}_n^* + \mathbf{I}_n\right)^{-1}, \quad \mathfrak{G}_{\nabla}(z) = \left(g_1(z)\mathbf{\Phi}_n + \mathbf{I}_n\right)^{-1},$$

where $g_1(z)$ is defined in Theorem 1.3 and $\Phi_n = \mathbf{B}_n^*\mathbf{B}_n$. Also, denote $L_1 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \partial\left\{\left[\mathbf{e}_j^T\mathfrak{G}_{\triangle}(z)\mathbf{e}_j\right]^2\right\}/\partial z \doteq \lim_{n\to\infty} L_{1,n}$,

$$L_2 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \partial\left\{\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\mathfrak{G}_{\triangle}(z)\mathbf{e}_j\right\}/\partial z \doteq \lim_{n\to\infty} L_{2,n},$$

and

$$L_3 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \partial\left[\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathfrak{G}_{\nabla}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathfrak{G}_{\nabla}(z)\mathbf{B}_n^*\mathbf{e}_j\right]/\partial z \doteq \lim_{n\to\infty} L_{3,n},$$

where $\mathbf{e}_j$ is the j-th column of an identity matrix. Furthermore, let

$$\begin{aligned} d_1(z) = {} & \frac{c^2 d_3(z)}{z^4 g_1(z)}\int \frac{x}{(1+xg_2(z))^2}dH_\Sigma(x)\left[\int\frac{t^2}{(g_1(z)t+1)^3}dH_\Phi(t)\right. \\ & \left.+\frac{d_2(z)}{zg_2(z)}\int\frac{t}{(g_1(z)t+1)^2}dH_\Phi(t)\right] \\ & +\frac{c}{z^4 g_2(z)}d_2(z)\int\frac{x^2}{(1+xg_2(z))^3}dH_\Sigma(x)\int\frac{t}{(g_1(z)t+1)^2}dH_\Phi(t), \\ d_2(z) = {} & \int\frac{t^2}{(g_1(z)t+1)^2}dH_\Phi(t), \quad d_3(z) = \int\frac{x^2}{(1+xg_2(z))^2}dH_\Sigma(x), \\ d_4(z) = {} & \frac{1}{c}\left(\frac{3}{4}\tau - \frac{5}{2}\right)\frac{\partial(z\underline{m}(z))}{\partial z} + \frac{(\tau-6)L_1}{4c} - \frac{(\tau+2)L_2}{4c} + \frac{(\tau+2)L_3}{4c}, \\ d_5(z) = {} & \frac{5}{4c}\frac{\partial(z\underline{m}(z))}{\partial z} + \frac{3L_1}{4c} + \frac{L_2}{4c} - \frac{L_3}{4c}, \end{aligned} \tag{1.3}$$

and

$$d(z_1,z_2) = \frac{1}{z_1z_2}\frac{z_1g_1(z_1) - z_2g_1(z_2)}{g_2(z_1)-g_2(z_2)}\frac{z_1g_2(z_1)-z_2g_2(z_2)}{g_1(z_1)-g_1(z_2)}.$$

Theorem 1.4. *Suppose the assumptions* (a)–(d) *hold and suppose* $\mathbf{B}_n$ *is invertible. Denote by* $\lambda_1 \geq \cdots \geq \lambda_p$ *the eigenvalues of* $\mathbf{\Sigma}_n$. *Let* $f_1, \cdots, f_\kappa$ *be functions on* $\mathbb{R}$ *analytic on an open interval containing*

$$\left[\liminf_n \lambda_p \lambda_{\min}^{\mathbf{\Phi}_n} I_{(0,1)}(c)\left(1-\sqrt{c}\right)^2, \limsup_n \lambda_1\lambda_{\max}^{\mathbf{\Phi}_n}\left(1+\sqrt{c}\right)^2\right]. \tag{1.4}$$

Then

$$p\left(\int f_1(x)dG_n(x), \cdots, \int f_\kappa(x)dG_n(x)\right) \tag{1.5}$$

converges weakly to a Gaussian vector $(X_{f_1}, \cdots, X_{f_\kappa})$ *with mean*

$$\begin{aligned} \operatorname{E} X_f = &-\frac{2-\beta}{2\pi i}\oint_{\mathcal{C}} \frac{f(z)d_1(z)}{\left(1-cz^{-2}d_2(z)d_3(z)\right)^2}dz \\ &-\frac{1}{2\pi i}\oint_{\mathcal{C}} f(z)d_4(z)dz - \frac{\beta-1}{2\pi i}\oint_{\mathcal{C}} f(z)d_5(z)dz, \end{aligned} \tag{1.6}$$

and covariance function

(1.7)

$$\begin{aligned} &\operatorname{Cov}\Big(X_f, X_g\Big) \\ &= -\frac{3-\beta}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f(z_1)g(z_2)\frac{\partial^2}{\partial z_2 \partial z_1}\int_0^{d(z_1,z_2)} \frac{1}{1-z}dz dz_1 dz_2 \\ &\quad - \lim_{n\to\infty}\frac{\tau-3+\beta}{4\pi^2 p}\sum_{j=1}^n \oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f(z_1)g(z_2)\frac{\partial^2 \mathbf{e}_j^T \mathbb{G}_\triangle(z_1)\mathbf{e}_j\mathbf{e}_j^T\mathbb{G}_\triangle(z_2)\mathbf{e}_j}{\partial z_2 \partial z_1}dz_1 dz_2, \end{aligned}$$

where $f, g \in \{f_1, \cdots, f_\kappa\}$, $\beta = 1$ *when* $\mathbf{u} \sim U_{\mathbb{R}}(S^{p-1})$ *and* $\beta = 2$ *when* $\mathbf{u} \sim U_{\mathbb{C}}(S^{p-1})$. *The contours in* (1.6) *and* (1.7) *(two contours in* (1.7)*, which we may assume to be non-overlapping) are closed and are taken in the positive direction in the complex plane, each enclosing the support of* $F^{c,H_{\mathbf{\Sigma}},H_{\mathbf{\Phi}}}$.

Remark 1.5. To fully express such kind of CLT above, it is necessary to invoke three separate limits, namely L_1, L_2 and L_3 defined above. Generally speaking, these limits may not always exist. These kinds of assumptions are due to the non-negligible influences of the population eigenvectors and are prevalent in the descriptions of CLT for LSS under non-Gaussian populations. It is worth noting that in cases where the underlying distribution is Gaussian or under certain diagonal assumptions, these limits naturally exist. In practice, we can treat the sequences $L_{1,n}$, $L_{2,n}$ and $L_{3,n}$ as parameters in the expressions of "limiting" means and variances of specific LSS. Then, one finds that the normalized LSS

$$\frac{\left(p\int f(x)dG_n(x) - \widetilde{\operatorname{E} x_f}\right)}{\sqrt{\widetilde{\operatorname{Cov}}(x_f, x_f)}},$$

where $\widetilde{\operatorname{E} x_f}$ and $\widetilde{\operatorname{Cov}}(x_f, x_f)$ are obtained by replacing L_1, L_2 and L_3 in $\operatorname{E} x_f$ and $\operatorname{Cov}(x_f, x_f)$ with $L_{1,n}, L_{2,n}$ and $L_{3,n}$, will convergence weakly to a standard Gaussian variable. The joint distributions can be obtained in a similar fashion.

Remark 1.6. The invertible condition of $\mathbf{B}_n$ is just for technical reason. It is critical for our proof strategy but is not necessary for the CLT to hold.

Remark 1.7. It is seen that the first term in (1.6) will disappear in the complex case ($\beta = 2$), and the same phenomenon was also observed by [4] for populations under ICS. The rest of the terms related to $d_4(z)$ are due to the dependence in the elliptical population case.

Remark 1.8. We mainly consider elliptical populations in this paper due to their widely applications in practice. It will be seen in the proof we use the fact that $\mathbf{u}$ can be written as a normalized Gaussian vector. However, if we assume B_n is diagonal and the vector $\mathbf{u}$ is a normalized random vector whose entries are non-Gaussian but with the same fourth moment $\mu_4 < \infty$, one can obtain the same theoretical results on LSD and CLT for LSS by taking the same procedures as in this paper.

It is worth noting that the proof of the above theorem is very different from the one in [4]. Under ICS, in [4] the CLT for LSS of separable covariance matrices is derived by two steps. The first step is to prove the result when $\mathbf{X}_n$ are Gaussian (also a special case of our theorem when $\tau = 2$). The second step is to compare the characteristic functions of LSSs under the Gaussian case and general case (under the assumption that the fourth moment of the underlying distribution equals to 3). However, this strategy is no longer valid under elliptical distribution case. Thus, in this paper, we need some new strategies to establish the above theorem. To be specific, by using the properties of elliptical distributions and the orthogonal property of Gaussian random vector, we find that the difference between separable sample covariance matrices under elliptical population and Gaussian population is on account of the influences of two random diagonal matrices. Thus, by the expansion of the random diagonal matrix, we successfully decompose the objective Stieltjes transform into several parts and point out their contributions. This strategy is also different with the one used in [9], which is based on a newly established formula for quadratic forms under elliptical distribution when $\mathbf{B} = \mathbf{I}_n$.

The rest of this paper is organized as follows. Section 2 contains some applications of main results. The main theorems are proved in Section 3. Some useful lemmas are presented in Section 4. Appendix collects some technical proofs.

2. Applications

In this section, we give some applications of our main results. The first application is on realized sample covariance matrices and the second one is on Kronecker channel model.

2.1. **Application on realized sample covariance matrices.**

2.1.1. *The realized sample covariance matrices for diffusion processes.* Consider the diffusion processes, which are widely used in economics to model financial asset price processes. Suppose $\mathbf{x}_t^{(j)}, 1 \le j \le p$, are the log price processes of p stocks. Denote $\mathbf{x}_t = \left(\mathbf{x}_t^{(1)}, \cdots, \mathbf{x}_t^{(p)}\right)^T$. A widely used model for $\mathbf{x}_t$ is

$$d\mathbf{x}_t = \mu_t dt + \Theta_t d\mathbf{w}_t. \tag{2.1}$$

Here μ_t and $\mathbf{w}_t$ are p-dimensional drift process and p-dimensional standard Brownian motion respectively. Θ_t, a $p \times p$ matrix for any t, is called the co-volatility process. Under the situation that one can observe the processes $\mathbf{x}_t$ at high frequency synchronously, one is usually interested in the so-called integrated covariance (ICV) matrix defined as

$$\Sigma^{ICV} := \int_0^1 \Theta_t \Theta_t^T dt,$$

as well as its estimator, realized covariance (RCV) matrix defined as

$$\Sigma^{RCV} := \sum_{l=1}^n \Delta\mathbf{x}_l \left(\Delta\mathbf{x}_l\right)^T,$$

where $\Delta \mathbf{x}_l := \begin{pmatrix} \mathbf{x}^{(1)}_{\tau_{n,l}} - \mathbf{x}^{(1)}_{\tau_{n,l-1}} \\ \vdots \\ \mathbf{x}^{(p)}_{\tau_{n,l}} - \mathbf{x}^{(p)}_{\tau_{n,l-1}} \end{pmatrix}$, with $\tau_{n,l}, 1 \le l \le n$ being the observed time points.

It is explored in [19] that the LSD of Σ^{RCV} depends not only on the LSD of ICV matrix, but also on how the co-volatility process evolves over time. Furthermore, under the assumption that the process belongs to class $\mathcal{C}$ to be defined below, and some other assumptions, [19] demonstrates clearly how the time-variability of the co-volatility process affects the LSD of Σ^{RCV}.

Definition 2.1. Suppose that the process $\mathbf{x}_t$ is a p-dimensional process satisfying (2.1) and Θ_t is c$\grave{a}$dl$\grave{a}$g. If, almost surely, there exist $(\gamma_t) \in D([0,1]; R)$ and a $p \times p$ matrix Λ_n satisfying $\Theta_t = \gamma_t \Lambda_n$ and $\operatorname{tr}(\Lambda_n \Lambda_n^T) = p$, where $D(I; S)$ stands for the space of c$\grave{a}$dl$\grave{a}$g functions from I to S, then we say that $\mathbf{x}_t$ belongs to class $\mathcal{C}$.

The next question is, how the co-volatility process and the LSD of ICV matrix affect the second order limits of the eigenvalues of RCV matrix. Such kind of results is useful when making statistical inference on ICV matrix. Here, by applying our main results of this paper, we shall give a theorem on the LSS of RCV matrix. To the specific, given functions $f_1, \cdots, f_\kappa$ on $\mathbb{R}$, we will derive the fluctuation properties of the linear spectral statistics specified by these test functions. We first present some necessary conditions.

Assume that the diffusion process $\mathbf{x}_t$ belongs to class $\mathcal{C}$, the drift process $\mu_t \equiv 0$ and $\tau_{n,l}$'s and γ_t are both nonrandom and independent of $\mathbf{w}_t$. Further assume that

(C1) $\breve{\Sigma}_n = \Lambda_n \Lambda_n^T$ satisfies assumptions (c) and (d) in Definition 1.1. And there exists $C_1 < \infty$ such that $|\gamma_t| \le C_1$ for all $t \in [0,1]$.
(C2) $c_n = p/n \to c \in (0, \infty)$ as $n \to \infty$;
(C3) The observation times $\tau_{n,l}$ are independent of $\mathbf{x}_t$; moreover, there exists $\kappa < \infty$ such that the observation durations $\Delta\tau_{n,l} := \tau_{n,l} - \tau_{n,l-1}$ satisfy

$$\max_n \max_{l=1,\cdots,n} (n \Delta \tau_{n,l}) \le \kappa;$$

additionally, almost surely, there exists a process $\nu_s \in C([0,1]; \mathbb{R}_+)$ such that

$$\tau_{n,[ns]} \to \Upsilon_s := \int_0^s \nu_r dr$$

as $n \to \infty$ for all $0 \le s \le 1$, where for any x, $[x]$ stands for its integer part.

The above assumption (C3) on $\tau_{n,l}$ indicates that as $n \to \infty$, the largest observation duration should tend to 0 with rate n^{-1}. When the observation times are equally spaced, this assumption is trivially satisfied. We then have the following results.

Lemma 2.2. *Under* (C1)-(C3)*, as $n \to \infty$, with probability* 1*, the empirical spectral distribution $F^{\Sigma^{RCV}}$ converges weakly to a non-random probability distribution function $F^{c, H_{\breve{\Sigma}}, \omega}$ as specified in Theorem* 1.3 *with $\int \frac{1}{1+q_1(z)y} dH_\Phi(y)$ replaced by $\int_0^1 \frac{1}{1+q_1(z)(\gamma \Upsilon_s)^2 \nu_s} ds$.*

Lemma 2.2 is consistent with Proposition 5. in [19].

Furthermore, let $F^{p/n,H_{\breve{\Sigma}_n},H_{\Phi_n}}$ denote the one obtained from $F^{c,H_{\breve{\Sigma}},\omega}$ with $\left(c,H_{\breve{\Sigma}}\right)$ replaced by $\left(c_n,H_{\breve{\Sigma}_n}\right)$ and $\int_0^1 \frac{1}{1+q_1(z)(\gamma\Upsilon_s)^2\nu_s}ds$ replaced by $\int \frac{1}{1+q_1(z)x}dH_{\Phi_n}(x)$ with H_{Φ_n} the empirical distribution of $\{\int_{\tau_{n,l-1}}^{\tau_{n,l}} n\gamma_t^2dt\}_{l=1}^n$. Then we have the CLT.

Theorem 2.3. *Assume the following additional condition*

(C4) *Denote $\lambda_1 \geq \cdots \geq \lambda_p$ the eigenvalues of $\breve{\Sigma}$. $f_1, \cdots, f_\kappa$ are analytic on an open interval containing*

$$\left[\liminf_n \lambda_p \left(1-\sqrt{c}\right)^2 I_{(0,1)}(c), \limsup_n \kappa C_1^2 \lambda_1 \left(1+\sqrt{c}\right)^2\right].$$

Consider the process

$$\mathcal{G}_n := p\left(\int f_1(x)dG_n(x), \cdots, \int f_\kappa(x)dG_n(x)\right),$$

where

$$G_n(x) = F^{\Sigma^{RCV}}(x) - F^{p/n, H_{\breve{\Sigma}_n}, H_{\Phi_n}}(x).$$

Under assumptions (C1)–(C4), *as $n \to \infty$, $\mathcal{G}_n$ converges weakly to a Gaussian vector $(X_{f_1}, \cdots, X_{f_\kappa})$ with mean*

$$\mathrm{E}\, X_f = -\frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{f(z)\widetilde{d}_1(z)}{\left(1 - cz^{-2}\widetilde{d}_2(z)\widetilde{d}_3(z)\right)^2} dz \tag{2.2}$$

and covariance function

$$\mathrm{Cov}\Big(X_f, X_g\Big) = -\frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f(z_1)g(z_2)\frac{\partial^2}{\partial z_2 \partial z_1}\int_0^{\widetilde{d}(z_1,z_2)} \frac{1}{1-z} dz dz_1 dz_2, \tag{2.3}$$

where $f, g \in \{f_1, \cdots, f_\kappa\}$, $\widetilde{d}_1(z), \widetilde{d}_2(z), \widetilde{d}_3(z)$ and $\widetilde{d}(z_1, z_2)$ are specified in (1.3) *with t replaced by $(\gamma_{\Upsilon_s})^2\nu_s$ and $dH_\Phi(t)$ replaced by ds.*

2.1.2. *Simulation studies.* We demonstrate in this subsection some simulations to investigate the influence of the time variability of the covolatility process on the fluctuation of the LSS. It will be shown that even sharing the same ICV matrix, the LSS of different RCV matrices can obey very different CLTs, depending on the time variability of the covolatility process.

Following [19], we assume in the simulation below that $\Lambda_n = \mathbf{I}_p$. That is, $\mathbf{x}_t$ follows (2.1) with γ_t a deterministic (scalar) process, and $\mathbf{W}_t$ a p-dimensional standard Brownian motion. The observation times are taken to be equidistant: $\tau_{n,l} = l/n, l = 0, \cdots, n$. We set two designs as follows.

- Model I: The volatility path follows piecewise constants. More specifically, γ_t is set to be

  $$\gamma_t = \begin{cases} \sqrt{a} \times 10^{-2}, & t \in [0, 1/4) \cup [3/4, 1], \\ \sqrt{b} \times 10^{-2}, & t \in [1/4, 3/4), \end{cases} \quad \text{where } a + b = 8.$$

  When a and b take different values, the covolatility processes are different, but they share the same ICV matrix.
- Model II: The volatility processes have continuous sample paths. In this case, γ_t is set to be

  $$\gamma_t = \sqrt{0.0004 + d\cos(2\pi t)}, \quad t \in [0, 1].$$

  When d takes different values, the covolatility processes are different, but they share the same ICV matrix.

For each model, we compare the empirical distributions of two LSSs of RCV matrices when the parameters take different values. The simulation results are presented in Figs. 1-4.

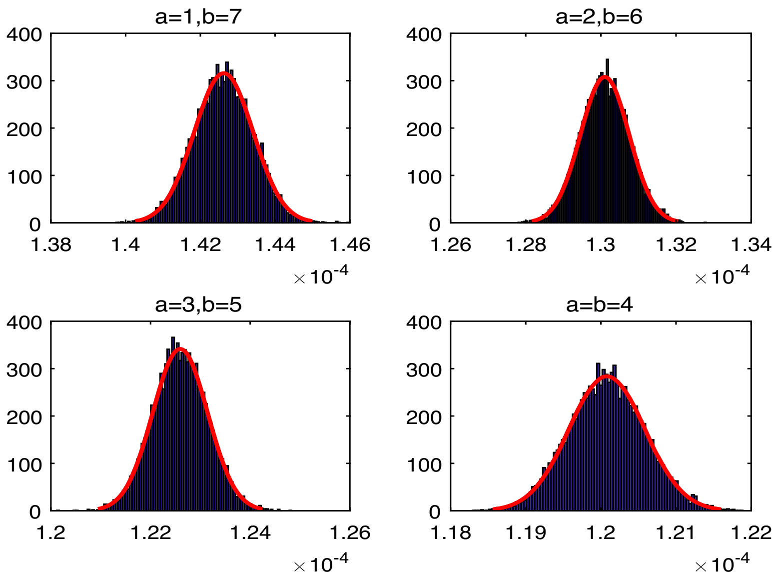

FIGURE 1. Empirical distributions of traces of square of RCV matrices when a and b take different values under model I. $p = 500$, $n = 1000$. The results are based on repetitions of 10000 times.

It can be seen that the time variability of the covolatility process could affect the fluctuation of LSS of RCV matrix heavily. The CLT always holds but the limiting parameters may be different.

2.2. **Application on Kronecker channel model.** In wireless communication, the Kronecker-based stochastic model (KBSM) is an important channel model for massive multiple-input multiple-output (MIMO) systems, see [15]. In KBSM, the channel matrix $\mathbf{H}$ can be expressed as $\mathbf{H}_n = \mathbf{A}_n\mathbf{X}_n\mathbf{B}_n$, where $\mathbf{X}_n$ is a $p\times n$ matrix with i.i.d. complex entries with zero-mean and unit-variance, $\mathbf{A}_n\mathbf{A}_n^*$ and $\mathbf{B}_n^*\mathbf{B}_n$ are overall spatial correlation matrices at the receiver and transmitter, respectively. We are interested in the capacity of the channel, which is characterized by Shannon's mutual information $I(\rho)$ between the n-dimensional input and p-dimensional output signals. Assuming that the input signal is circularly symmetric complex Gaussian with covariance matrix $(\rho/n)\mathbf{I}_n$, $\rho \in (0,\infty)$, the mutual information is then given by

$$I(\rho) = \log\det(\frac{\rho}{n}\mathbf{H}_n\mathbf{H}_n^* + \mathbf{I}_p).$$

The asymptotic behaviors of $I(\rho)$ under the framework that n and p tend to infinity at the same rate have been considered by many authors. For instance, when $\mathbf{A}_n = \mathbf{I}_p, \mathbf{B} = \mathbf{I}_n$ and the entries of $\mathbf{X}_n$ follow normal distribution, the fluctuation of $I(\rho)$ is established in [10]. For non-Gaussian matrix $\mathbf{X}_n$, we refer the reader to [2, 5]. The case where $\mathbf{A}_n$ and $\mathbf{B}_n$ are both diagonal is considered in [7]. Recently, [8] obtained the fluctuation of $I(\rho)$ when $\mathbf{B}_n = \mathbf{I}_n$ and the columns of $\mathbf{X}_n$ follow

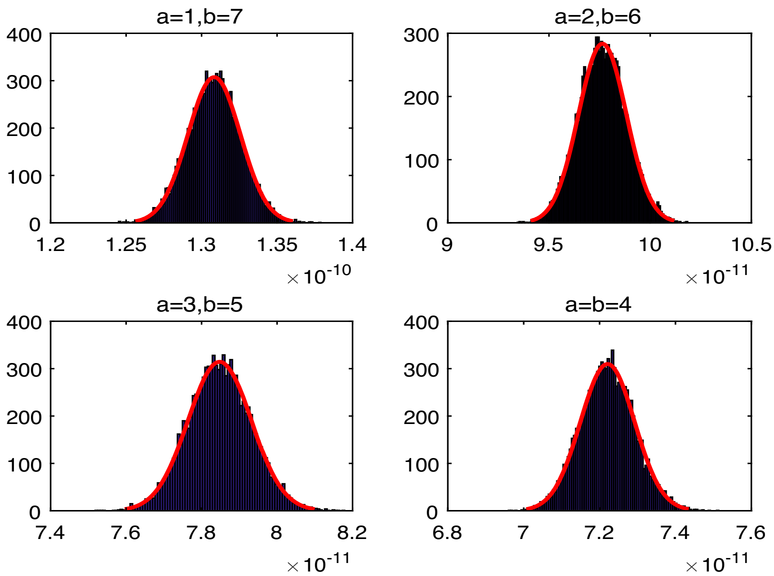

FIGURE 2. Empirical distributions of traces of the fourth-power of RCV matrices when a and b take different values under model I. $p = 500$, $n = 1000$. The results are based on repetitions of 10000 times.

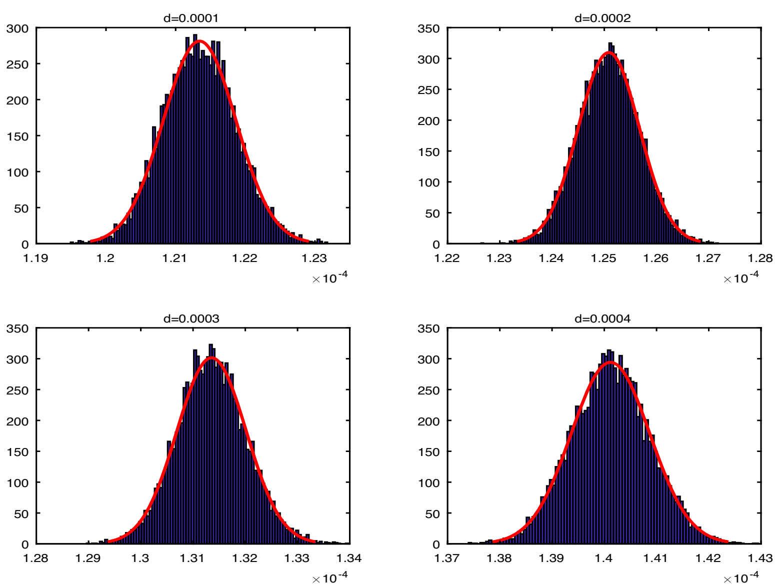

FIGURE 3. Empirical distributions of traces of the fourth-power of RCV matrices when d takes different values under model II. $p = 500$, $n = 1000$. The results are based on repetitions of 10000 times.

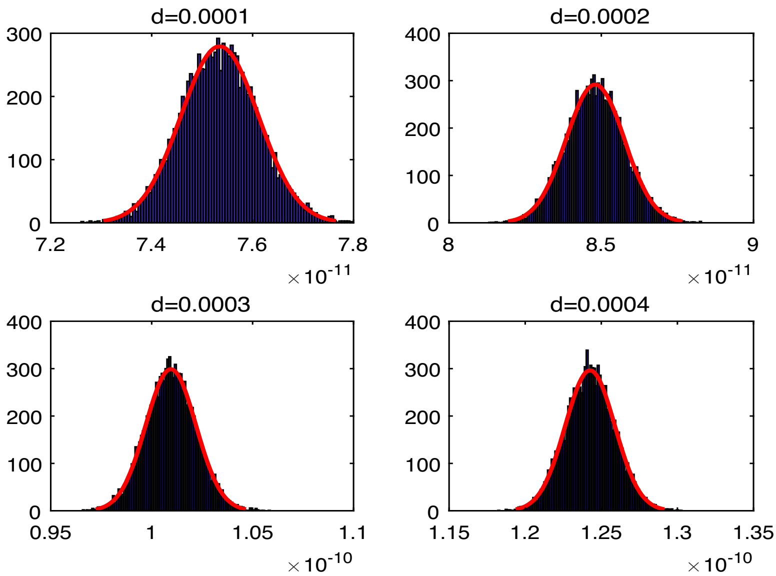

FIGURE 4. Empirical distributions of traces of the fourth-power of RCV matrices when d takes different values under model II. $p = 500$, $n = 1000$. The results are based on repetitions of 10000 times.

elliptical distribution. In this section, by virtue of our theoretical results, we shall give a theorem on $I(\rho)$ without the assumption that $\mathbf{B}_n = \mathbf{I}_n$. This is a significant improvement both on theoretical and practice as we allow more general correlation structure between the entries of the channel matrix $\mathbf{H}_n$. Formally, we have Theorem 2.4.

Theorem 2.4. *Assume that* (a)-(d) *are satisfied with* $\mathbf{u}_j \overset{i.i.d.}{\sim} U_{\mathbb{C}}(S^{p-1})$, *and define the central term*

$$C(\rho) := p \int \log(1 + \rho x) dF^{c_n, H_{\Sigma_n}, H_{\Phi_n}}.$$

We have

$$I(\rho) - C(\rho) \xrightarrow{D} N(\mu_{\rho,H_\Sigma,H_\Phi}, \sigma^2_{\rho,H_\Sigma,H_\Phi}),$$

where the limiting mean

$$\mu_{\rho,H_\Sigma,H_\Phi} = -\frac{1}{2\pi i} \oint_{\mathcal{C}} \log(1 + \rho z) d_4(z) dz - \frac{1}{2\pi i} \oint_{\mathcal{C}} \log(1 + \rho z) d_5(z) dz,$$

and

$$\begin{aligned}
\sigma^2_{\rho,H_\Sigma,H_\Phi} = &-\frac{1}{4\pi^2} \oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \log(1 + \rho z_1) \log(1 + \rho z_2) \frac{\partial^2}{\partial z_2 \partial z_1} \int_0^{d(z_1,z_2)} \frac{1}{1 - z} dz dz_1 dz_2 \\
&- \lim_{n\to\infty} \frac{\tau - 1}{4\pi^2 p} \sum_{j=1}^{n} \oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \log(1 + \rho z_1) \log(1 + \rho z_2) \frac{\partial}{\partial z_1} \left(\mathbf{e}_j^T \mathbb{G}_\triangle(z_1) \mathbf{e}_j \right) \\
&\cdot \frac{\partial}{\partial z_2} \left(\mathbf{e}_j^T \mathbb{G}_\triangle(z_2) \mathbf{e}_j \right) dz_1 dz_2.
\end{aligned}$$

This theorem is an easy consequence of Theorem 1.4, thus we omit its proof.

3. Proof of theoretical results

This section is devoted to the proofs of main theorems. The proof of Theorem 1.3 will be presented in Section 3.1. And the rest part of this section will give the proof of Theorem 1.4.

3.1. **Proof of Theorem 1.3.** We first deal with the proof of the theorem on LSD.

By definition, the random vectors $\mathbf{u}_j$ can be expressed as $\mathbf{u}_j = \frac{\mathbf{y}_j}{\|\mathbf{y}_j\|}$ with $\mathbf{y}_j \sim N(0, \mathbf{I}_p)$. Thus, we have $\mathbf{x}_j = \frac{\rho_j \mathbf{y}_j}{\|\mathbf{y}_j\|}$. Let $\mathbf{Y}_n = (\mathbf{y}_1, \cdots, \mathbf{y}_n)$, so

$$\mathbf{X}_n = \mathbf{Y}_n \operatorname{Diag}\left(\frac{\rho_1}{\|\mathbf{y}_1\|}, \cdots, \frac{\rho_n}{\|\mathbf{y}_n\|}\right) \triangleq \mathbf{Y}_n \mathbb{C}_n.$$

Note the spectral decomposition of $\mathbf{A}_n^*\mathbf{A}_n = \mathbf{U}_n^* \mathbf{\Lambda}_n \mathbf{U}_n$, where $\mathbf{\Lambda}_n = \operatorname{Diag}(\lambda_1, \cdots, \lambda_p)$ with $\lambda_j, 1 \le j \le p$ being the eigenvalues of $\mathbf{A}_n^*\mathbf{A}_n$. By the orthogonal property of Gaussian random vector, we shall redefine and investigate the properties of

$$\mathbb{S}_n \overset{d}{=} \frac{1}{n}\mathbf{B}_n^*\mathbf{X}_n^*\mathbf{\Lambda}_n\mathbf{X}_n\mathbf{B}_n = \frac{1}{n}\mathbf{B}_n^*\mathbb{C}_n\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbb{C}_n\mathbf{B}_n.$$

The proof of the theorem on LSD is mainly based on comparing the eigenvalues of $\mathbb{S}_n$ with those of $\frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n$ since the LSD of the latter one is known to us. Notice that by Lemma 4.2, $\|\mathbb{C}_n^2 - \mathbf{I}\| \to 0, a.s.$ Thus we have

$$\begin{aligned}
&\|\mathbb{S}_n - \frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n\| \\
\le &\|\mathbb{C}_n - \mathbf{I}\|\|\frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n\|\|\mathbb{C}_n\| + \|\mathbb{C}_n - \mathbf{I}\|\|\frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n\| \xrightarrow{a.s.} 0.
\end{aligned}$$

Then this theorem follows from Weyl's inequality and Theorem 1.2.1 of [18].

Then, we shall present the proof of the theorem on CLT for LSS. The main strategy is to apply the CLT for martingale.

3.2. **The proof of Theorem 1.4.** Mow, we shall present the proof of the theorem on CLT for LSS. The proof is based on the procedure developed in [2] and the main strategy is to apply the CLT for martingale.

Rewrite for $z \in \mathbb{C}^+$, $M_n(z) = n[\underline{m}_n(z) - \underline{m}_n^0(z)]$, where

$$\begin{aligned}
m_n(z) = m_{F^{\mathbf{S}_n}}(z),\ m_n^0(z) = m_{F^{c_n, H_{\Sigma_n}, H_{\Phi_n}}}(z), \\
\underline{m}_n(z) = m_{F^{\mathbb{S}_n}}(z),\ \underline{m}_n^0(z) = m_{\mathbb{F}^{c_n, H_{\Sigma_n}, H_{\Phi_n}}}(z).
\end{aligned}$$

Moreover $\left(m_n^0(z), g_{1n}^0(z), g_{2n}^0(z)\right)$ satisfies equation (1.2) with (c, H_1, H_2) replaced by $(c_n, H_{\Sigma_n}, H_{\Phi_n})$. In other words, we have

$$\begin{cases}
\underline{m}_n^0(z) & = -z^{-1}\int \frac{1}{1+g_{1n}^0(z)y} dH_\Phi(y), \\
m_n^0(z) & = -z^{-1}\int \frac{1}{1+g_{2n}^0(z)x} dH_\Sigma(x), \\
m_n^0(z) & = -z^{-1} - c_n^{-1} g_{1n}^0(z) g_{2n}^0(z).
\end{cases} \tag{3.1}$$

Furthermore, one has

$$zg_{1n}^0(z) = -c_n \int \frac{x}{1+g_{2n}^0(z)x} dH_\Sigma(x), \tag{3.2}$$

$$zg_{2n}^0(z) = -\int \frac{y}{1+g_{1n}^0(z)y} dH_\Phi(y). \tag{3.3}$$

Let v_0 be any positive number, $x_r \in (\limsup_n \lambda_1 \lambda_{\max}^{\mathbf{\Phi}_n} (1+\sqrt{c})^2, \infty)$. Let x_l be any negative number if the left end point of interval (1.4) is zero. Otherwise choose $x_l \in \left(0, \liminf_n \lambda_p \lambda_{\min}^{\mathbf{\Phi}_n} I_{(0,1)}(c)(1-\sqrt{c})^2\right)$. Let $\mathcal{C}_u = \{x+iv_0 : x \in [x_l, x_r]\}$. Define the contour $\mathcal{C} = \{x_l + iv : v \in [0, v_0]\} \cup \mathcal{C}_u \cup \{x_r + iv : v \in [0, v_0]\}$. To avoid dealing with the small $\Im z$, we truncate $M_n(z)$ on a contour $\mathcal{C}$ of the complex plane. We define now the subset $\mathcal{C}_n$ of $\mathcal{C}$ on which $M_n(\cdot)$ agrees with $\widehat{M}_n(\cdot)$. Choose a sequence $\{\varepsilon_n\}$ decreasing to zero satisfying for some $\alpha \in (0,1)$, $\varepsilon_n \geq n^{-\alpha}$. Let $\mathcal{C}_l = \left\{x_l + iv : v \in [n^{-1}\varepsilon_n, v_0]\right\}$ and $\mathcal{C}_r = \left\{x_r + iv : v \in [n^{-1}\varepsilon_n, v_0]\right\}$. Then $\mathcal{C}_n = \mathcal{C}_l \cup \mathcal{C}_u \cup \mathcal{C}_r$. For $z = x+iv$, the process $\widehat{M}_n(\cdot)$ can now be defined as

$$\widehat{M}_n(\cdot) = \begin{cases} M_n(z), & \text{for } z \in \mathcal{C}_n, \\ M_n(x_l + in^{-1}\varepsilon_n), & \text{for } x = x_l, v \in [0, n^{-1}\varepsilon_n], \\ M_n(x_r + in^{-1}\varepsilon_n), & \text{for } x = x_r, v \in [0, n^{-1}\varepsilon_n]. \end{cases} \tag{3.4}$$

We are going to present the following CLT for the truncated process $\widehat{M}_n(\cdot)$.

Lemma 3.1. *Under the assumptions of Theorem* 1.4, $\widehat{M}_n(z)$ *converges weakly to a two-dimensional Gaussian process* $M(\cdot)$ *satisfying for* $z \in \mathcal{C}$,

$$\mathrm{E}\, M(z) = (2-\beta)\frac{d_1(z)}{\left(1 - cz^{-2}d_2(z)d_3(z)\right)^2} + d_4(z).$$

And for $z_1, z_2 \in \mathcal{C} \cup \overline{\mathcal{C}}$ *with* $\overline{\mathcal{C}} = \{\bar{z} : z \in \mathcal{C}\}$,

$$\begin{aligned} \mathrm{Cov}\left(M(z_1), M(z_2)\right) &= (3-\beta)\frac{\partial^2}{\partial z_2 \partial z_1}\int_0^{d(z_1,z_2)} \frac{1}{1-z}dz \\ &\quad + \lim_{n\to\infty} \frac{\tau - 3 + \beta}{p}\sum_{j=1}^n \frac{\partial}{\partial z_1}\left(\mathbf{e}_j^T \mathbb{G}_\triangle(z_1)\mathbf{e}_j\right)\frac{\partial}{\partial z_2}\left(\mathbf{e}_j^T \mathbb{G}_\triangle(z_2)\mathbf{e}_j\right). \end{aligned}$$

It will be shown in the next subsection that Theorem 1.4 will directly follow once Lemma 3.1 is verified.

To prove Lemma 3.1, we decompose the Stieltjes transform into several parts. One of these parts is coincident with the ICS case and the other parts are due to the dependence in elliptical distribution. Denote $\mathbf{M}_n = \mathbf{B}_n^{-1}\mathbb{C}_n^{-1}\mathbf{B}_n$,

$$\mathbf{J}(z) = \frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n - z\mathbf{I}_n,\ \mathbf{H}(z) = \frac{1}{n}\mathbf{B}_n^*\mathbf{Y}_n^*\mathbf{\Lambda}_n\mathbf{Y}_n\mathbf{B}_n - z\mathbf{M}_n^*\mathbf{M}_n.$$

Notice that for any invertible matrices $\mathbf{A}_0$ and $\mathbf{B}_0$,

$$\mathbf{A}_0^{-1} - \mathbf{B}_0^{-1} = \mathbf{A}_0^{-1}(\mathbf{B}_0 - \mathbf{A}_0)\mathbf{B}_0^{-1}. \tag{3.5}$$

Then we obtain that

$$\begin{aligned} &\mathbf{H}^{-1}(z) - \mathbf{J}^{-1}(z) \\ &\quad = z\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\mathbf{J}^{-1}(z) + z^2\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^2\mathbf{H}^{-1}(z) \\ &\quad = z\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\mathbf{J}^{-1}(z) + z^2\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z) \\ &\qquad + z^3\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^3\mathbf{H}^{-1}(z). \end{aligned}$$

Hence, it yields that

$$
\begin{aligned}
nm_{F^{\mathbb{S}_n}}(z) &= \operatorname{tr}(\mathbb{S}_n - z\mathbf{I}_n)^{-1} = \operatorname{tr}\left(\mathbf{H}^{-1}\right) + \operatorname{tr}\left(\mathbf{H}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) \\
&= \operatorname{tr}\left(\mathbf{J}^{-1}(z)\right) + \operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) + z\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right) \\
&\quad + z\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right]^2 + z^2\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z)\right] \\
&\quad + z^2\operatorname{tr}\left\{\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right]^2\mathbf{H}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right\} \\
&\quad + z^3\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^3\mathbf{H}^{-1}(z)\right] \\
&=: \mathcal{I}(z) + \mathcal{II}(z),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{I}(z) &= \operatorname{tr}\left(\mathbf{J}^{-1}(z)\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right) + \operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) \\
&\quad - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) \\
&\quad + z\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right) \\
&\quad - z\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{II}(z) &= \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right) + \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) + z\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right) \\
&\quad + z\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right]^2 + z^2\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z)\right] \\
&\quad + z^2\operatorname{tr}\left\{\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right]^2\mathbf{H}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right\} \\
&\quad + z^3\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)^3\mathbf{H}^{-1}(z)\right].
\end{aligned}
$$

Then we have that

$$M_n(z) = \mathcal{I}(z) + \mathcal{II}(z) - n\underline{m}_n^0(z).$$

As will be shown, $\mathcal{I}(z)$ determines the Gaussian fluctuation of $M_n(z)$ and $\mathcal{II}(z) - n\underline{m}_n^0(z)$ yields the expectation of $M_n(z)$.

In the rest of this subsection, we will firstly truncate the variable, then show that $\mathcal{II}(z) - n\underline{m}_n^0(z)$ converges to a limit and find the limiting distribution of $\mathcal{I}(z)$. Finally, we finish the proof of Theorem 1.4.

3.2.1. *Truncation.* We first truncate the variable at a proper order to control the high order moments. The procedure is similar to Section 5.2 in [9]. However, we give the sketch of this step here for the convenience of readers. Firstly, by the assumption $\sup_p \operatorname{E}|(\rho^2 - p)/\sqrt{p}|^{2+\varepsilon} < \infty$, for some positive constant ε, we shall choose a sequence δ_n with

$$\delta_n \to 0,\ \ \delta_n p^{1/2} \to \infty.$$

Then let $\hat{\rho}_j = \rho_j I(|\rho_j^2 - p| \le \delta_n p),\ j = 1, \cdots, n$, We have

$$\operatorname{P}\left((\rho_1, \cdots, \rho_n) \neq (\hat{\rho}_1, \cdots, \hat{\rho}_n)\right) \le n\delta_n^{-2}p^{-2}\operatorname{E}\left(\left(\rho_j^2 - p\right)^2 I\left(|\rho_j^2 - p| > \delta_n p\right)\right) \to 0.$$

Secondly, let $\sigma_n^2 = \operatorname{E}(\hat{\rho}_1^2)/p$ and define $\tilde{\rho}_j = \hat{\rho}_j = \hat{\rho}_j/\sigma_n$ for $j = 1, \cdots, n$. One finds $\operatorname{E}(\tilde{\rho}_1^2) = p$ and $\operatorname{E}(\tilde{\rho}_j^4) = p^2 + \tau p + o(p)$. Also, we can show that the limiting distribution of $M_n(z)$ will not be affected by replacing $(\hat{\rho}_1, \cdots, \hat{\rho}_n)$ with $(\tilde{\rho}_1, \cdots, \tilde{\rho}_n)$.

To conclude, in the rest of the proof, we shall assume that the variable $\rho_j, q \leq j \leq n$ has been truncated and satisfies

$$|\rho_j^2 - p| \leq \delta_n p. \tag{3.6}$$

3.2.2. *The limit of $\mathcal{II}(z) - n\underline{m}_n^0(z)$.* We now introduce some necessary notation that is useful in the proof below. Denote $\mathbf{q}_k = \mathbf{B}_n^* \mathbf{Y}_n^* \mathbf{e}_k$,

$$\mathbf{D} = \mathbf{B}_n^{-1} \left[\operatorname{diag} \left(\frac{\|\mathbf{y}_1\|^2}{\rho_1^2}, \cdots, \frac{\|\mathbf{y}_n\|^2}{\rho_n^2} \right) - \mathbf{I}_n \right] \mathbf{B}_n = \sum_{k=1}^{p} \mathbb{D}_k, \ \mathbf{J}_k(z) = \mathbf{J}(z) - \frac{1}{n} \lambda_k \mathbf{q}_k \mathbf{q}_k^*,$$

$$\mathbf{D}_k = \mathbf{D} - \mathbb{D}_k, \ \mathbb{D}_k = \mathbf{B}_n^{-1} \left[\operatorname{diag} \left(\frac{y_{k1}^2}{\rho_1^2}, \cdots, \frac{y_{kn}^2}{\rho_n^2} \right) - \frac{1}{p} \mathbf{I}_n \right] \mathbf{B}_n.$$

Define $g_{2n}(z) = \frac{1}{n} \operatorname{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{\Phi}_n \right)$,

$$\varepsilon_k(z) = \mathbf{q}_k^* \mathbf{J}_k^{-1}(z) \mathbf{q}_k - \operatorname{tr}(\mathbf{J}_k^{-1}(z) \mathbf{\Phi}_n), \quad \gamma_k(z) = \mathbf{q}_k^* \mathbf{J}_k^{-2}(z) \mathbf{q}_k - \operatorname{tr}(\mathbf{J}_k^{-2}(z) \mathbf{\Phi}_n),$$

$$\beta_k(z) = \frac{1}{1 + n^{-1} \lambda_k \mathbf{q}_k^* \mathbf{J}_k^{-1}(z) \mathbf{q}_k}, \qquad \widetilde{\beta}_k(z) = \frac{1}{1 + n^{-1} \lambda_k \operatorname{tr}(\mathbf{J}_k^{-1}(z) \mathbf{\Phi}_n)},$$

$$b_k(z) = \frac{1}{1 + n^{-1} \lambda_k \operatorname{E} \operatorname{tr}(\mathbf{J}_k^{-1}(z) \mathbf{\Phi}_n)}, \qquad \psi_k(z) = \frac{1}{1 + \lambda_k \operatorname{E} g_{2n}(z)}.$$

To obtain the limit of $\mathcal{II}(z) - n\underline{m}_n^0(z)$, we rewrite that

$$\begin{aligned}
&\mathcal{II}(z) - n\underline{m}_n^0(z) \\
&= \operatorname{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \right) - n\underline{m}_n^0(z) + \operatorname{E} \operatorname{tr} \left(\mathbf{J}^{-1} \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right) + z \operatorname{E} \operatorname{tr} \left(\mathbf{J}^{-2}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right) \\
&\quad + z \operatorname{tr} \left[\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right]^2 + z^2 \operatorname{tr} \left[\left(\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right)^2 \mathbf{J}^{-1}(z) \right] \\
&+ z^2 \operatorname{tr} \left\{ \left[\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right]^2 \mathbf{H}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right\} \\
&\quad + z^3 \operatorname{tr} \left[\left(\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right)^3 \mathbf{H}^{-1}(z) \right].
\end{aligned}$$

Firstly, we will deal with the last two terms of the right-hand side of the above equality. We shall prove that the moments of $\beta_k(z)$, $\widetilde{\beta}_k(z)$, $\|\mathbf{J}^{-1}(z)\|$, and $\|\mathbf{J}_k^{-1}(z)\|$ are all bounded in n with $z \in \mathcal{C}_n$. It can also be verified that $\left\|\mathbf{H}^{-1}(z)\right\| \leq C$ in probability, and $|b_k(z)| \leq C$. Using Taylor's expansion, we get

$$\mathbf{M}_n = \mathbf{I}_n + \frac{1}{2} \mathbf{D} - \frac{1}{8} \mathbf{D}^2 (1 + o(1)).$$

This implies that

$$\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n = \frac{1}{2} \mathbf{D} + \frac{1}{2} \mathbf{D}^* - \frac{1}{8} \mathbf{D}^2 (1 + o(1)) + \frac{1}{4} \mathbf{D}^* \mathbf{D} (1 + o(1)) - \frac{1}{8} \left(\mathbf{D}^* \right)^2 (1 + o(1)). \tag{3.7}$$

By Lemma 4.2 and (3.7), we have

$$\begin{aligned}
&\left| z^2 \operatorname{tr} \left\{ \left[\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right]^2 \mathbf{H}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right\} \right. \\
&\qquad \left. + z^3 \operatorname{tr} \left[\left(\mathbf{J}^{-1}(z) \left(\mathbf{M}_n^* \mathbf{M}_n - \mathbf{I}_n \right) \right)^3 \mathbf{H}^{-1}(z) \right] \right| \\
&\leq Cn \left\| \mathbf{D} \right\|^3 (1 + o(1)) \leq \frac{C}{\sqrt{n}} \to 0.
\end{aligned}$$

Next, we will show that

$$\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2-\operatorname{E}\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2\xrightarrow{p}0.$$

It follows from (3.7) that

$$\begin{aligned}&\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2\\&\quad=\frac{1}{4}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2(1+o(1))+\frac{1}{4}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)^2(1+o(1))\\&\qquad+\frac{1}{2}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\mathbf{J}^{-1}(z)\mathbf{D}^*\right)(1+o(1)).\end{aligned}$$

Hence, it suffices to show that

$$\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2-\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2\xrightarrow{p}0. \tag{3.8}$$

The proof is postponed to Appendix A. Moreover, following the same lines, we can get the similar conclusion

$$\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z)\right]-\operatorname{E}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z)\right]\xrightarrow{p}0.$$

Hence, $\mathcal{II}$ can be rewritten as

$$\begin{aligned}&\mathcal{II}(z)\\&=\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right)-n\underline{m}_n^0(z)+\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right)\\&\quad+z\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right)\\&\quad+z\operatorname{E}\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2\\&\quad+z^2\operatorname{E}\operatorname{tr}\left[\left(\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right)^2\mathbf{J}^{-1}(z)\right]+o_p(1)\\&=\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right)-n\underline{m}_n^0(z)+\mathcal{II}_1(z)+\mathcal{II}_2(z)+\mathcal{II}_3(z)+\mathcal{II}_4(z)+o_p(1).\end{aligned} \tag{3.9}$$

In [4] it is shown that

$$\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right)-n\underline{m}_n^0(z)\to(2-\beta)\frac{d_1(z)}{\left[1-cz^{-2}d_2(z)d_3(z)\right]^2}. \tag{3.10}$$

Thus the analysis of $\mathcal{II}(z)$ can be decomposed into the following four steps.

Step 1 (The limit of $\mathcal{II}_1(z)$). By (3.7), we get

$$\begin{aligned}\mathcal{II}_1(z)=&\frac{1}{2}\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)+\frac{1}{2}\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)-\frac{1}{8}\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right)(1+o(1))\\&+\frac{1}{4}\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right)(1+o(1))-\frac{1}{8}\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right)(1+o(1)).\end{aligned}$$

Using

$$\mathbf{J}^{-1}(z)=\mathbf{J}_k^{-1}(z)-\frac{1}{n}\lambda_k\beta_k(z)\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z), \tag{3.11}$$

one can derive that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&=\sum_{k=1}^{p}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbb{D}_k\right)\\
&=\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&\quad-\frac{1}{n}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\mathrm{E}\left(\lambda_k\beta_k(z)|y_{kj}|^2\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)\\
&\quad-\frac{1}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left[\lambda_k\beta_k(z)\left(|y_{kj}|^2-1\right)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right].
\end{aligned}$$

Applying the Taylor expansion to get that when $(x-p)\to 0$,

$$\frac{1}{x}=\frac{1}{p}-\frac{x-p}{p^2}+\frac{(x-p)^2}{p^3}(1+o(1)), \tag{3.12}$$

we have

$$\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)=\frac{1}{p^3}\mathrm{E}\left(\rho_j^2-p\right)^2(1+o(1))=\frac{\tau+o(1)}{p^2}, \tag{3.13}$$

and

$$\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)^2=\frac{1}{p^4}\mathrm{E}\left(\rho_j^2-p\right)^2(1+o(1))=\frac{\tau+o(1)}{p^3}. \tag{3.14}$$

By (3.13), one can derive that

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\mathrm{E}\left(\lambda_k\beta_k(z)|y_{kj}|^2\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)\right|\\
&\le\frac{C}{n}\sum_{k=1}^{p}\sum_{j=1}^{n}\left|\mathrm{E}\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\right|\mathrm{E}\left|\beta_k(z)|y_{kj}|^2\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|\\
&\le\frac{C}{np^2}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}^{1/2}\left|\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^2\mathrm{E}^{1/4}\left|y_{kj}\right|^8\mathrm{E}^{1/4}\left|\beta_k(z)\right|^4\\
&\le\frac{C}{n}\to 0.
\end{aligned}$$

Hence, we see that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&=-\frac{1}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left[\lambda_k\beta_k(z)\left(|y_{kj}|^2-1\right)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right]\\
&\quad+\frac{\tau}{p}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right)+o(1).
\end{aligned}$$

Using (5.17) in [4],

$$\mathrm{E}\left|\beta_k(z)-\psi_k(z)\right|^2=O(n^{-1}), \tag{3.15}$$

and

$$\mathrm{E}\left(\mathbf{e}_k^T\mathbf{Y}_n\mathbf{R}_1\mathbf{Y}_n^*\mathbf{e}_k - \mathrm{tr}\left(\mathbf{R}_1\right)\right)\left(\mathbf{e}_k^T\mathbf{Y}_n\mathbf{R}_2\mathbf{Y}_n^*\mathbf{e}_k - \mathrm{tr}\left(\mathbf{R}_2\right)\right)$$
$$= \mathrm{tr}\left(\mathbf{R}_1\mathbf{R}_2\right) + |\,\mathrm{E}\, y_{k1}^2|\,\mathrm{tr}\left(\mathbf{R}_1\mathbf{R}_2^T\right),$$

one obtains that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&= -\frac{1}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k\psi_k(z)\,\mathrm{E}\left[\left(|y_{kj}|^2-1\right)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right]\\
&\quad + \frac{\tau}{p}\,\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right) + o(1)\\
&= -\frac{1+|\,\mathrm{E}\,y_{11}^2|}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k\psi_k(z)\,\mathrm{E}\left[\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\right]\\
&\quad + \frac{\tau}{p}\,\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right)\\
&\quad - \frac{1+|\,\mathrm{E}\,y_{11}^2|}{pn^2}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k^2\psi_k(z)\,\mathrm{E}\left[\beta_k(z)\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\right]\\
&\quad + o(1)\\
&= -\frac{1+|\,\mathrm{E}\,y_{11}^2|}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k\psi_k(z)\,\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&\quad + \frac{\tau}{p}\,\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right) + o(1),
\end{aligned}$$

where the first equality is due to the fact that

$$\begin{aligned}
&\left|\frac{1}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k\,\mathrm{E}\left[\left(\beta_k(z)-\psi_k(z)\right)\left(|y_{kj}|^2-1\right)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right]\right|\\
&\le \frac{C}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}^{1/2}\left|\beta_k(z)-\psi_k(z)\right|^2\mathrm{E}^{1/2}\left||y_{kj}|^2-1\right|^4 \le \frac{C}{\sqrt{n}} \to 0,
\end{aligned}$$

and the last second equality is from the fact that

$$\begin{aligned}
&\left|\frac{2}{pn^2}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k^2\psi_k(z)\,\mathrm{E}\left[\beta_k(z)\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\right]\right|\\
&\le \frac{C}{pn^2}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}^{1/2}\left|\mathbf{q}_k^*\mathbf{J}_k^{-1}(\bar{z})\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^2\mathrm{E}^{1/4}\left|\beta_k(z)\right|^4\mathrm{E}^{1/4}\left\|\mathbf{J}_k^{-1}(z)\right\|^4\\
&\le \frac{C}{n} \to 0.
\end{aligned}$$

Hence, we get that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&\quad = -\frac{1+|\,\mathrm{E}\,y_{11}^2|}{pn}\sum_{k=1}^{p}\sum_{j=1}^{n}\lambda_k\psi_k(z)\,\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&\qquad + \frac{\tau}{p}\,\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right) + o(1).
\end{aligned}$$

Let $\mathbf{W}(z)=\frac{1}{n}\sum_{j=1}^n\lambda_j\psi_j(z)\mathbf{\Phi}_n-z\mathbf{I}_n$. It has been proved that $\left\|\mathbf{W}^{-1}(z)\right\|$ is uniformly bounded on $\mathcal{C}_n$ and

$$|\psi_k(z)-\frac{1}{1+\lambda_k g_{2n}^0(z)}|=o(1). \tag{3.16}$$

It follows from (3.16) that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&\quad=\frac{zg_{1n}^0(z)\left(1+|\mathrm{E}\,y_{11}^2|\right)}{p}\sum_{j=1}^n\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&\qquad+\frac{\tau}{p}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\right)+o(1).
\end{aligned}$$

From the Supplement A of [4], we know that

$$\begin{aligned}
&\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&\qquad=\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j+o(1).
\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\\
&\to\lim_{n\to\infty}\frac{g_1(z)\left(1+|\mathrm{E}\,y_{11}^2|\right)}{zp}\sum_{j=1}^n\mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_\triangle(z)\mathbf{e}_j\mathbf{e}_j^T\mathbb{G}_\triangle(z)\mathbf{e}_j+\tau c^{-1}\underline{m}(z)\\
&=\lim_{n\to\infty}\frac{1+|\mathrm{E}\,y_{11}^2|}{zp}\mathrm{tr}\,\mathbb{G}_\triangledown(z)-\lim_{n\to\infty}\frac{1+|\mathrm{E}\,y_{11}^2|}{zp}\sum_{j=1}^n\left[\mathbf{e}_j^T\mathbb{G}_\triangle(z)\mathbf{e}_j\right]^2+\tau c^{-1}\underline{m}(z)\\
&=\left(\tau-1-|\mathrm{E}\,y_{11}^2|\right)c^{-1}\underline{m}(z)-\lim_{n\to\infty}\frac{1+|\mathrm{E}\,y_{11}^2|}{zp}\sum_{j=1}^n\left[\mathbf{e}_j^T\mathbb{G}_\triangle(z)\mathbf{e}_j\right]^2.
\end{aligned}$$

By the same argument, we can also get that

$$\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)\\
&\qquad=\left(\tau-1-|\mathrm{E}\,y_{11}^2|\right)c^{-1}\underline{m}(z)-\lim_{n\to\infty}\frac{1+|\mathrm{E}\,y_{11}^2|}{zp}\sum_{j=1}^n\left[\mathbf{e}_j^T\mathbb{G}_\triangle(z)\mathbf{e}_j\right]^2.
\end{aligned}$$

Appendix B shows that

$$\begin{aligned}
&-\frac{1}{8}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right)+\frac{1}{4}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right)-\frac{1}{8}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right)\\
&\to-\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4c}\underline{m}(z)\\
&\quad-\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4zp}\sum_{j=1}^n\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_\triangle(z)\mathbf{e}_j.
\end{aligned}$$

Together with the above three equalities, we find that

$$\mathcal{II}_1(z) \to \left(\frac{3}{4}\tau - \frac{5}{4}\left(1 + |\operatorname{E} y_{11}^2|\right)\right) c^{-1}\underline{m}(z) - \lim_{n\to\infty} \frac{1 + |\operatorname{E} y_{11}^2|}{zp} \sum_{j=1}^{n} \left[\mathbf{e}_j^T \mathbb{G}_{\triangle}(z)\mathbf{e}_j\right]^2$$
$$- \lim_{n\to\infty} \frac{\tau + 1 + |\operatorname{E} y_{11}^2|}{4zp} \sum_{j=1}^{n} \mathbf{e}_j^T \left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1} \mathbf{e}_j\mathbf{e}_j^T \mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_{\triangle}(z)\mathbf{e}_j.$$

Step 2 (The limit of $\mathcal{II}_2(z)$). It is obvious that

$$\begin{aligned}
\mathcal{II}_2(z) = {} & z\frac{\partial \mathcal{II}_1(z)}{\partial z} \\
= {} & \frac{z}{c}\left(\frac{3}{4}\tau - \frac{5}{4}\left(1 + |\operatorname{E} y_{11}^2|\right)\right) \frac{\partial \underline{m}(z)}{\partial z} + \lim_{n\to\infty} \frac{1 + |\operatorname{E} y_{11}^2|}{pz} \sum_{j=1}^{n} \left[\mathbf{e}_j^T \mathbb{G}_{\triangle}(z)\mathbf{e}_j\right]^2 \\
& - \lim_{n\to\infty} \frac{\left(1 + |\operatorname{E} y_{11}^2|\right)}{p} \sum_{j=1}^{n} \frac{\partial\left\{\left[\mathbf{e}_j^T \mathbb{G}_{\triangle}(z)\mathbf{e}_j\right]^2\right\}}{\partial z} \\
& + \lim_{n\to\infty} \frac{\tau + 1 + |\operatorname{E} y_{11}^2|}{4zp} \sum_{j=1}^{n} \mathbf{e}_j^T \left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1} \mathbf{e}_j\mathbf{e}_j^T \mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_{\triangle}(z)\mathbf{e}_j \\
& - \lim_{n\to\infty} \frac{\tau + 1 + |\operatorname{E} y_{11}^2|}{4p} \sum_{j=1}^{n} \frac{\partial\left\{\mathbf{e}_j^T \left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1} \mathbf{e}_j\mathbf{e}_j^T \mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_{\triangle}(z)\mathbf{e}_j\right\}}{\partial z}.
\end{aligned}$$

Step 3 (The limit of $\mathcal{II}_3(z)$). Rewrite

$$\begin{aligned}
\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right]^2 = {} & \Big[\frac{1}{4}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 + \frac{1}{4}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)^2 \\
& + \frac{1}{2}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\mathbf{J}^{-1}(z)\mathbf{D}^*\right)\Big](1 + o(1)).
\end{aligned}$$

By (3.11), we have

$$\begin{aligned}
& \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 \\
& = \sum_{k=1}^{p} \operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right) + \sum_{k=1}^{p} \operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right) \\
& \quad - \frac{1}{n}\sum_{k=1}^{p} \lambda_k \operatorname{E}\left(\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right) \\
& \quad - \frac{1}{n}\sum_{k=1}^{p} \lambda_k \operatorname{E}\left(\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right) \\
& \quad + \frac{1}{n^2}\sum_{k=1}^{p} \lambda_k^2 \operatorname{E}\left(\beta_k^2(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right).
\end{aligned}$$

It follows from (A.1) that

$$\operatorname{E}\|\mathbb{D}_k\|^\ell = \max_{j=1,\cdots,n} \operatorname{E}\left|\frac{|y_{kj}|^2}{\rho_j^2} - \frac{1}{p}\right|^\ell \le \frac{C}{p^\ell}. \tag{3.17}$$

Using Lemma 4.1, Lemma 4.5, and (3.17), one gets that

$$\left|\frac{1}{n}\sum_{k=1}^{p}\lambda_k \operatorname{E}\left(\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\left(\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}+\mathbf{D}\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)\right|$$
$$\le \frac{C}{n}\sum_{k=1}^{p}\operatorname{E}\left(|\beta_k(z)\mathbf{q}_k^*\mathbf{q}_k|\left\|\mathbf{J}_k^{-3}(z)\mathbb{D}_k\mathbf{D}\right\|\right)$$
$$\le \frac{C}{n}\sum_{k=1}^{p}\operatorname{E}^{1/4}|\beta_k(z)|^4\operatorname{E}^{1/4}|\mathbf{q}_k^*\mathbf{q}_k|^4\operatorname{E}^{1/4}\left\|\mathbf{J}_k^{-1}(z)\right\|^{12}\operatorname{E}^{1/8}\|\mathbb{D}_k\|^8\operatorname{E}^{1/8}\|\mathbf{D}\|^8\to 0,$$

and

$$\left|\frac{1}{n^2}\sum_{k=1}^{p}\lambda_k^2\operatorname{E}\left(\beta_k^2(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)\right|$$
$$\le \frac{C}{n^2}\sum_{k=1}^{p}\operatorname{E}^{1/4}|\beta_k(z)|^8\operatorname{E}^{1/4}|\mathbf{q}_k^*\mathbf{q}_k|^8\operatorname{E}^{1/4}\left\|\mathbf{J}_k^{-1}(z)\right\|^{16}\operatorname{E}^{1/8}\|\mathbb{D}_k\|^8\operatorname{E}^{1/8}\|\mathbf{D}\|^8\to 0.$$

Hence, it yields that

(3.18)
$$\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2$$
$$=\sum_{k=1}^{p}\operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)+\sum_{k=1}^{p}\operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)+o(1).$$

By (3.13), (3.14), and Lemma 4.6, we have

$$\sum_{k=1}^{p}\operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)$$
$$=\frac{1+|\operatorname{E}y_{11}^2|}{p}\sum_{j=1}^{n}\operatorname{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)+o(1),$$

and

$$\sum_{k=1}^{p}\operatorname{E}\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)$$
$$=\frac{\tau}{p}\sum_{j=1}^{n}\operatorname{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)+o(1),$$

where the details can be found in the proof of (B.1). From the supplement of [4], one can get that

$$\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2=\!\!\to\lim_{n\to\infty}\frac{\tau+1+|\operatorname{E}y_{11}^2|}{pz^2}\sum_{j=1}^{n}\left(\mathbf{e}_j^T\mathbb{G}_{\triangle}(z)\mathbf{e}_j\right)^2.$$

Thus

$$\operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)^2\to\lim_{n\to\infty}\frac{\tau+1+|\operatorname{E}y_{11}^2|}{pz^2}\sum_{j=1}^{n}\left(\mathbf{e}_j^T\mathbb{G}_{\triangle}(z)\mathbf{e}_j\right)^2,$$

and

$$\begin{aligned}&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\mathbf{J}^{-1}(z)\mathbf{D}^*\right)\\&\qquad\to \lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{z^2p}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathbb{G}_{\bigtriangledown}(z)\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^*\mathbf{e}_j.\end{aligned}$$

Consequently, we get that

$$\begin{aligned}\mathrm{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2 &\to \lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2pz^2}\sum_{j=1}^{n}\left(\mathbf{e}_j^T\mathbb{G}_{\bigtriangleup}(z)\mathbf{e}_j\right)^2\\&+\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2z^2p}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^*\mathbf{e}_j.\end{aligned}$$

Thus,

$$\begin{aligned}\mathcal{II}_3(z)\to&\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2pz}\sum_{j=1}^{n}\left(\mathbf{e}_j^T\mathbb{G}_{\bigtriangleup}(z)\mathbf{e}_j\right)^2\\&+\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2zp}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^*\mathbf{e}_j.\end{aligned}$$

Step 4 (The limit of $\mathcal{II}_4(z)$). We have

$$\begin{aligned}&\mathcal{II}_4(z)\\&=\frac{z^2}{2}\frac{\partial\left\{\mathrm{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n-\mathbf{I}_n\right)\right]^2\right\}}{\partial z}\\&=\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4p}\sum_{j=1}^{n}\frac{\partial\left\{\left(\mathbf{e}_j^T\mathbb{G}_{\bigtriangleup}(z)\mathbf{e}_j\right)^2\right\}}{\partial z}\\&\quad-\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2pz}\sum_{j=1}^{n}\left(\mathbf{e}_j^T\mathbb{G}_{\bigtriangleup}(z)\mathbf{e}_j\right)^2\\&\quad-\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{2zp}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathbb{G}_{\bigtriangledown}(z)\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^*\mathbf{e}_j\\&\quad+\lim_{n\to\infty}\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4p}\sum_{j=1}^{n}\frac{\partial\left[\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbb{G}_{\bigtriangledown}(z)\mathbf{B}_n^*\mathbf{e}_j\right]}{\partial z}.\end{aligned}$$

Combining (3.9), (3.10) and the argument above (Steps 1-4), we conclude that

$$\mathcal{II}(z)\to(2-\beta)\frac{d_1(z)}{\left[1-cz^{-2}d_2(z)d_3(z)\right]^2}+d_4(z)+(\beta-1)d_5(z).$$

3.2.3. *The limiting distribution of* $\mathcal{I}(z)$. Now we deal with the term $\mathcal{I}(z)$. Recall that

$$\begin{aligned}\mathcal{I}(z) = &\left[\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\right)\right]\\ &+ \left[\operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right)\right]\\ &+ z\left[\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-2}(z)\left(\mathbf{M}^*\mathbf{M} - \mathbf{I}_n\right)\right)\right]\\ =& \mathcal{I}_1(z) + \mathcal{I}_2(z) + \mathcal{I}_3(z).\end{aligned}$$

By (3.7), we have that

$$\begin{aligned}&\operatorname{tr}\left[\mathbf{J}^{-1}(z)\left(\mathbf{M}_n^*\mathbf{M}_n - \mathbf{I}_n\right)\right]\\ =&\frac{1}{2}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right) + \frac{1}{2}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right) - \frac{1}{8}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right)(1+o(1))\\ &+ \frac{1}{4}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right)(1+o(1)) - \frac{1}{8}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right)(1+o(1)).\end{aligned}$$

Then following the same procedures as in Appendix A, we shall obtain that

$$\begin{aligned}&\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right) = o_p(1),\\ &\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right) = o_p(1),\\ &\operatorname{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right) = o_p(1).\end{aligned}$$

Hence,

$$\begin{aligned}\mathcal{I}_2(z) =&\frac{1}{2}\left[\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)\right]\\ &+ \frac{1}{2}\left[\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right) - \operatorname{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\right)\right] + o_p(1)\\ =& \frac{1}{2}\mathcal{I}_{21}(z) + \frac{1}{2}\mathcal{I}_{22}(z) + o_p(1).\end{aligned}\tag{3.19}$$

Using (3.11), one can rewrite $\mathcal{I}_{21}$ as

$$\begin{aligned}\mathcal{I}_{21}(z) =& \sum_{k=1}^{p}\left(\operatorname{E}_k - \operatorname{E}_{k-1}\right)\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)\\ &- \frac{1}{n}\sum_{k=1}^{p}\left(\operatorname{E}_k - \operatorname{E}_{k-1}\right)\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\\ &- \frac{1}{n}\sum_{k=1}^{p}\left(\operatorname{E}_k - \operatorname{E}_{k-1}\right)\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\\ &+ \sum_{k=1}^{p}\sum_{j=1}^{n}\left(\operatorname{E}_0 - \operatorname{E}\right)\left(\frac{1}{\rho_j^2} - \frac{1}{p}\right)\operatorname{E}\left(|y_{kj}|^2\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\ \triangleq& \sum_{k=1}^{p}\left(\operatorname{E}_k - \operatorname{E}_{k-1}\right)\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right) - \sum_{k=1}^{p}\left(Q_{k1}(z) + Q_{k2}(z)\right)\\ &+ \frac{1}{p}\sum_{j=1}^{n}\left(\rho_j^2 - p\right)\operatorname{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\end{aligned}\tag{3.20}$$

$$+\frac{1}{p^2}\sum_{j=1}^{n}\left[\left(\rho_j^2-p\right)^2-\tau p\right]\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)(1+o_p(1)).$$

We shall assert that the last term in the right-hand side of the last equality tends to zero in probability. In fact, by (3.6),

$$\begin{aligned}\mathrm{E}&\left|\frac{1}{p^2}\sum_{j=1}^{n}\left[\left(\rho_j^2-p\right)^2-\tau p\right]\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)\right| \\ &\leq\frac{C}{p^4}\sum_{j=1}^{n}\mathrm{E}\left|\left(\rho_j^2-p\right)^2-\tau p\right|^2\leq\frac{C}{p^4}\sum_{j=1}^{n}\left[\delta_n^2p^2\,\mathrm{E}\left(\rho_j^2-p\right)^2+\tau^2p^2\right]\to 0.\end{aligned}\tag{3.21}$$

Now we are in position to show that $\sum_{k=1}^{p}Q_{k1}(z)$ and $\sum_{k=1}^{p}Q_{k2}(z)$ are $o_p(1)$. Using Lemma 4.8 and (3.17), one finds that

$$\begin{aligned}\mathrm{E}\left|\sum_{k=1}^{p}Q_{k2}(z)\right|^2&\leq\frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left[(\mathbf{q}_k^*\mathbf{q}_k)^2\left\|\mathbb{D}_k\right\|^2\right]\\&\leq\frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}^{1/2}\left(\mathbf{q}_k^*\mathbf{q}_k\right)^4\mathrm{E}^{1/2}\left\|\mathbb{D}_k\right\|^4\to 0.\end{aligned}$$

That is

$$\sum_{k=1}^{p}Q_{k2}(z)=o_p(1).\tag{3.22}$$

By

$$\beta_k(z)=\widetilde{\beta}_k(z)-\frac{1}{n}\lambda_k\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z),\tag{3.23}$$

it follows that

$$\begin{aligned}\sum_{k=1}^{p}Q_{k1}(z)=&\frac{1}{n}\sum_{k=1}^{p}\mathrm{E}_k\,\lambda_k\widetilde{\beta}_k(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)\\&-\frac{1}{n^2}\sum_{k=1}^{p}\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\lambda_k^2\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k.\end{aligned}$$

Applying Lemma 4.1 and Lemma 4.5, one has that

$$\mathrm{E}\left|\frac{1}{n}\sum_{k=1}^{p}\mathrm{E}_k\,\lambda_k\widetilde{\beta}_k(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)\right|^2\to 0,$$

and

$$\mathrm{E}\left|\frac{1}{n^2}\sum_{k=1}^{p}\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\lambda_k^2\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^2\to 0.$$

Hence, we obtain that

$$\sum_{k=1}^{p}Q_{k1}(z)=o_p(1).\tag{3.24}$$

Combining (3.21), (3.22), (3.24), with (3.20), we deduce that

$$\mathcal{I}_{21}(z) = \frac{1}{p}\sum_{k=1}^{p}\sum_{j=1}^{n} \mathrm{E}_k \left(\mathbf{e}_j^T \mathbf{B}_n \mathbf{J}_k^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j \left(|y_{kj}|^2 - 1\right)\right) + \frac{1}{p}\sum_{j=1}^{n} \left(\rho_j^2 - p\right) \mathbf{e}_j^T \mathbf{B}_n \, \mathrm{E}\, \mathbf{J}^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j + o_p(1).$$

By the same argument, we can get that

$$\mathcal{I}_{22}(z) = \frac{1}{p}\sum_{k=1}^{p}\sum_{j=1}^{n} \mathrm{E}_k \left(\mathbf{e}_j^T \left(\mathbf{B}_n^*\right)^{-1} \mathbf{J}_k^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j \left(|y_{kj}|^2 - 1\right)\right) + \frac{1}{p}\sum_{j=1}^{n} \left(\rho_j^2 - p\right) \mathbf{e}_j^T \mathbf{B}_n \, \mathrm{E}\, \mathbf{J}^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j + o_p(1).$$

Hence,

(3.25)

$$\begin{aligned}\mathcal{I}_2(z) &= \frac{1}{2p}\sum_{k=1}^{p}\sum_{j=1}^{n} \mathrm{E}_k \left[\left(|y_{kj}|^2 - 1\right)\left(\mathbf{e}_j^T \mathbf{B}_n \mathbf{J}_k^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j + \mathbf{e}_j^T \left(\mathbf{B}_n^*\right)^{-1} \mathbf{J}_k^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j\right)\right] \\ &\quad + \frac{1}{p}\sum_{j=1}^{n} \left(\rho_j^2 - p\right) \mathbf{e}_j^T \mathbf{B}_n \, \mathrm{E}\, \mathbf{J}^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j + o_p(1) \\ &\triangleq \sum_{k=1}^{p} Q_{k3}(z) - Q_4(z) + o_p(1),\end{aligned}$$

where

$$Q_4(z) = \frac{1}{zp}\sum_{j=1}^{n} \left(\rho_j^2 - p\right) \mathbf{e}_j^T \left(g_{1n}^0(z) \mathbf{B}_n \mathbf{B}_n^* + \mathbf{I}_n\right)^{-1} \mathbf{e}_j + o(1).$$

Because $\mathcal{I}_3 = z\frac{\partial}{\partial z}\mathcal{I}_2$, we have

$$\mathcal{I}_3(z) = z\sum_{k=1}^{p} \frac{\partial}{\partial z} Q_{k3}(z) - z\frac{\partial}{\partial z} Q_4(z) + o_p(1). \tag{3.26}$$

In [4] it is shown that

$$\mathcal{I}_1(z) = \operatorname{tr} \mathbf{J}^{-1}(z) - \mathrm{E} \operatorname{tr} \mathbf{J}^{-1}(z) = \sum_{k=1}^{p} W_k(z) + o_p(1), \tag{3.27}$$

where $W_k(z) = -\frac{1}{n}\mathrm{E}_k\left(\frac{\partial}{\partial z}\lambda_k \widetilde{\beta}_k(z) \varepsilon_k(z)\right)$. They also show that $\operatorname{tr} \mathbf{J}^{-1}(z) - \mathrm{E} \operatorname{tr} \mathbf{J}^{-1}(z)$ converges to a Gaussian process with mean zero and the following covariance function

$$(3 - \beta)\frac{\partial^2}{\partial z_2 \partial z_1}\int_0^{d(z_1, z_2)} \frac{1}{1 - z} dz.$$

Combining (3.25), (3.26), and (3.27), one obtains that

$$\mathcal{I}(z) = \sum_{k=1}^{p}\left(W_k(z) + Q_{k3}(z) + z \cdot \frac{\partial}{\partial z} Q_{k3}(z)\right) - Q_4(z) - z\frac{\partial}{\partial z} Q_4(z) + o_p(1).$$

Next, we will show for any positive integer $r > 0$ the sum

$$\sum_{j=1}^{r} \alpha_j \mathcal{I}(z_j), \quad \Im(z_j) > 0$$

will converge in distribution to a Gaussian random variable.

For any $z_1, \dots, z_r \in \mathbb{C}_+$, $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ and any $\varepsilon > 0$, we have

$$\begin{aligned}
&\sum_{k=1}^{p} \mathrm{E}\left(\left|\sum_{\ell=1}^{r} \alpha_\ell \left(Q_{k3}(z_\ell) - \frac{1}{p} Q_4(z_\ell)\right)\right|^2 I\left(\left|\sum_{\ell=1}^{r} \alpha_\ell \left(Q_{k3}(z_\ell) - \frac{1}{p} Q_4(z_\ell)\right)\right| \geq \varepsilon\right)\right) \\
\leq &\frac{C}{\varepsilon^2} \sum_{k=1}^{p} \sum_{\ell=1}^{r} \alpha_\ell^4 \,\mathrm{E}\left|Q_{k3}(z_\ell) - \frac{1}{p} Q_4(z_\ell)\right|^4 \\
\leq &\frac{C}{\varepsilon^2 p^3} \sum_{k=1}^{p} \sum_{\ell=1}^{r} \sum_{j=1}^{n} \alpha_\ell^4 \,\mathrm{E}\left|\mathbf{e}_j^T \mathbf{B}_n \mathbf{J}_k^{-1}(z) \mathbf{B}_n^{-1} \mathbf{e}_j + \mathbf{e}_j^T \left(\mathbf{B}_n^*\right)^{-1} \mathbf{J}_k^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j\right|^4 \\
&\quad + \frac{C}{\varepsilon^2 p^6} \sum_{\ell=1}^{r} \sum_{j=1}^{n} \alpha_\ell^4 \,\mathrm{E}\left|\rho_j^2 - p\right|^4 \to 0.
\end{aligned}$$

This implies the satisfaction of Lindberger condition. Next, we derive the limit of the quadratic variation process, which is a sum involving the following eight processes:

$$\sum_{k=1}^{p} \mathrm{E}_{k-1}\left(W_k(z_1) W_k(z_2)\right), \ \sum_{k=1}^{p} \mathrm{E}_{k-1}\left(Q_{k3}(z_1) Q_{Q3}(z_2)\right),$$

$$\sum_{k=1}^{p} \mathrm{E}_{k-1}\left(\frac{\partial Q_{k3}(z_1)}{\partial z_1} \frac{\partial Q_{k3}(z_2)}{\partial z_2}\right),$$

$$\sum_{k=1}^{p} \mathrm{E}_{k-1}\left(W_k(z_1) Q_{k3}(z_2)\right),$$

$$\sum_{k=1}^{p} \mathrm{E}_{k-1}\left(W_k(z_1) \frac{\partial Q_{k3}(z_2)}{\partial z_2}\right), \ \sum_{k=1}^{p} \mathrm{E}_{k-1}\left(Q_{k3}(z_1) \frac{\partial Q_{k3}(z_2)}{\partial z_2}\right),$$

and

$$\mathrm{E}\left(Q_4(z_1) Q_4(z_2)\right), \quad \mathrm{E}\left(Q_4(z_1) \frac{\partial Q_4(z_2)}{\partial z_2}\right).$$

In fact, it is easy to see that deriving the limits of the following terms is enough

$$\sum_{k=1}^{p} \mathrm{E}_{k-1}\left(W_k(z_1) Q_{k3}(z_2)\right), \sum_{k=1}^{p} \mathrm{E}_{k-1}\left(Q_{k3}(z_1) Q_{k3}(z_2)\right), \ \mathrm{E}\left(Q_4(z_1) Q_4(z_2)\right).$$

First, let $\mathbf{W}_k(z) = \frac{1}{n}\sum_{j\neq k=1}^n \lambda_j\psi_j(z)\mathbf{\Phi}_n - z\mathbf{I}_n$. It has been proved in [4] that $|\widetilde{\beta}_k(z_1) - \frac{1}{1+\lambda_k g_{2n}^0(z)}| = o_p(1)$. Then by the above equality and (3.16), we have that

$$\sum_{k=1}^p \mathrm{E}_{k-1}\left(W_k(z_1)Q_{k3}(z_2)\right)$$
$$\xrightarrow{p} \lim_{n\to\infty} \frac{1+|\mathrm{E}\, y_{11}^2|}{pz_2}\frac{\partial}{\partial z_1}\left(g_1(z_1)\sum_{j=1}^n \mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_{\triangle}(z_1)\mathbf{e}_j\mathbf{e}_j^T\left(g_1(z_2)\mathbf{B}_n\mathbf{B}_n^*+\mathbf{I}_n\right)^{-1}\mathbf{e}_j\right).$$

Secondly, one finds that

$$\sum_{k=1}^p \mathrm{E}_{k-1}\left(Q_{k3}(z_1)Q_{k3}(z_2)\right) \xrightarrow{p} \lim_{n\to\infty}\frac{1+|\mathrm{E}\, y_{11}^2|}{pz_1z_2}\sum_{j=1}^n \mathbf{e}_j^T\mathbb{G}_{\triangle}(z_1)\mathbf{e}_j\mathbf{e}_j^T\mathbb{G}_{\triangle}(z_2)\mathbf{e}_j.$$

Next, we compute to derive that

$$\mathrm{E}\left(Q_4(z_1)Q_4(z_2)\right) \xrightarrow{p} \lim_{n\to\infty}\frac{\tau}{z_1z_2p}\sum_{j=1}^n \mathbf{e}_j^T\mathbb{G}_{\triangle}(z_1)\mathbf{e}_j\mathbf{e}_j^T\mathbb{G}_{\triangle}(z_2)\mathbf{e}_j.$$

Combining the above results and Lemma 4.4, we then obtain that $\mathcal{I}(z)$ converges weakly to a Gaussian process with mean zero and covariance function

$$\begin{aligned} v(z_1,z_2) = \lim_{n\to\infty}&\frac{\tau-3+\beta}{p}\sum_{j=1}^n\frac{\partial\left(\mathbf{e}_j^T\mathbb{G}_{\triangle}(z)\mathbf{e}_j\right)}{\partial z_1}\frac{\partial\left(\mathbf{e}_j^T\mathbb{G}_{\triangle}(z)\mathbf{e}_j\right)}{\partial z_2}\\ &+(3-\beta)\frac{\partial^2}{\partial z_2\partial z_1}\int_0^{d(z_1,z_2)}\frac{1}{1-z}dz. \end{aligned}$$

3.2.4. *The tightness of $M_n(z)$.* This part is devoted to the tightness of $\operatorname{tr}\mathbf{J}^{-1}(z) - \mathrm{E}\left(\operatorname{tr}\mathbf{J}^{-1}(z)\right)$, $\sum_{j=1}^n Q_{j3}(z)$ and $Q_4(z)$. Note that the tightness of $\operatorname{tr}\mathbf{J}^{-1}(z) - \mathrm{E}\left(\operatorname{tr}\mathbf{J}^{-1}(z)\right)$ has been proved in Bai et al. (2019) and thus we focus on the latter two terms.

Now we begin to prove the tightness of $\sum_{j=1}^n Q_{j3}(z)$. Actually we plan to prove that

$$\sup_{n,z_1,z_2\in\mathcal{C}^+}\mathrm{E}\left|\sum_{k=1}^p(Q_{k3}(z_1)-Q_{k3}(z_2))\right|^2/|z_1-z_2|^2 \le K$$

for some finite constant K. So it suffices to show that

$$\sup_{n,z_1,z_2\in\mathcal{C}_n}\mathrm{E}\left|\sum_{k=1}^p(Q_{k3}(z_1)-Q_{k3}(z_2))\right|^2/|z_1-z_2|^2 \le K.$$

Note that

$$\begin{aligned}&\sum_{k=1}^p(Q_{k3}(z_1)-Q_{k3}(z_2))/(z_1-z_2)\\ &=\frac{1}{2p}\sum_{j=1}^n\sum_{k=1}^p\left(|y_{kj}|^2-1\right)\\ &\cdot\left(\mathbf{e}_j^T\mathbf{B}_n\,\mathrm{E}_k\left(\mathbf{J}_k^{-1}(z_1)\mathbf{J}_k^{-1}(z_2)\right)\mathbf{B}_n^{-1}\mathbf{e}_j+\mathbf{e}_j^T\left(\mathbf{B}_n^*\right)^{-1}\mathrm{E}_k\left(\mathbf{J}_k^{-1}(z_1)\mathbf{J}_k^{-1}(z_2)\right)\mathbf{B}_n^*\mathbf{e}_j\right),\end{aligned}$$

which implies that

$$\mathrm{E}\left|\sum_{k=1}^{p}(Q_{k3}(z_1)-Q_{k3}(z_2))\right|^2/|z_1-z_2|^2 \le \frac{C}{p^2}\sum_{j=1}^{n}\sum_{k=1}^{p}\mathrm{E}\left\|\mathbf{J}_k^{-1}(z_1)\mathbf{J}_k^{-1}(z_2)\right\|^2 \le K.$$

Next, we shall show the tightness of $Q_4(z)$. It suffices to prove

$$\sup_{n,z_1,z_2\in\mathcal{C}_n}\mathrm{E}\left|Q_4(z_1)-Q_4(z_2)\right|^2/|z_1-z_2|^2 \le K.$$

After some calculations, we find that

$$\begin{aligned}\frac{\mathrm{E}\left|Q_4(z_1)-Q_4(z_2)\right|^2}{|z_1-z_2|^2} &= \mathrm{E}\left|\frac{1}{p}\sum_{j=1}^{n}\left(\rho_j^2-p\right)\mathbf{e}_j^T\mathbf{B}_n\,\mathrm{E}\left(\mathbf{J}^{-1}(z_1)\mathbf{J}^{-1}(z_2)\right)\mathbf{B}_n^{-1}\mathbf{e}_j\right|^2\\ &\le C\,\mathrm{E}\left\|\mathbf{J}^{-1}(z_1)\mathbf{J}^{-1}(z_2)\right\|^2 \le K.\end{aligned}$$

So the tightness is proved.

3.2.5. *Complete the proof.* It is known (see [9]) that for any l, $\eta_1>\lambda_1(1+\sqrt{c})^2$ and $\eta_2<\lambda_p(1-\sqrt{c})^2$, the probabilities $\mathrm{P}\left(\lambda_{\max}\left(\frac{1}{n}\mathbf{A}_n\mathbf{X}_n\mathbf{X}_n^*\mathbf{A}_n^*\right)\ge\eta_1\right)=o(n^{-l})$ and $\mathrm{P}\left(\lambda_{\min}(\frac{1}{n}\mathbf{A}_n\mathbf{X}_n\mathbf{X}_n^*\mathbf{A}_n^*)\le\eta_2\right)=o(n^{-l})$. Note the bounds $\lambda_{\max}^{\mathbf{AB}}\le\lambda_{\max}^{\mathbf{A}}\lambda_{\max}^{\mathbf{B}}$, and $\lambda_{\min}^{\mathbf{AB}}\ge\lambda_{\min}^{\mathbf{A}}\lambda_{\min}^{\mathbf{B}}$ that are valid for $n\times n$ nonnegative definite $\mathbf{A}$ and $\mathbf{B}$. Combining the definitions of x_l and x_r, we find that with probability 1, $\liminf_{n\to\infty}\min(x_r-\lambda_{\max}(\mathbf{S}_n),\lambda_{\min}(\mathbf{S}_n)-x_l)>0$. Thus, letting $\eta_r\in(\limsup_n\lambda_1\lambda_{\max}^{\mathbf{\Phi}_n}\left(1+\sqrt{c}\right)^2,x_r)$, we have for any $l>0$, $\mathrm{P}\left(\lambda_{\max}(\mathbf{S}_n)\ge\eta_r\right)=o(n^{-l})$. Likewise, we have $\mathrm{P}\left(\lambda_{\min}\left(\mathbf{S}_n\right)\le\eta_l\right)=o(n^{-l})$, where

$$\eta_l\in\begin{cases}(x_l,\liminf_n\lambda_p\lambda_{\min}^{\mathbf{\Phi}_n}I_{(0,1)}(c)\left(1-\sqrt{c}\right)^2), & \text{if } \liminf_n\lambda_p\lambda_{\min}^{\mathbf{\Phi}_n}I_{(0,1)}(c)>0,\\ (x_l,0), & \text{if } \liminf_n\lambda_p\lambda_{\min}^{\mathbf{\Phi}_n}I_{(0,1)}(c)\le 0.\end{cases}$$

By similar arguments in [4] and [9], combining the results in the above subsections 3.2.2–3.2.4, we complete the proof of Lemma 3.1 and thus Theorem 1.4 follows.

4. Auxiliary lemmas

Here we present some useful lemmas. The proofs of the first two lemmas are postponed to the appendix.

Lemma 4.1. *Under the conditions of Theorem* 1.4*, for any even $k\ge 2$, we have*

$$\mathrm{E}\left\|\mathbf{D}\right\|^k\to 0 \quad \textit{and} \quad \max_{j=1,\cdots,n}\mathrm{E}\left\|\mathbf{D}_j\right\|^k\to 0.$$

Lemma 4.2. *Under the conditions of Theorem* 1.4*, we have*

$$\|\mathbb{C}_n^2-\mathbf{I}\|\to 0\ a.s.,\quad \|\mathbf{D}\|=O_p(n^{-1/2}).$$

Lemma 4.3 (Theorem A.37 in [3])**.** *If $\mathbf{A}$ and $\mathbf{B}$ are two $n\times p$ matrices and λ_k and $\delta_k, k=1,2,\cdots,n$, denote their singular values respectively, then*

$$\min_{\pi}\sum_{k=1}^{n}|\lambda_k-\delta_{\pi(k)}|^2\le\mathrm{tr}\left[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*\right]\le\max_{\pi}\sum_{k=1}^{n}|\lambda_k-\delta_{\pi(k)}|^2.$$

If the singular values are arranged in descending order, then we have

$$\sum_{k=1}^{\nu} |\lambda_k - \delta_k|^2 \leq \mathrm{tr}\left[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*\right],$$

where $\nu = \min\{p, n\}$.

Lemma 4.4 (CLT for martingale)**.** *Suppose for each* n, $Y_{n1}, Y_{n2}, \cdots, Y_{nr_n}$ *with finite second moments is a real martingale difference sequence with respect to the increasing* σ*-fields* $\{\mathcal{F}_{nj}\}$. *If as* $n \to \infty$,

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

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

$$\text{(ii)} \qquad \sum_{j=1}^{r_n} \mathrm{E}(Y_{nj}^2 I(|Y_{nj} \geq \varepsilon|)) \to 0,$$

then

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

Lemma 4.5 (Lemma B.26 in [2])**.** *Let* $\mathbf{A} = (a_{jk})$ *be an* $n \times n$ *nonrandom matrix and* $\mathbf{x} = (x_1, \cdots, x_n)^T$ *be a random vector of independent entries. Assume that* $\mathrm{E}\, x_j = 0$, $\mathrm{E}\,|x_j|^2 = 1$ *and* $\mathrm{E}\,|x_j|^l \leq \nu_l$. *Then for* $k \geq 1$,

$$\mathrm{E}\,|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathrm{tr}\,\mathbf{A}|^k \leq C_k \left[(\nu_4\, \mathrm{tr}\, \mathbf{A}\mathbf{A}^*)^{k/2} + \nu_{2k}\, \mathrm{tr}\, (\mathbf{A}\mathbf{A}^*)^{k/2} \right],$$

where C_k *is a constant depending on* k *only.*

Lemma 4.6 (Inequality (4.8) in [2])**.** *Let* $\mathbf{M}$ *be an* $n \times n$ *nonrandom matrix. We have for* $j \in \{1, 2, \cdots, p\}$,

$$\mathrm{E}\left|\mathrm{tr}\, \mathbf{J}_j^{-1}\mathbf{M} - \mathrm{E}\,\mathrm{tr}\,\mathbf{J}_j^{-1}\mathbf{M}\right|^2 \leq C\|\mathbf{M}\|^2.$$

Lemmas 4.7 and 4.8 are trivial.

Lemma 4.7. *For rectangular matrices* $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, *we have*

$$|\mathrm{tr}\,(\mathbf{ABCD})| \leq \|\mathbf{A}\|\|\mathbf{C}\| \left[\mathrm{tr}\,(\mathbf{BB}^*)\right]^{1/2} \left[\mathrm{tr}\,(\mathbf{DD}^*)\right]^{1/2}.$$

Lemma 4.8. *For rectangular matrix* $\mathbf{A}$, *complex vectors* $\mathbf{a}$ *and* $\mathbf{b}$, *we have*

$$|\mathbf{a}^*\mathbf{A}\mathbf{b}| \leq \|\mathbf{A}\| \left(\mathbf{a}^*\mathbf{a}\right)^{1/2} \left(\mathbf{b}^*\mathbf{b}\right)^{1/2}.$$

Appendix A. The proof of (3.8)

By (3.11), we have that

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 - \mathrm{E}\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 \\
&= \sum_{k=1}^{p} (\mathrm{E}_k - \mathrm{E}_{k-1}) \left(\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 - \operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\right) + (\mathrm{E}_0 - \mathrm{E})\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 \\
&= \sum_{k=1}^{p} (\mathrm{E}_k - \mathrm{E}_{k-1}) \Bigg[\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)^2 + 2\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k(z)\right) \\
&\quad + \frac{1}{n^2}\lambda_k^2\beta_k^2(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)^2 + \frac{1}{n^2}\lambda_k^2\beta_k^2(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)^2 \\
&\quad + \frac{2}{n^2}\lambda_k^2\beta_k^2(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k \\
&\quad - \frac{2}{n}\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k \\
&\quad - \frac{2}{n}\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k \\
&\quad - \frac{2}{n}\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k \\
&\quad - \frac{2}{n}\lambda_k\beta_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\Bigg] + (\mathrm{E}_0 - \mathrm{E})\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2 \\
&\triangleq \sum_{k=1}^{p}(P_{k1} + 2P_{k2} + P_{k3} + P_{k4} + 2P_{k5} - 2P_{k6} - 2P_{k7} - 2P_{k8} - 2P_{k9}) \\
&\quad + (\mathrm{E}_0 - \mathrm{E})\operatorname{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2.
\end{aligned}$$

Firstly, by (3.6), it is apparent that

$$\begin{aligned}
\left|\frac{|y_{kj}|^2}{\rho_j^2} - \frac{1}{p}\right| &\leq |y_{kj}|^2\left|\frac{1}{\rho_j^2} - \frac{1}{p}\right| + \frac{1}{p}\left||y_{kj}|^2 - 1\right| \\
&\leq \frac{\delta_n}{(1-\delta_n)p}|y_{kj}|^2 + \frac{1}{p}\left||y_{kj}|^2 - 1\right| \\
&\leq \frac{1}{p}|y_{kj}|^2 + \frac{1}{p}\left||y_{kj}|^2 - 1\right|. \qquad \text{(A.1)}
\end{aligned}$$

Applying Lemma 4.7 and the above inequality, one has that

$$\begin{aligned}
\mathrm{E}\left|\sum_{k=1}^{p} P_{k1}\right|^2 &= \sum_{k=1}^{p}\mathrm{E}\left|\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)^2\right|^2 \leq Cn^2\sum_{k=1}^{p}\mathrm{E}\left\|\mathbb{D}_k\right\|^4 \\
&\leq Cn^2\max_{j=1,\cdots,n}\sum_{k=1}^{p}\mathrm{E}\left(\frac{|y_{kj}|^2}{\rho_j^2} - \frac{1}{p}\right)^4 \\
&\leq \frac{C}{n} + \frac{Cn^2}{p^4}\max_{j=1,\cdots,n}\sum_{k=1}^{p}\mathrm{E}\left(|y_{kj}|^2 - 1\right)^4 \to 0.
\end{aligned}$$

This yields that

$$\sum_{k=1}^{p} P_{k1} = O_p(1).$$

Secondly, we deal with $\sum_{k=1}^{p} P_{k2}$. It follows from Lemma 4.1 and (3.6) that

$$\begin{aligned}
&\mathrm{E}\left|\sum_{k=1}^{p} P_{k2}\right|^2 \\
&\le C\sum_{k=1}^{p}\mathrm{E}\left|\sum_{j=1}^{n}\left(\frac{|y_{kj}|^2}{\rho_j^2}-\frac{1}{p}\right)\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right|^2 \\
&\le C\sum_{k=1}^{p}\sum_{j=1}^{n}\sum_{\ell=1}^{n}\mathrm{E}\left(|y_{kj}|^2|y_{k\ell}|^2\right)\mathrm{E}\left|\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\left(\frac{1}{\rho_\ell^2}-\frac{1}{p}\right)\right|\mathrm{E}\left\|\mathbf{D}_k\right\|^2 \\
&\quad+\frac{C}{p}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left||y_{kj}|^2\left(|y_{kj}|^2-1\right)\right|\mathrm{E}\left|\frac{1}{\rho_j^2}-\frac{1}{p}\right|\mathrm{E}\left\|\mathbf{D}_k\right\|^2 \\
&\quad+\frac{C}{p^2}\sum_{k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left(|y_{kj}|^2-1\right)^2\mathrm{E}\left\|\mathbf{D}_k\right\|^2 \\
&\le \frac{C}{p}\sum_{k=1}^{p}\mathrm{E}\left\|\mathbf{D}_k\right\|^2 \to 0.
\end{aligned}$$

Thus, we obtain that $\sum_{k=1}^{p} P_{k2} = O_p(1)$.

Thirdly, by (3.23), we get that

$$\begin{aligned}
&\sum_{k=1}^{p} P_{k3} \\
&= \frac{1}{n^2}\sum_{k=1}^{p}\lambda_k^2\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\Big[\widetilde{\beta}_k^2(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right. \\
&\quad\left.-\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)^2\Big] \\
&\quad+\frac{2}{n^2}\sum_{k=1}^{p}\lambda_k^2\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\Big[\widetilde{\beta}_k^2(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right) \\
&\qquad\qquad\cdot\operatorname{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\Big] \\
&\quad-\frac{2}{n^3}\sum_{k=1}^{p}\lambda_k^3\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\left(\beta_k(z)\widetilde{\beta}_k^2(z)\varepsilon_k(z)\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)^2\right) \\
&\quad+\frac{1}{n^4}\sum_{k=1}^{p}\lambda_k^4\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\left(\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right)^2 \\
&\triangleq \mathcal{J}_1+2\mathcal{J}_2-2\mathcal{J}_3+\mathcal{J}_4.
\end{aligned}$$

Due to Lemma 4.1 and Lemma 4.5, we have that

$$\begin{aligned}\mathrm{E}\,|\mathcal{J}_1|^2 &\le \frac{C}{n^4}\sum_{k=1}^{p}\mathrm{E}\left|\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right|^4\\ &\le \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\,\|\mathbf{D}_k\|^4\to 0\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}\,|\mathcal{J}_2|^2 &\le \frac{C}{n^3}\sum_{k=1}^{p}\mathrm{E}\left|\|\mathbf{D}_k\|\left(\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)\right|^2\\ &\le \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\,\|\mathbf{D}_k\|^4\to 0.\end{aligned}$$

By the Cauchy-Schwarz inequality and Lemma 4.5,

$$\begin{aligned}\mathrm{E}\,|\mathcal{J}_3|^2 &\le \frac{C}{n^6}\sum_{k=1}^{p}\mathrm{E}\left[|\varepsilon_k(z)|^2\left|\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^4\right]\\ &\le \frac{C}{n}\sum_{k=1}^{p}\mathrm{E}^{1/2}\,\|\mathbf{D}_k\|^8\to 0\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}\,|\mathcal{J}_4|^2 &\le \frac{C}{n^8}\sum_{k=1}^{p}\mathrm{E}\left|\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^4\\ &\le \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}^{1/2}\,\|\mathbf{D}_k\|^8\to 0.\end{aligned}$$

Consequently, one has

$$\sum_{k=1}^{p}P_{k3}=O_p(1).$$

Now, we deal with $P_{k4}-P_{k8}$. Use Lemma 4.5 to get for any $\ell>0$

$$\mathrm{E}\left(\mathbf{q}_k^*\mathbf{q}_k\right)^{\ell}\le C_\ell\,\mathrm{E}\left(\mathbf{q}_k^*\mathbf{q}_k-\mathrm{tr}\left(\mathbf{\Phi}_n\right)\right)^{\ell}+C_\ell\,\mathrm{E}\left(\mathrm{tr}\left(\mathbf{\Phi}_n\right)\right)^{\ell}\le C_\ell n^{\ell}. \tag{A.2}$$

Using Lemma 4.8, (A.1), and (A.2), we get that

$$\begin{aligned}\mathrm{E}\left|\sum_{k=1}^{p}P_{k4}\right| &\le \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left|\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^2\\ &\le \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left(\|\mathbb{D}_k\|\,\mathbf{q}_k^*\mathbf{q}_k\right)^2\le\frac{C}{n^2}\max_{j=1,\cdots,n}\sum_{k=1}^{p}\mathrm{E}\left[\left(\mathbf{q}_k^*\mathbf{q}_k\right)^2\left(\frac{|y_{kj}|^2}{\rho_k^2}-\frac{1}{p}\right)^2\right]\\ &\le \frac{C}{n^2}\max_{j=1,\cdots,n}\sum_{k=1}^{p}\mathrm{E}^{1/2}\left(\mathbf{q}_k^*\mathbf{q}_k\right)^4\mathrm{E}^{1/2}\left(\frac{|y_{kj}|^2}{\rho_j^2}-\frac{1}{p}\right)^4\le\frac{C}{n}\to 0,\end{aligned}$$

and

$$
\begin{aligned}
&\mathrm{E}\left|\sum_{k=1}^{p}\left(2P_{k5}-2P_{k6}-2P_{k7}-2P_{k8}\right)\right| \\
&\leq \frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left(\left\|\mathbf{D}_k\right\|\left\|\mathbb{D}_k\right\|\left(\mathbf{q}_k^*\mathbf{q}_k\right)^2\right) \\
&\quad+\frac{C}{n}\sum_{k=1}^{p}\mathrm{E}\left(\left\|\mathbf{D}_k\right\|\left\|\mathbb{D}_k\right\|\mathbf{q}_k^*\mathbf{q}_k\right)+\frac{C}{n}\sum_{k=1}^{p}\mathrm{E}\left(\left\|\mathbb{D}_k\right\|^2\mathbf{q}_k^*\mathbf{q}_k\right) \\
&\leq C\sum_{k=1}^{p}\mathrm{E}\left(\left\|\mathbf{D}_k\right\|\mathrm{E}_{(j)}^{1/2}\left\|\mathbb{D}_k\right\|^2\right)+C\sum_{k=1}^{p}\mathrm{E}^{1/2}\left\|\mathbb{D}_k\right\|^4 \\
&\leq \frac{C}{n}\sum_{k=1}^{p}\mathrm{E}\left\|\mathbf{D}_k\right\|+\frac{C}{n}\to 0,
\end{aligned}
$$

where $\mathrm{E}_{(j)}=\mathrm{E}\left(\cdot\mid\mathcal{F}_j\right)$ and

$$
\mathcal{F}_j=\sigma\{\mathbf{q}_1,\cdots,\mathbf{q}_{j-1},\mathbf{q}_{j+1},\cdots,\mathbf{q}_p,\rho_1,\cdots,\rho_n\}.
$$

Then, we see that

$$
\sum_{k=1}^{p}\left(P_{k4}+2P_{k5}-2P_{k6}-2P_{k7}-2P_{k8}\right)=o_p(1).
$$

Now, we are in position to show that

$$
\sum_{k=1}^{p}P_{k9}=o_p(1).
$$

It follows from (3.23) that

$$
\begin{aligned}
&\sum_{k=1}^{p}P_{k9} \\
&=\frac{1}{n}\sum_{k=1}^{p}\mathrm{E}_k\left[\lambda_k\widetilde{\beta}_k(z)\left(\mathbf{q}_k^*\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\operatorname{tr}\left(\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)\right] \\
&\quad-\frac{1}{n^2}\sum_{k=1}^{p}\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\lambda_k^2\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k.
\end{aligned}
$$

By Lemma 4.1 and Lemma 4.5, one finds that

$$
\begin{aligned}
&\mathrm{E}\left|\frac{1}{n}\sum_{k=1}^{p}\mathrm{E}_k\left[\lambda_k\widetilde{\beta}_k(z)\left(\mathbf{q}_k^*\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\operatorname{tr}\left(\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right)\right]\right|^2 \\
&\leq\frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left|\mathbf{q}_k^*\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{q}_k-\operatorname{tr}\left(\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\right)^2\mathbf{J}_k^{-1}(z)\mathbf{\Phi}_n\right)\right|^2 \\
&\leq\frac{C}{n}\sum_{k=1}^{p}\mathrm{E}\left\|\mathbf{D}_k\right\|^4\to 0,
\end{aligned}
$$

and

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{n^2}\sum_{k=1}^{p}\left(\mathrm{E}_k-\mathrm{E}_{k-1}\right)\lambda_k^2\beta_k(z)\widetilde{\beta}_k(z)\varepsilon_k(z)\mathbf{q}_k^*\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbf{q}_k\right|^2\\
&\le \frac{C}{n^4}\sum_{k=1}^{p}\mathrm{E}\left|\|\mathbf{D}_k\|^2\varepsilon_k(z)\left(\mathbf{q}_k^*\mathbf{q}_k-\mathrm{tr}\left(\mathbf{\Phi}_n\right)\right)\right|^2+\frac{C}{n^2}\sum_{k=1}^{p}\mathrm{E}\left|\|\mathbf{D}_k\|^2\varepsilon_k(z)\right|^2\\
&\le \frac{C}{n}\sum_{k=1}^{p}\mathrm{E}\,\|\mathbf{D}_k\|^4\to 0.
\end{aligned}$$

Hence,

$$\sum_{k=1}^{p}P_{k9}=o_p(1).$$

Combining the above results, we conclude that

$$\mathrm{tr}(\mathbf{J}^{-1}(z)\mathbf{D})^2-\mathrm{E}\,\mathrm{tr}(\mathbf{J}^{-1}(z)\mathbf{D})^2=\mathrm{E}_0\,\mathrm{tr}(\mathbf{J}^{-1}(z)\mathbf{D})^2-\mathrm{E}(\mathbf{J}^{-1}(z)\mathbf{D})^2+o_p(1). \tag{A.3}$$

Similarly to the proof of (3.18), we get that

$$\begin{aligned}
(\mathrm{E}_0-\mathrm{E})\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}\right)^2&=\sum_{k=1}^{p}(\mathrm{E}_0-\mathrm{E})\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)\\
&\quad+\sum_{k=1}^{p}(\mathrm{E}_0-\mathrm{E})\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)+o_p(1).
\end{aligned} \tag{A.4}$$

Using (3.12), one can obtain that

$$\begin{aligned}
&\sum_{k=1}^{p}(\mathrm{E}_0-\mathrm{E})\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)\\
&=\frac{1}{p^3}\sum_{j\neq\ell=1}^{n}\mathrm{E}_0\left[\left(\rho_j^2-p\right)\left(\rho_\ell^2-p\right)\right]\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j(1+o_p(1))\\
&\quad+\frac{3}{p^3}\sum_{j=1}^{n}(\mathrm{E}_0-\mathrm{E})\left(\rho_j^2-p\right)^2\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)^2(1+o_p(1))\\
&\quad+\frac{4}{p^2}\sum_{j=1}^{n}\mathrm{E}_0\left(\rho_j^2-p\right)\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right)^2(1+o_p(1))+o_p(1).
\end{aligned}$$

After some calculations, one finds that

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{p^3}\sum_{j\neq\ell=1}^{n}\mathrm{E}_0\left[\left(\rho_j^2-p\right)\left(\rho_\ell^2-p\right)\right]\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^{-1}\mathbf{e}_j\right|^2\\
&\le\frac{C}{p^6}\sum_{j\neq\ell=1}^{n}\mathrm{E}\left(\rho_j^2-p\right)^2\mathrm{E}\left(\rho_\ell^2-p\right)^2\le\frac{C}{p^2}\to 0.
\end{aligned}$$

Other terms can be handled similarly. Hence,

$$\sum_{k=1}^{p}(\mathrm{E}_0-\mathrm{E})\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k\mathbf{J}_k^{-1}(z)\mathbb{D}_k\right)=o_p(1). \tag{A.5}$$

Similarly it can be verified that

$$\sum_{k=1}^{p} (\mathrm{E}_0 - \mathrm{E}) \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbf{D}_k \mathbf{J}_k^{-1}(z) \mathbb{D}_k \right) = o_p(1). \tag{A.6}$$

Together with (A.3)–(A.6), we see that

$$\operatorname{tr}(\mathbf{J}^{-1}(z)\mathbf{D})^2 - \mathrm{E} \operatorname{tr}(\mathbf{J}^{-1}(z)\mathbf{D})^2 \xrightarrow{p} 0.$$

Appendix B. The limit of $-\frac{1}{8} \mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^2 \right) + \frac{1}{4} \mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^* \mathbf{D} \right) - \frac{1}{8} \mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \left(\mathbf{D}^* \right)^2 \right)$

To begin with, we shall analyse $\mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^* \mathbf{D} \right)$. Note that

$$\begin{aligned}
\mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^* \mathbf{D} \right) &= \sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^* \mathbb{D}_k \right) \\
&= \sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbb{D}_k^* \mathbb{D}_k \right) + \sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbf{D}_k^* \mathbb{D}_k \right) \\
&\quad - \frac{1}{n} \sum_{k=1}^{p} \lambda_k \, \mathrm{E} \left(\beta_k(z) \mathbf{q}_k^* \mathbf{J}_k^{-1}(z) \mathbf{D}^* \mathbb{D}_k \mathbf{J}_k^{-1}(z) \mathbf{q}_k \right).
\end{aligned}$$

Using Lemma 4.1, Lemma 4.5, and (3.17), one finds that

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{k=1}^{p} \lambda_k \, \mathrm{E} \left(\beta_k(z) \mathbf{q}_k^* \mathbf{J}_k^{-1}(z) \mathbf{D}^* \mathbb{D}_k \mathbf{J}_k^{-1}(z) \mathbf{q}_k \right) \right| \\
&\leq \frac{C}{n} \sum_{k=1}^{p} \mathrm{E} \left(|\beta_k(z)| \left\| \mathbf{J}_k^{-2}(z) \mathbf{D} \mathbb{D}_k \right\| \mathbf{q}_k^* \mathbf{q}_k \right) \\
&\leq \frac{C}{n} \sum_{k=1}^{p} \mathrm{E}^{1/4} \left| \beta_k(z) \right|^4 \mathrm{E}^{1/4} \left| \mathbf{q}_{k_1}^* \mathbf{q}_{k_1} \right|^4 \mathrm{E}^{1/4} \left\| \mathbb{D}_k \right\|^4 \mathrm{E}^{1/8} \left\| \mathbf{J}_k^{-2}(z) \right\|^8 \mathrm{E}^{1/8} \left\| \mathbf{D} \right\|^8 \to 0.
\end{aligned}$$

Hence, we get that

$$\mathrm{E} \operatorname{tr} \left(\mathbf{J}^{-1}(z) \mathbf{D}^* \mathbf{D} \right) = \sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbb{D}_k^* \mathbb{D}_k \right) + \sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbf{D}_k^* \mathbb{D}_k \right) + o(1).$$

By (3.13) and (3.14),

$$\begin{aligned}
&\sum_{k=1}^{p} \mathrm{E} \operatorname{tr} \left(\mathbf{J}_k^{-1}(z) \mathbb{D}_k^* \mathbb{D}_k \right) \\
&= \sum_{k=1}^{p} \sum_{j=1}^{n} \mathrm{E} \left(\left(\frac{|y_{kj}|^2}{\rho_j^2} - \frac{1}{p} \right)^2 \mathbf{e}_j^T \left(\mathbf{B}_n^* \right)^{-1} \mathbf{B}_n^{-1} \mathbf{e}_j \mathbf{e}_j^T \mathbf{B}_n \mathbf{J}_k^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j \right) \\
&\quad + \sum_{k=1}^{p} \sum_{j \neq \ell = 1}^{n} \mathrm{E} \left(\frac{1}{\rho_j^2} - \frac{1}{p} \right) \mathrm{E} \left(\frac{1}{\rho_\ell^2} - \frac{1}{p} \right) \mathrm{E} \left(\mathbf{e}_j^T \left(\mathbf{B}_n^* \right)^{-1} \mathbf{B}_n^{-1} \mathbf{e}_\ell \mathbf{e}_\ell^T \mathbf{B}_n \mathbf{J}_k^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j \right) + o(1) \\
&= \frac{1 + |\mathrm{E}\, y_{11}^2|}{p} \sum_{j=1}^{n} \mathrm{E} \left(\mathbf{e}_j^T \left(\mathbf{B}_n \mathbf{B}_n^* \right)^{-1} \mathbf{e}_j \mathbf{e}_j^T \mathbf{B}_n \mathbf{J}^{-1}(z) \mathbf{B}_n^* \mathbf{e}_j \right) + o(1),
\end{aligned}$$

and

$$
\begin{aligned}
&\sum_{k=1}^{p} \mathrm{E}\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k^*\mathbb{D}_k\right)\\
&= \sum_{k=1}^{p}\sum_{j=1}^{n} \mathrm{E}\left(\left(\frac{1}{\rho_j^2}-\frac{1}{p}\right)\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{D}_k^*\mathbf{B}_n^{-1}\mathbf{e}_j\right)\\
&= \frac{\tau}{p^3}\sum_{k=1}^{p}\sum_{\ell\neq k=1}^{p}\sum_{j=1}^{n}\mathrm{E}\left(|y_{\ell j}|^2\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}_k^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\right)\\
&\quad+\frac{\tau}{p^3}\sum_{k=1}^{p}\sum_{\ell\neq k=1}^{p}\sum_{t=1}^{n}\mathrm{E}\left(\left(|y_{\ell t}|^2-1\right)\mathbf{e}_t^T\left(\mathbf{B}_n^*\right)^{-1}\mathbf{J}_k^{-1}(z)\mathbf{B}_n^*\mathbf{e}_t\right)+o(1)\\
&= \frac{\tau}{p}\sum_{j=1}^{n}\mathrm{E}\left(\mathbf{e}_j^T\mathbf{B}_n\mathbf{J}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\right)+o(1).
\end{aligned}
$$

Together with the above three equalities, we see that

$$
\begin{aligned}
&\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right)\\
&= \frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{p}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{W}^{-1}(z)\mathbf{B}_n^*\mathbf{e}_j+o(1)\\
&\to -\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{zp}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\left(g_1(z)\mathbf{B}_n\mathbf{B}_n^*+\mathbf{I}\right)^{-1}\mathbf{e}_j.
\end{aligned}
\tag{B.1}
$$

Similarly to the proof of (B.1), we can get that

$$
\begin{aligned}
\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right) &= \sum_{k=1}^{p}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbb{D}_k^2\right)+\sum_{k=1}^{p}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}_k^{-1}(z)\mathbf{D}_k\mathbb{D}_k\right)+o(1)\\
&= \frac{\tau+2}{p}\,\mathrm{tr}\left(\mathbf{W}^{-1}(z)\right)+o(1)\\
&\to \left(\tau+1+|\mathrm{E}\,y_{11}^2|\right)c^{-1}\underline{m}(z),
\end{aligned}
$$

and

$$
\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right)\to\left(\tau+1+|\mathrm{E}\,y_{11}^2|\right)c^{-1}\underline{m}(z).
$$

Combining the above results, we conclude that

$$
\begin{aligned}
&-\frac{1}{8}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^2\right)+\frac{1}{4}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\mathbf{D}^*\mathbf{D}\right)-\frac{1}{8}\mathrm{E}\,\mathrm{tr}\left(\mathbf{J}^{-1}(z)\left(\mathbf{D}^*\right)^2\right)\\
&\to -\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4c}\underline{m}(z)-\frac{\tau+1+|\mathrm{E}\,y_{11}^2|}{4zp}\sum_{j=1}^{n}\mathbf{e}_j^T\left(\mathbf{B}_n\mathbf{B}_n^*\right)^{-1}\mathbf{e}_j\mathbf{e}_j^T\mathbf{B}_n\mathbf{B}_n^*\mathbb{G}_{\triangle}(z)\mathbf{e}_j.
\end{aligned}
$$

Appendix C. The proof of Theorem 2.3

When the diffusion process $\mathbf{x}_t$ belongs to class $\mathcal{C}$, the drift process $\mu_t \equiv 0$ and $\tau_{n,l}$'s and γ_t are both nonrandom and independent of $\mathbf{w}_t$, we know that $\Delta\mathbf{x}_l$ follows the same distribution as

$$
\sqrt{\int_{\tau_{n,l-1}}^{\tau_{n,l}}\gamma_t^2\,dt}\cdot\breve{\Sigma}^{1/2}\mathbf{z}_l,
$$

where $\{\mathbf{z}_1, \cdots, \mathbf{z}_n\}$ follow the p-dimensional standard normal distribution. Then denoting $\omega_l^n = \int_{\tau_{n,l-1}}^{\tau_{n,l}} \gamma_t^2 dt$, we shall rewrite the RCV matrix as

$$\Sigma^{RCV} \stackrel{d}{=} \sum_{l=1}^{n} \omega_l^n \breve{\Sigma}^{1/2} \mathbf{z}_l \mathbf{z}_l^T \breve{\Sigma}^{1/2} = \frac{1}{n} \mathbf{A}_n \mathbf{Z}_n \mathbf{B}_n \mathbf{B}_n^T \mathbf{Z}_n^T \mathbf{A}_n^T,$$

where $\mathbf{A}_n = \breve{\Sigma}^{1/2}$, $\mathbf{B}_n \mathbf{B}_n^T = \mathrm{diag}(n\omega_1^n, \cdots, n\omega_n^n)$ and $\mathbf{Z} = (\mathbf{z}_1, \cdots, \mathbf{z}_n)$. This is in fact a special case of separable sample covariance matrix from elliptical population with $\tau = 2$. Then this theorem follows from Theorem 1.4.

Appendix D. The proof of Lemmas 4.1 and 4.2

D.1. **The proof of Lemma 4.1.** To prove the first conclusion of this lemma, it suffices to prove that for any even $k \geq 2$,

$$\mathrm{E} \| \mathrm{Diag}(\|\mathbf{y}_1\|^2/\rho_1^2, \cdots, \|\mathbf{y}_\mathrm{n}\|^2/\rho_\mathrm{n}^2) - \mathbf{I}_n \|^k \to 0$$

as $n \to \infty$. To achieve this, we only need to prove that for any $1 \leq j \leq n$, $\mathrm{E}\left(\|\mathbf{y}_j\|^2/\rho_j^2 - 1\right)^k$ is sufficiently small. Note that by assumption (a) and (3.6), we have

$$\begin{aligned} \mathrm{E}\left(\|\mathbf{y}_j\|^2/\rho_j^2 - 1\right)^k &\leq \frac{C_k \left(\mathrm{E}\left(\|\mathbf{y}_j\|^2 - p\right)^k + \mathrm{E}\left(\rho_j^2 - p\right)^k\right)}{(1-\delta_n)^k p^k} \\ &\leq \frac{C_k \left(p^{k/2} + p^{k-1}\right)}{(1-\delta_n)^k p^k} = o(p^{-1}). \end{aligned} \tag{D.1}$$

This leads to the first conclusion. Taking the same procedure, we find that $\mathrm{E}\|\mathbb{D}_s\|^k = o(p^{-1})$ for any $s = 1, \cdots, n$ and thus the second conclusion follows.

D.2. **The proof of Lemma 4.2.** The second conclusion follows from the conclusion of Lemma 4.1 with $k = 2$. Now we deal with the first conclusion. Note that

$$\begin{aligned} &\mathrm{P}\left(\left|\frac{\rho_1^2}{\|y_1\|^2} - 1\right| > \epsilon\right) \\ &\leq \mathrm{P}\left(\|y_1\|^2 < p/2\right) + \mathrm{P}\left(\|y_1\|^2 \geq p/2, \left|\frac{\rho_1^2}{\|y_1\|^2} - 1\right| > \epsilon\right) \\ &\leq \mathrm{P}\left(\|y_1\|^2 < p/2\right) + p^{-l}(\epsilon/2)^{-l}\,\mathrm{E}\left|\frac{\rho_1^2}{\|y_1\|^2} - 1\right|^l \\ &\leq \mathrm{P}\left(\|y_1\|^2 < p/2\right) + 2p^{-l}(\epsilon/2)^{-l}\left(\mathrm{E}\left|\rho_1^2 - p\right|^l + \mathrm{E}\left|\|y_1\|^2 - p\right|^l\right). \end{aligned} \tag{D.2}$$

From the Chernoff bounds for chi-square random variable and the moments for chi-square random variable, we have that $\mathrm{P}\left(\|y_1\|^2 < p/2\right) = o(n^{-t})$ and $2p^{-l}(\epsilon/2)^{-l}\,\mathrm{E}\left|\|y_1\|^2 - p\right|^l = o(n^{-t})$ for any given $t > 0$. Also, we have that

$$2p^{-l}(\epsilon/2)^{-l}\,\mathrm{E}\left|\rho_1^2 - p\right|^l \leq 2\left(\frac{\epsilon}{2\delta_n}\right)^{-l},$$

which is $o(n^{-t})$ for any $t > 0$ by choosing $l = \log n$.

Acknowledgments

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

References

[1] T. W. Anderson, *An introduction to multivariate statistical analysis*, 3rd ed., Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2003. MR1990662

[2] Z. D. Bai and Jack W. Silverstein, *CLT for linear spectral statistics of large-dimensional sample covariance matrices*, Ann. Probab. **32** (2004), no. 1A, 553–605, DOI 10.1214/aop/1078415845. MR2040792

[3] Zhidong Bai and Jack W. Silverstein, *Spectral analysis of large dimensional random matrices*, 2nd ed., Springer Series in Statistics, Springer, New York, 2010, DOI 10.1007/978-1-4419-0661-8. MR2567175

[4] Zhidong Bai, Huiqin Li, and Guangming Pan, *Central limit theorem for linear spectral statistics of large dimensional separable sample covariance matrices*, Bernoulli **25** (2019), no. 3, 1838–1869, DOI 10.3150/18-BEJ1038. MR3961233

[5] Zhigang Bao, Guangming Pan, and Wang Zhou, *Asymptotic mutual information statistics of MIMO channels and CLT of sample covariance matrices*, IEEE Trans. Inform. Theory **61** (2015), no. 6, 3413–3426, DOI 10.1109/TIT.2015.2421894. MR3352506

[6] Arjun K. Gupta, Tamas Varga, and Taras Bodnar, *Elliptically contoured models in statistics and portfolio theory*, 2nd ed., Springer, New York, 2013, DOI 10.1007/978-1-4614-8154-6. MR3112145

[7] Walid Hachem, Malika Kharouf, Jamal Najim, and Jack W. Silverstein, *A CLT for information-theoretic statistics of non-centered Gram random matrices*, Random Matrices Theory Appl. **1** (2012), no. 2, 1150010, 50, DOI 10.1142/S2010326311500109. MR2934716

[8] Jiang Hu, Weiming Li, and Wang Zhou, *Central limit theorem for mutual information of large MIMO systems with elliptically correlated channels*, IEEE Trans. Inform. Theory **65** (2019), no. 11, 7168–7180, DOI 10.1109/TIT.2019.2913760. MR4030880

[9] Jiang Hu, Weiming Li, Zhi Liu, and Wang Zhou, *High-dimensional covariance matrices in elliptical distributions with application to spherical test*, Ann. Statist. **47** (2019), no. 1, 527–555, DOI 10.1214/18-AOS1699. MR3909941

[10] M. Ajith Kamath and Brian L. Hughes, *The asymptotic capacity of multiple-antenna Rayleigh-fading channels*, IEEE Trans. Inform. Theory **51** (2005), no. 12, 4325–4333, DOI 10.1109/TIT.2005.858966. MR2243163

[11] Bo Li, Marc G. Genton, and Michael Sherman, *Testing the covariance structure of multivariate random fields*, Biometrika **95** (2008), no. 4, 813–829, DOI 10.1093/biomet/asn053. MR2461213

[12] Huiqin Li, Yanqing Yin, and Shurong Zheng, *Central limit theorem for linear spectral statistics of general separable sample covariance matrices with applications*, J. Statist. Plann. Inference **211** (2021), 80–89, DOI 10.1016/j.jspi.2020.06.003. MR4111759

[13] M. W. Mitchell and M. L. Gumpertz, *Spatio-temporal prediction inside a free-air CO2 enrichment system*, J. Agric. Biol. Environ. Stat. **8** (2003), no. 3, 310–327.

[14] Debashis Paul and Jack W. Silverstein, *No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix*, J. Multivariate Anal. **100** (2009), no. 1, 37–57, DOI 10.1016/j.jmva.2008.03.010. MR2460475

[15] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, *Fading correlation and its effect on the capacity of multielement antenna systems*, IEEE Trans. Commun. **48** (2000), no. 3, 502–513.

[16] S. Verdú, *Multiuser detection*, Cambridge University Press, 1998.

[17] Sergio Verdú, *Spectral efficiency in the wideband regime*, IEEE Trans. Inform. Theory **48** (2002), no. 6, 1319–1343, DOI 10.1109/TIT.2002.1003824. Special issue on Shannon theory: perspective, trends, and applications. MR1908977

[18] Lixin Zhang, *Spectral analysis of large dimentional random matrices*, Ph.D. Thesis, National University of Singapore, 2006.

[19] Xinghua Zheng and Yingying Li, *On the estimation of integrated covariance matrices of high dimensional diffusion processes*, Ann. Statist. **39** (2011), no. 6, 3121–3151, DOI 10.1214/11-AOS939. MR3012403

School of Mathematics and Statistics, Chongqing University, Chongqing, People's Republic of China

Email address: `lihq118@nenu.edu.cn`

Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

Email address: `gmpan@ntu.edu.sg`

School of Mathematics and Statistics, Chongqing University, Chongqing, People's Republic of China

Email address: `yinyq@cqu.edu.cn`

Department of Statistics and Data Science, National University of Singapore, 6 Science Drive 2, 117546 Singapore, Singapore

Email address: `wangzhou@nus.edu.sg`