The Annals of Statistics
2021, Vol. 49, No. 3, 1569–1593
https://doi.org/10.1214/20-AOS2013

CENTRAL LIMIT THEOREM FOR LINEAR SPECTRAL STATISTICS OF LARGE DIMENSIONAL KENDALL'S RANK CORRELATION MATRICES AND ITS APPLICATIONS

BY ZENG LI[1], QINWEN WANG[2] AND RUNZE LI[3]

[1]*Department of Statistics and Data Science, Southern University of Science and Technology, liz9@sustech.edu.cn*

[2]*School of Data Science, Fudan University, wqw@fudan.edu.cn*

[3]*Department of Statistics, Pennsylvania State University, rzli@psu.edu*

This paper is concerned with the limiting spectral behaviors of large dimensional Kendall's rank correlation matrices generated by samples with independent and continuous components. The statistical setting in this paper covers a wide range of highly skewed and heavy-tailed distributions since we do not require the components to be identically distributed, and do not need any moment conditions. We establish the central limit theorem (CLT) for the linear spectral statistics (LSS) of the Kendall's rank correlation matrices under the Marchenko–Pastur asymptotic regime, in which the dimension diverges to infinity proportionally with the sample size. We further propose three nonparametric procedures for high dimensional independent test and their limiting null distributions are derived by implementing this CLT. Our numerical comparisons demonstrate the robustness and superiority of our proposed test statistics under various mixed and heavy-tailed cases.

1. Introduction. Most well-established statistics in classical multivariate analysis can be presented as linear functionals of eigenvalues of sample covariance or correlation matrix models. Such linear functionals of eigenvalues are termed as *linear spectral statistics (LSS)* in the literature of *Random Matrix Theory (RMT)*. Studies of asymptotic properties of the LSS are particularly important in multivariate analysis of variance, multivariate linear models, canonical correlation analysis and factor analysis, etc. Analysis of high dimensional data calls for new theoretical framework since classical multivariate analysis theory may become invalid under high-dimensional regime. While RMT, which studies asymptotic behaviors of eigenvalues of large random matrices with certain structures, serves as an effective tool to deal with high dimensional problems, particularly under the *Marchenko–Pastur asymptotic regime*, where the dimension proportionally diverges to infinity with the sample size.

The study of the fluctuation of LSS for different types of random matrix models has received considerable attention in the past decades. In particular, Bai and Silverstein (2004) established CLT of LSS for large dimensional *Sample Covariance Matrices (SCM)*; other popular random matrix models include Bai and Yao (2005) for Wigner matrix, Zheng (2012) for the Fisher matrices, Yang and Pan (2015) and Gao et al. (2017) for canonical and Pearson correlation matrices. Such results have broad applications in various fields such as wireless communications and finance, etc.; see recent monographs (Bai and Silverstein (2010), Couillet and Debbah (2011), Yao, Zheng and Bai (2015)) and survey papers (Tulino and Verdú (2004), Johnstone (2006), Paul and Aue (2014)). Although the seminal result in Bai and Silverstein (2004) of LSS for SCM is universal in the sense that the CLT does not depend on any particular distribution assumption of the data, it does require the same kurtosis

Received June 2019; revised February 2020.
MSC2020 subject classifications. Primary 62H10; secondary 62H15.
Key words and phrases. Kendall's rank correlation matrices, linear spectral statistics, central limit theorem, random matrix theory, high dimensional independent test.

as Gaussian distribution. Such moment restrictions have been relaxed in subsequent developments like Pan and Zhou (2008), Zheng, Bai and Yao (2015), finite fourth or even higher order moments are still generally required in Pearson-type models. This kind of constraints significantly limit the applicability for real data analysis, but certainly have stimulated the investigation of nonparametric methods which are indeed free of moment restrictions. For example, Bao et al. (2015) established the asymptotic normality of polynomial functions of the spectrum of Spearman's rank correlation matrices. Bandeira, Lodhia and Rigollet (2017) obtained the limiting spectral distribution of Kendall's rank correlation matrices, which appeared to be a variation of the Marchenko–Pastur law. Bao (2019a) further proved the Tracy–Widom limit for largest eigenvalues of Kendall's rank correlation matrices.

In this paper, we focus on a nonparametric matrix model, the Kendall's rank correlation matrices denoted by $\mathbf{K}_n$, which is generated by samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ from a p-dimensional random vector $\mathbf{x}$. The primary goal of this paper is to establish the CLT for general LSS of $\mathbf{K}_n$ under the Marchenko–Pastur asymptotic regime when $\mathbf{x}$ consists of p independent components. From technical point of view, the LSS of $\mathbf{K}_n$ is actually the type of U-statistics. To the best of our knowledge, there are no such CLTs for LSS related to U-statistics been established in the current literature. Moreover, the structure of $\mathbf{K}_n$ is very different from most existing random matrix models. Although $\mathbf{K}_n$ can be written as Wishart-type product of certain data matrix $\boldsymbol{\Theta}$, that is, $\mathbf{K}_n = \boldsymbol{\Theta}\boldsymbol{\Theta}^\top$ (see more details in (4.6)), the elements within each row of $\boldsymbol{\Theta}$ is nonlinearly correlated and such correlation structure cannot be incorporated into any random matrix models studied so far. In fact, the most commonly used model assumption is the *independent component model (ICM)* assumption: there exist constant vector $\boldsymbol{\mu}$ and nonnegative definite matrix $\boldsymbol{\Sigma}_p$ such that $\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\Sigma}_p^{1/2}\mathbf{z}$, where all elements within $\mathbf{z}$ are i.i.d. with zero mean and unit variance (Bai and Silverstein (2004), Pan and Zhou (2008), Zheng (2012)). Indeed, techniques for the ICM cannot accommodate the nonlinear correlation structure in $\boldsymbol{\Theta}$ and novel mathematical developments are thus needed. First, for each row of $\boldsymbol{\Theta}$, derivation of explicit formula for expectation of its quadratic form is one of the key steps for deriving the CLT and the result is established in Lemma S1.5. Inequalities related to certain quadratic form's higher order moment and its centralized version also need to be reconsidered and reestablished, which are obtained in Lemmas S1.3 and S1.4. Second, our model is structured as $\mathbf{K}_n = \boldsymbol{\Theta}\boldsymbol{\Theta}^\top$ where $\boldsymbol{\Theta}$ is a $p \times M$ ($M = n(n-1)/2$) matrix with $p/M \to 0$ while in the commonly studied ICM structure, the data matrix is $p \times n$ ($p/n = O(1)$), this difference gives an extra term in our limiting mean term. As far as we know, such phenomenon has never occurred in previous works where the appearance of such term is irrelevant with the aforementioned results on quadratic forms. Third, since the LSS is expressed as a contour integral in our proof, bounds for both the smallest and largest eigenvalues of $\mathbf{K}_n$ are needed to ensure that the contour encompasses the entire spectrum of $\mathbf{K}_n$ with high probability. Existing works commonly use matrix inequalities to ascribe the desired matrix models to the ICM structure and directly apply the spectral norm bound for SCV established in Bai and Silverstein (1998). While such tricks cannot be adopted here due to the inapplicability of ICM structure. To overcome this, we have proved that no eigenvalues will reside outside the support of LSD of $\mathbf{K}_n$ for all large n. This by-product is summarized in Proposition 4.1.

To demonstrate the potential of the newly established CLT, we study high dimensional independent tests without any moment restrictions. Under Gaussian assumption, such tests of independence have been well studied in the low dimensional framework (Anderson (1984), Muirhead (2009)), where the *Likelihood Ratio Test statistic (LRT)* has a limiting chi-square null distribution. To cope with high dimensionality, various methods have been proposed, including the modifications of LRT (Jiang, Bai and Zheng (2013), Jiang and Qi (2015)); Frobenius-norm-type statistics based on sample correlation matrices (Gao et al. (2017)), sample canonical correlation coefficients (Yang and Pan (2015)) and maximum-norm-type statistics (Jiang (2004), Zhou (2007), Cai and Jiang (2011)). However, these tests are in general infeasible for heavy-tailed distributions since these approaches are based on Pearson-type covariance or correlation matrices which all require strong moment conditions. On the other hand, nonparametric approaches are another line of work in this direction, which can get rid of such moment restrictions and robustness are thus achieved by replacing the Pearson-type correlations with rank-based versions. Recent works on this topic include Bao et al. (2015) for tests based on the polynomial functions of the spectrum of Spearman's rank correlation matrices, Han, Chen and Liu (2017) for the maximum type statistics including Kendall's tau and Spearman's rho, and Leung and Drton (2018) for a class of nonparametric U-statistics. In this paper, we propose three test statistics for high dimensional independent test. The first two are based on the second and fourth spectral moments of $\mathbf{K}_n$ while the third is based on the *entropy loss (EL)* between $\mathbf{K}_n$ and its population counterpart, which is motivated by (James and Stein (1961), Muirhead (2009), Zheng et al. (2019)). All the three statistics can be represented in certain forms of LSS of $\mathbf{K}_n$, thus their limiting null distributions can be fully derived using our newly established CLT. Numerical studies demonstrate robustness of the three tests for mixed and heavy-tailed distributions. Their performances are very satisfactory for distinguishing various settings of dependent alternatives including both linear and nonlinear cases.

The paper is organized as follows. Section 2 first introduces some preliminary results on high dimensional Kendall's rank correlation matrices and then establishes our main result on the CLT for LSS of $\mathbf{K}_n$. Section 3 develops three statistics for high dimensional independent test, with their limiting null distributions explicitly derived. Simulation experiments are conducted to compare the finite sample performance of our test statistics with other existing ones under various model settings and alternatives. Section 4 contains the proof of our main result. Technical lemmas and all the remaining proofs are relegated to the Supplementary Material (Li, Wang and Li (2021)).

Throughout this paper, we use bold Roman capital letters, for example, $\mathbf{A}$ to represent matrices. $\operatorname{Tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. Scalars are in lowercase letters. Vectors are bold letters in lowercase like $\mathbf{v}$. $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ represent the sets of natural, real and complex numbers. $\mathbb{C}^+$ denotes the upper complex plane. $\mathbb{E}(\cdot)$ means taking expectation and $\Im(\cdot)$ taking imaginary part of complex numbers. $\mathbb{1}(\cdot)$ stands for indicator function. T for transpose, $*$ for conjugate transpose of vectors or matrices.

2. CLT for LSS of high-dimensional Kendall's rank correlation matrices.

Suppose $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})^\mathsf{T}, i = 1, \dots, n$ is a random sample from a p dimensional vector $\mathbf{x}$ with independent and continuous components. Denote $\mathbf{X}_n = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ as the $p \times n$ data matrix. The Kendall's rank correlation matrix of $\mathbf{X}_n$ is defined to be a $p \times p$ matrix $\mathbf{K}_n = (\tau_{k\ell})$, whose (k, ℓ)th entry is the empirical Kendall's rank correlation coefficient between the kth and ℓth components of $\mathbf{x}$ (the kth and ℓth row of $\mathbf{X}_n$), that is,

$$\tau_{k\ell} = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \operatorname{sign}(x_{ki} - x_{kj}) \cdot \operatorname{sign}(x_{\ell i} - x_{\ell j}), \quad 1 \le k, \ell \le p. \tag{2.1}$$

For any analytic function $f(x)$, a general LSS of $\mathbf{K}_n$ is defined as $\sum_{i=1}^p f(\lambda_i)/p$, where $\{\lambda_i, 1 \le i \le p\}$ are eigenvalues of $\mathbf{K}_n$. In this section, we are interested in the CLT for such LSS of $\mathbf{K}_n$ under the Marchenko–Pastur asymptotic regime: $(p, n) \to \infty$, $c_n := p/n \to c \in (0, \infty)$.

2.1. *Preliminary results on the limiting spectrum of* $\mathbf{K}_n$.

The *empirical spectral distribution* (*ESD*) of $\mathbf{K}_n$ is referred as a random measure $F^{\mathbf{K}_n}$ such that

$$F^{\mathbf{K}_n}(x) = \frac{1}{p} \sum_{i=1}^{p} \delta_{\lambda_i}, \tag{2.2}$$

where δ_{λ_i} is the Dirac mass at point λ_i. It is proven in Bandeira, Lodhia and Rigollet (2017) that $F^{\mathbf{K}_n}$ converges in probability to a limit $F_c := 2Y/3 + 1/3$ as $(p,n) \to \infty$, $p/n \to c \in (0,\infty)$, where Y follows the standard Marchenko–Pastur law with parameter c; see Marčenko and Pastur (1967). This cumulative distribution function F_c is called *limiting spectral distribution* (*LSD*) of $\mathbf{K}_n$ with its explicit form of density dF_c given by

$$\begin{aligned} dF_c(x) = &\frac{9}{4\pi c(3x-1)}\sqrt{(d_{c,+}-x)(x-d_{c,-})} \\ &+ \left(1-\frac{1}{c}\right)\delta_{\frac{1}{3}}\mathbb{1}_{\{c>1\}}, \quad d_{c,-} \le x \le d_{c,+}, \end{aligned} \tag{2.3}$$

where

$$d_{c,-} = \frac{1}{3} + \frac{2}{3}(1-\sqrt{c})^2, \qquad d_{c,+} = \frac{1}{3} + \frac{2}{3}(1+\sqrt{c})^2. \tag{2.4}$$

Moreover, the *Stieltjes transform* $m_{F_c}(z)$ corresponding to $F_c(x)$, which is defined to be $m_{F_c}(z) = \int \frac{1}{x-z}\,\mathrm{d}F_c(x)$, $z \in \mathbb{C}^+$, is the unique solution of the following equation:

$$\frac{2}{3}c\left(z-\frac{1}{3}\right)m_{F_c}^2(z) + \left(z - 1 + \frac{2}{3}c\right)m_{F_c}(z) + 1 = 0 \tag{2.5}$$

such that $\Im(z)\cdot\Im(m_{F_c}(z)) > 0$. It can also be expressed explicitly as a function of the limiting dimension-to-sample size ratio c, that is,

$$m_{F_c}(z) = \frac{1-\frac{2}{3}c - z + \sqrt{(z-1-\frac{2}{3}c)^2 - \frac{16}{9}c}}{\frac{4}{3}c(z-\frac{1}{3})}.$$

Replacing c with c_n, we denote by $F^{c_n}(x)$, $m_{F^{c_n}}(z)$, $d_{c_n,-}$, $d_{c_n,+}$ the analogues for $F_c(x)$, $m_{F_c}(z)$, $d_{c,-}$, $d_{c,+}$, respectively.

2.2. *Main results*. Denote the eigenvalues of $\mathbf{K}_n$ as $\lambda_1,\dots,\lambda_p$ in the descending order. Of interest is the asymptotic behavior of $\sum_{i=1}^p f(\lambda_j)/p$, the LSS of $\mathbf{K}_n$, where $f(x)$ is an analytic function. Note that

$$\frac{1}{p}\sum_{i=1}^{p} f(\lambda_j) = \int f(x)\,\mathrm{d}F^{\mathbf{K}_n}(x). \tag{2.6}$$

As has already been established that the ESD of $\mathbf{K}_n$ converges in probability to F_c, then asymptotically, the quantity (2.6) will tend to $\int f(x)\,\mathrm{d}F_c(x)$ almost surely. In the following, we will show that the convergence rate of $\int f(x)\,\mathrm{d}F^{\mathbf{K}_n}(x) - \int f(x)\,\mathrm{d}F^{c_n}(x)$ is essentially $1/p$. To this end, define

$$G_n(x) = p\big(F^{\mathbf{K}_n}(x) - F^{c_n}(x)\big),$$

our main result is stated in the following theorem.

THEOREM 2.1. *Suppose* $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ *is a random sample from a* p*-dimensional population* $\mathbf{x}$, *where* $\mathbf{x}$ *has* p *independent components*, *all of which are absolutely continuous with respect to the Lebesgue measure. Let* $f_1,\dots,f_k$ *be functions on* $\mathbb{R}$ *and analytic on an open interval containing the support of* $dF_c(x)$ (*defined in* (2.3)), *then as* $(p,n)\to\infty$, $p/n \to c \in (0,\infty)$, *the random vector*

$$\left(\int f_1(x)\,\mathrm{d}G_n(x), \dots, \int f_k(x)\,\mathrm{d}G_n(x)\right)$$

forms a tight sequence in n and converges weakly to a Gaussian random vector $(X_{f_1}, \dots, X_{f_k})$, *with means*

$$
\begin{aligned}
\mathbb{E}X_f = &-\frac{1}{2\pi i}\oint_{\gamma} f(z)\left\{\frac{36cm^3(z)(1+\frac{2}{3}cm(z))}{[-9(1+\frac{2}{3}cm(z))^2+4cm^2(z)]^2}\right. \\
&\left.-\frac{2c^2m^3(z)[(1+\frac{2}{3}cm(z))^2+6+\frac{4}{3}cm(z)]}{-9(1+\frac{2}{3}cm(z))^2+4cm^2(z)}\right\}\mathrm{d}z \\
&-\frac{1}{2\pi i}\oint_{\gamma} f(z)\frac{8cm^3(z)}{(1+\frac{2}{3}cm(z))[-9(1+\frac{2}{3}cm(z))^2+4cm^2(z)]}\mathrm{d}z
\end{aligned}
\tag{2.7}
$$

and covariance functions

$$
\begin{aligned}
\mathrm{Cov}(X_{f_i}, X_{f_j}) = &-\frac{1}{2\pi^2}\oint\!\!\oint f_i(z_i)f_j(z_j)\frac{m'(z_i)m'(z_j)}{(m(z_i)-m(z_j))^2}\mathrm{d}z_i\,\mathrm{d}z_j \\
&+\frac{1}{\pi^2}\oint\!\!\oint f_i(z_i)f_j(z_j)\frac{2cm'(z_i)m'(z_j)}{9(1+\frac{2}{3}cm(z_i))^2(1+\frac{2}{3}cm(z_j))^2}\mathrm{d}z_i\,\mathrm{d}z_j,
\end{aligned}
\tag{2.8}
$$

where $1 \le i \ne j \le k$. *The contours in* (2.7) *and* (2.8) (*two in* (2.8) *are assumed to be nonoverlapping*) *are closed and taken in the positive direction in the complex plane, each enclosing the support of* $dF_c(x)$.

REMARK 2.1. In Theorem 2.1, independence and continuity are the only two assumptions imposed on the components of $\mathbf{x}$. There are no moment constraints and, furthermore, the p components of $\mathbf{x}$ are not necessarily to be identically distributed.

REMARK 2.2. Consider a special case of the ICM such that $\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\Sigma}_p^{1/2}\mathbf{z}$, where $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma}_p = \mathbf{I}_p$. The components of such $\mathbf{x}$ are thus uncorrelated or independent if the underlying distribution is Gaussian. Results in Bai and Silverstein (2004) demonstrate the CLT of LSS for SCM generated by such $\mathbf{x}$, and the corresponding limiting mean and variance are given by

$$
\mathbb{E}X_f = -\frac{1}{2\pi i}\oint_{\gamma} f(z)\frac{c\underline{m}^3(z)}{(1+\underline{m}(z))[(1+\underline{m}(z))^2-c\underline{m}^2(z)]}\mathrm{d}z,
\tag{2.9}
$$

$$
\mathrm{Cov}(X_{f_i}, X_{f_j}) = -\frac{1}{2\pi^2}\oint\!\!\oint f_i(z_i)f_j(z_j)\frac{\underline{m}'(z_i)\underline{m}'(z_j)}{(\underline{m}(z_i)-\underline{m}(z_j))^2}\mathrm{d}z_i\,\mathrm{d}z_j,
\tag{2.10}
$$

where $\underline{m}(z)$ is the companion Stieltjes transform of the LSD of SCV. From first-order asymptotic results, we know that the LSD of $\mathbf{K}_n$ (generated by such $\mathbf{x}$) is an affine transformation of the LSD of SCM, which is Marchenko–Pastur law. However, it can be seen from the explicit expressions of the two CLTs ((2.7), (2.8) and (2.9), (2.10)) that such affine transformation can only explain part of the second-order asymptotic fluctuations. The existence of extra terms (e.g., the first two terms in (2.7); the second term in (2.8)) in the CLT of LSS for $\mathbf{K}_n$ should be attributed to higher order dependence structure among the columns of $\boldsymbol{\Theta}$.

For ease of computation, the following corollary provides an alternative form of contour integrations for calculating the limiting mean and covariance, which only depends on a complex number ξ that runs counterclockwise along the unit circle. Its proof is relegated into the Supplementary Material (Li, Wang and Li (2021)) by implementing change of variable.

COROLLARY 2.1. *Under the same assumptions as in Theorem 2.1, the limiting mean and covariance function can be expressed as follows*:

$$\begin{aligned}\mathbb{E}X_f = {} & \lim_{r\downarrow 1}\frac{1}{2\pi i}\oint_{|\xi|=1} f\left(\frac{1}{3}+\frac{2}{3}|1+\sqrt{c}\xi|^2\right)\\ & \cdot\left(\frac{1}{(r^2\xi^2-1)\xi}-\frac{2}{\xi^3}+\frac{c}{2(\sqrt{c}+r\xi)^3}\right)\mathrm{d}\xi\\ & + \lim_{r\downarrow 1}\frac{1}{2\pi i}\oint_{|\xi|=1} f\left(\frac{1}{3}+\frac{2}{3}|1+\sqrt{c}\xi|^2\right)\\ & \cdot\left(\frac{3c}{(\sqrt{c}+r\xi)\xi^2}-\frac{c\sqrt{c}}{(\sqrt{c}+r\xi)^2\xi^2}\right)\mathrm{d}\xi\end{aligned}\tag{2.11}$$

and

$$\begin{aligned}\mathrm{Cov}(X_{f_i}, X_{f_j}) = {} & -\lim_{r\downarrow 1}\frac{1}{2\pi^2}\oint\!\!\oint f_i\left(\frac{1}{3}+\frac{2}{3}|1+\sqrt{c}\xi_i|^2\right) f_j\left(\frac{1}{3}+\frac{2}{3}|1+\sqrt{c}\xi_j|^2\right)\\ & \cdot\left(\frac{1}{(\xi_i-r\xi_j)^2}-\frac{1}{\xi_i^2\xi_j^2}\right)\mathrm{d}\xi_i\,\mathrm{d}\xi_j.\end{aligned}\tag{2.12}$$

Here, ξ (*also the two* ξ_i, ξ_j) *moves along the unit circle* $|\xi|=1$ *on the complex plane in the positive direction.*

REMARK 2.3. Here, we require $r>1$ to ensure that after the change of variable from z to $r\xi$, when z runs counterclockwise along γ in (2.7) and (2.8), ξ runs in the same direction along the unit circle in (2.11) and (2.12). Moreover, while in Theorem 2.1, the contour γ in (2.7) and (2.8) actually can be arbitrary chosen only to enclose all the poles, the integral will remain the same. Therefore, any value of $r>1$ will lead to the same result eventually; indeed, it can be checked that after the change of variable from z to $r\xi$, the contour integral is independent of the exact value of r as long as it is larger than one. Here, in Corollary 2.1, in order to ease the computation, we simply take the value of r to stay very close to 1.

REMARK 2.4. Note that the diagonal elements of $\mathbf{K}_n$ are all one, which leads to the fact that $\mathrm{Tr}(\mathbf{K}_n)=p$ is deterministic. Actually, by taking $f(x)=x$ and via some nontrivial calculations, we can show that, the centering term $\int x\,\mathrm{d}F^{c_n}(x)=p$, and the limiting mean $\mathbb{E}X_f$ and variance $\mathrm{Var}(X_f)$ are both zero using Corollary 2.1, which further proves the validity of Theorem 2.1 (also Corollary 2.1) in this degenerate case.

3. Application to test of independence.

Let $\mathbf{x}=(x_1,x_2,\ldots,x_p)^\mathsf{T}$ be a p-dimensional random vector. Of interest is to test

$$H_0: x_1, x_2, \ldots, x_p \text{ are independent}\tag{3.1}$$

based on n samples under the regime $(p,n)\to\infty$, $p/n\to c\in(0,\infty)$.

3.1. *Test statistics and their limiting null distributions.*

Here, we consider the following three test statistics based on the Kendall's rank correlation matrix $\mathbf{K}_n$,

$$Q_{\tau,2}=\mathrm{Tr}(\mathbf{K}_n^2),\qquad Q_{\tau,4}=\mathrm{Tr}(\mathbf{K}_n^4),\qquad Q_{\tau,\log}=\log|\mathbf{K}_n|.$$

The first two test statistics $Q_{\tau,2}$ and $Q_{\tau,4}$ are polynomial functions (of order two and order four, resp.) of the pairwise rank correlations among all the p components and naturally we reject H_0 when their values are too large. The third one $Q_{\tau,\log}$ is motivated by the entropy loss between the Kendall's rank correlation matrix $\mathbf{K}_n$ and its population version under H_0. To be more specific, when all the components of $\mathbf{x}$ are independent, it is straightforward to verify that $\mathbb{E}\mathbf{K}_n = \mathbf{I}_p$. So the entropy loss between $\mathbf{K}_n$ and $\mathbb{E}\mathbf{K}_n$, which is defined as $L(\mathbf{K}_n, \mathbb{E}\mathbf{K}_n) = \operatorname{Tr}\mathbf{K}_n(\mathbb{E}\mathbf{K}_n)^{-1} - \log(|\mathbf{K}_n(\mathbb{E}\mathbf{K}_n)^{-1}|) - p$ (James and Stein (1961), Muirhead (2009), Zheng et al. (2019)) reduces to $-\log|\mathbf{K}_n|$ under H_0, and we reject H_0 when the entropy loss is too large, or equivalently when $Q_{\tau,\log}$ is too small.

On the other hand, all the three statistics $Q_{\tau,k}(k=2,4)$ and $Q_{\tau,\log}$ can be directly linked to particular forms of LSS of $\mathbf{K}_n$ by taking $f(x) = x^k$, $k = 2, 4$ and $f(x) = \log(x)$, respectively, that is,

$$Q_{\tau,k} = p \int x^k \, \mathrm{d}F^{\mathbf{K}_n}(x), \quad k = 2, 4 \quad \text{and} \quad Q_{\tau,\log} = p \int \log(x) \, \mathrm{d}F^{\mathbf{K}_n}(x),$$

with their asymptotic fluctuation behaviors under H_0 fully characterized by implementing the CLT for general LSS of $\mathbf{K}_n$ in Theorem 2.1 or Corollary 2.1. Through some nontrivial calculations, the limiting distributions for the three statistics under H_0 are given in following theorem.

THEOREM 3.1. *Assuming that conditions in Theorem 2.1 hold, under H_0, we have, as $(p,n) \to \infty$, $p/n \to c \in (0,\infty)$,*

$$\begin{aligned}
&Q_{\tau,2} - p - \frac{4p^2}{9n} \xrightarrow{d} \mathcal{N}\left(\frac{14}{9}c^2 - \frac{4}{9}c, \frac{64}{81}c^2\right), \\
&Q_{\tau,4} - p - \frac{8p^2}{3n} - \frac{128p^3}{n^2} - \frac{16p^4}{81n^3} \xrightarrow{d} \mathcal{N}(\mu_{\tau,4}, \sigma^2_{\tau,4}), \\
&Q_{\tau,\log} + \frac{b}{a}\sqrt{pn} - (p+n)\log a + (n-p)\log(a - b\sqrt{p/n}) \\
&\quad \xrightarrow{d} \mathcal{N}(\mu_{\tau,\log}, \sigma^2_{\tau,\log}), \quad \textit{where} \\
&\mu_{\tau,4} = -\frac{8}{3}c + \frac{140}{27}c^2 + \frac{608}{81}c^3 + \frac{112}{81}c^4, \\
&\sigma^2_{\tau,4} = 4\left(\frac{8}{3}c + \frac{352}{81}c^2 + \frac{32}{27}c^3\right)^2 + 6\left(\frac{32}{27}c^{3/2} + \frac{64}{81}c^{5/2}\right)^2 + \frac{2048}{81^2}c^4, \qquad (3.2) \\
&a = \frac{\sqrt{d_{c,+}} + \sqrt{d_{c,-}}}{2}, \qquad b = \frac{\sqrt{d_{c,+}} - \sqrt{d_{c,-}}}{2}, \\
&\mu_{\tau,\log} = -2\log a + \frac{1}{2}\log(a^2 - b^2) + \log(a - b\sqrt{c}) \\
&\qquad + \frac{2b}{a}\sqrt{c} - \frac{4ab\sqrt{c} - 3cb^2}{4(a - b\sqrt{c})^2} + \frac{b^2}{a^2} \\
&\qquad + \left\{\frac{3b^2 - 2ab\sqrt{c}}{4(a\sqrt{c} - b)^2} - \log\left(a - \frac{b}{\sqrt{c}}\right)\right\}\mathbb{1}_{\{c>1\}}, \\
&\sigma^2_{\tau,\log} = 2\log\frac{a^2}{a^2 - b^2} - \frac{2b^2}{a^2}.
\end{aligned}$$

The proof of this theorem is postponed to the Supplementary Material (Li, Wang and Li (2021)).

REMARK 3.1. Leung and Drton (2018) introduced three types of test statistics that are constructed as sums or sums of squares of pairwise rank correlations, including Kendall's τ as a special case. In fact, Theorem 4.1 in Leung and Drton (2018) shows that under H_0, when $p, n \to \infty$,

$$\frac{n[\frac{1}{2}\operatorname{Tr}(\mathbf{K}_n^2) - \frac{p}{2} - \binom{p}{2}\mu_h]}{4p/9} \xrightarrow{d} \mathcal{N}(0,1), \tag{3.3}$$

where $\mu_h = \frac{2(2n+5)}{9n(n-1)} = \frac{4}{9n} + O(n^{-2})$. Under the Marchenko–Pastur asymptotic regime $(p,n) \to \infty$, $p/n \to c \in (0,\infty)$, (3.3) is consistent with our results for the limiting null distribution of $Q_{\tau,2}$ in Theorem 3.1.

3.2. *Simulation experiments.* In this section, we conduct numerical comparisons to examine the finite sample performance of the three proposed test statistics $Q_{\tau,2}$, $Q_{\tau,4}$ and $Q_{\tau,\log}$ with some existing ones. Let Z_α be the upper-α quantile of the standard normal distribution at level α. Based on Theorem 3.1, we obtain three procedures for testing the null hypothesis in (3.1) as follows:

$$\text{Reject } H_0 \quad \text{if} \left\{Q_{\tau,2} - p - \frac{4p^2}{9n} > \frac{8p}{9n}Z_\alpha + \frac{14p^2}{9n^2} - \frac{4p}{9n}\right\} \quad \text{or}$$
$$\left\{Q_{\tau,4} - p - \frac{8p^2}{3n} - \frac{128p^3}{n^2} - \frac{16p^4}{81n^3} > \widehat{\sigma}_{\tau,4}Z_\alpha + \widehat{\mu}_{\tau,4}\right\} \quad \text{or}$$
$$\left\{Q_{\tau,\log} + \frac{b}{a}\sqrt{pn} - (p+n)\log a + (n-p)\log(a - b\sqrt{p/n}) < \widehat{\sigma}_{\tau,\log}Z_{1-\alpha} + \widehat{\mu}_{\tau,\log}\right\},$$

where $\widehat{\mu}_{\tau,4}$, $\widehat{\sigma}^2_{\tau,4}$, $\widehat{\mu}_{\tau,\log}$, $\widehat{\sigma}^2_{\tau,\log}$ are the ones by replacing the limiting value c in the terms $\mu_{\tau,4}$, $\sigma^2_{\tau,4}$, $\mu_{\tau,\log}$, $\sigma^2_{\tau,\log}$ in (3.2) with its finite sample counterpart $c_n = p/n$.

As for comparison, Bao (2019a) proposed a test statistic $Q_{\tau,1}$, which is based on the largest eigenvalue $\lambda_1(\mathbf{K}_n)$ of $\mathbf{K}_n$. They have shown that, under similar assumptions as in Theorem 2.1,

$$Q_{\tau,1} := \frac{3}{2}n^{\frac{2}{3}}c_n^{\frac{1}{6}}d_{+,c_n}^{-\frac{2}{3}}(\lambda_1(\mathbf{K}_n) - \lambda_{+,c_n}) \xrightarrow{d} \mathrm{TW}_1,$$

where $d_{+,c_n} = (1+\sqrt{c_n})^2$, $\lambda_{+,c_n} = \frac{1}{3} + \frac{2}{3}d_{+,c_n}$ and TW_1 stands for the Tracy–Widom law of type I. In addition to Kendall's rank correlation matrix model, there are some other testing procedures based on Spearman and Pearson-type correlation matrices, denoted by $\mathbf{S}_n$ and $\mathbf{R}_n$, respectively. Here, both $\mathbf{S}_n = (s_{k\ell})$ and $\mathbf{R}_n = (\rho_{k\ell})$ are $p \times p$ matrices where $s_{k\ell}$ and $\rho_{k\ell}$ are the Spearman and Pearson correlation of the kth and ℓth row of $\mathbf{X}_n$ with

$$s_{k\ell} = \frac{\sum_{i=1}^n (r_{ki} - \overline{r}_k)(r_{\ell i} - \overline{r}_\ell)}{\sqrt{\sum_{i=1}^n (r_{ki} - \overline{r}_k)^2}\sqrt{\sum_{i=1}^n (r_{\ell i} - \overline{r}_\ell)^2}}, \qquad \overline{r}_k = \frac{1}{n}\sum_{i=1}^n r_{ki} = \frac{n+1}{2},$$

$$\rho_{k\ell} = \frac{\sum_{i=1}^n (x_{ki} - \overline{x}_k)(x_{\ell i} - \overline{x}_\ell)}{\sqrt{\sum_{i=1}^n (x_{ki} - \overline{x}_k)^2}\sqrt{\sum_{i=1}^n (x_{\ell i} - \overline{x}_\ell)^2}}, \qquad \overline{x}_k = \frac{1}{n}\sum_{i=1}^n x_{ki},$$

where r_{ki} is the rank of x_{ki} among $(x_{k1}, \dots, x_{kn})$. Test statistics based on $\mathbf{S}_n$ and $\mathbf{R}_n$ include:

1. $Q_{R,1} = \frac{n\lambda_1(\mathbf{R}_n) - (p^{1/2}+n^{1/2})^2}{(p^{1/2}+n^{1/2})(p^{-1/2}+n^{-1/2})^{1/3}}$ (Bao, Pan and Zhou (2012), Pillai and Yin (2012));
2. $Q_{R,2} = \operatorname{Tr}(\mathbf{R}_n\mathbf{R}_n^\top) - p - \frac{p^2}{n}$ (Gao et al. (2017));
3. $Q_{R,\max} = n(\max_{1\le i<j\le p}|\mathbf{R}_{ij}|)^2 - 4\log n + \log\log n$ (Jiang (2004));

4. $Q_{S,1} = n^{\frac{2}{3}} c_n^{\frac{1}{6}} d_{+,c_n}^{-\frac{2}{3}} (\lambda_1(\mathbf{S}_n) - d_{+,c_n})$ (Bao (2019b));
5. $Q_{S,2} = \frac{n^2}{p^2} \operatorname{Tr}(\mathbf{S}_n \mathbf{S}_n^{\mathsf{T}}) - \frac{n^2}{n-1} - \frac{n^2}{p} + \frac{n}{p}$ (Bao et al. (2015));
6. $Q_{S,4} = \frac{n^4}{p^4} \operatorname{Tr}(\mathbf{S}_n^4) - \frac{n^4}{(n-1)^3} - \frac{n^4}{p^3} - \frac{6n^4}{(n-1)p^2} - \frac{6n^4}{p(n-1)^2}$ (Bao et al. (2015));
7. $Q_{S,\max} = n(\max_{1 \le i < j \le p} |\mathbf{S}_{ij}|)^2 - 4\log p + \log\log p$ (Zhou (2007)).

To evaluate the finite sample performance of these test statistics, data are generated from various model scenarios for different (p, n) combinations. To examine Type I error rate, three models in the following are used with different distributions for $\mathbf{X}_n = (x_{ij})_{p \times n}$:

(I) *Mixed case*: $x_{ij} \sim \text{Gamma}(\text{shape} = 4, \text{scale} = 0.5)$ i.i.d. for $1 \le i \le [p/2]$, $1 \le j \le n$, $x_{ij} \sim t(5)$ i.i.d. for $[p/2] < i \le p$, $1 \le j \le n$;
(II) *Mixed case*: $x_{ij} \sim \text{Cauchy}(\text{location} = 0, \text{scale} = 1)$ i.i.d. for $1 \le i \le [p/2]$, $1 \le j \le n$, $x_{ij} \sim t(5)$ i.i.d. for $[p/2] < i \le p$, $1 \le j \le n$;
(III) *Heavy-tail case*: $x_{ij} \sim \text{Cauchy}(\text{location} = 0, \text{scale} = 1)$ i.i.d. for $1 \le i \le p$, $1 \le j \le n$.

To examine their empirical power, both linear and nonlinear alternatives are considered. A matrix $\mathbf{Z}_n$ with independent components is generated first following (I)$\sim$(III), then the data matrix $\mathbf{X}_n$ is constructed as follows:

(IV) *Toeplitz*: $\mathbf{X}_n = \mathbf{A}\mathbf{Z}_n$, $\mathbf{A} = (a_{ij})_{p \times p}$, $a_{ii} = 1$, for some $1 < k_0 < p$, $\forall 1 \le i \le p$, $1 \le k \le k_0$, $a_{i,i\pm k} = \rho_s^k$, $0 < \rho_s < 1$; $\forall 1 \le i \le p$, $k_0 < k < p$, $a_{i,i\pm k} = 0$;
(V) *Nonlinear correlation*: $x_{ij} = r_1 z_{ij} + r_2 z_{i+1,j}^2 + r_3 z_{i+2,j}^2 + r_4 e_{ij}$, $1 \le i \le p$, $1 \le j \le n$ where $e_{ij} \sim N(0, 1)$ i.i.d.

Here, we set $k_0 = [p/100]$, while for each distribution, ρ_s and $r_1 \sim r_4$ are set differently to accommodate different degrees of dependence. All empirical statistics are obtained using 1000 independent replicates. Note that for Model (II) and (III), the elements of $\mathbf{x}$ do not have finite fourth order moments, which fails to meet the requirement for Pearson-type test statistics, therefore, we eliminate all the corresponding results.

Table 1 shows empirical sizes of all the test statistics under different distribution models (I)$\sim$(III). It can be seen that for the nominal level $\alpha = 5\%$, all the Frobenius-norm-type test statistics ($Q_{\tau,2}$, $Q_{\tau,4}$, $Q_{S,4}$, $Q_{S,2}$, $Q_{R,2}$) have very accurate sizes close to 5%, while spectral-norm-type ($Q_{\tau,1}$, $Q_{S,1}$, $Q_{S,1}$) and maximum-norm-type ($Q_{S,\max}$, $Q_{R,\max}$) test statistics are a little bit undersized. Although such bias shrinks when the sample size becomes larger, it still exists even for very large n due to the slow convergence of extreme eigenvalues to the Tracy–Widom distribution. Moreover, $Q_{R,1}$, $Q_{R,2}$ and $Q_{R,\max}$ work only for model (I) due to moment restrictions. It's observed that our test statistic $Q_{\tau,\log}$ is a little oversized in the $p > n$ cases when p, n are relatively small. However, such bias significantly reduces when p, n increase.

As for the linear correlated alternatives, Table 2 presents the empirical powers of all test statistics under the Toeplitz population matrix Model (IV) for various distributions. It can be seen that our test statistics $Q_{\tau,\log}$ and $Q_{\tau,2}$ perform best under Model (IV). Moreover, among all the three Frobenius-norm-type statistics, our test statistic $Q_{\tau,2}$ demonstrates superiority over the other two across all scenarios and (p, n) combinations. $Q_{S,2}$ always has inferior power and $Q_{R,2}$ does not work for heavy-tailed distributions. As for the nonlinear correlated structure for Model (V), empirical powers of all the test statistics are shown in Table 3. It is obvious that our proposed test statistics $Q_{\tau,2}$, $Q_{\tau,4}$ and $Q_{\tau,\log}$ all demonstrate significant superiority over all the others under heavy-tailed distributions.

TABLE 1
Empirical sizes for all test statistics under different distribution Models (I)∼(III)

p	n	c	$Q_{\tau,\log}$	$Q_{\tau,4}$	$Q_{\tau,2}$	$Q_{S,4}$	$Q_{S,2}$	$Q_{R,2}$	$Q_{\tau,1}$	$Q_{S,1}$	$Q_{R,1}$	$Q_{S,\max}$	$Q_{R,\max}$
100	200	0.5	0.058	0.042	0.056	0.038	0.054	0.036	0.020	0.014	0.016	0.027	0.066
200	400	0.5	0.053	0.051	0.042	0.050	0.042	0.048	0.023	0.018	0.022	0.032	0.062
300	600	0.5	0.065	0.070	0.048	0.067	0.050	0.053	0.030	0.023	0.027	0.030	0.090
100	100	1	0.071	0.054	0.057	0.053	0.053	0.054	0.026	0.017	0.013	0.026	0.052
300	300	1	0.051	0.047	0.039	0.048	0.036	0.046	0.027	0.016	0.028	0.025	0.076
500	500	1	0.047	0.056	0.051	0.058	0.052	0.059	0.027	0.019	0.035	0.046	0.122
200	100	2	0.077	0.068	0.054	0.064	0.049	0.057	0.033	0.017	0.018	0.017	0.050
400	200	2	0.054	0.045	0.049	0.036	0.042	0.044	0.038	0.021	0.025	0.022	0.135
600	300	2	0.048	0.042	0.038	0.038	0.039	0.045	0.037	0.023	0.030	0.026	0.157
For Distribution Model (I)													
100	200	0.5	0.049	0.051	0.057	0.052	0.059	–	0.020	0.009	–	0.035	–
200	400	0.5	0.058	0.048	0.050	0.045	0.050	–	0.019	0.016	–	0.039	–
300	600	0.5	0.047	0.046	0.047	0.043	0.046	–	0.027	0.023	–	0.042	–
100	100	1	0.069	0.060	0.060	0.050	0.053	–	0.020	0.016	–	0.025	–
300	300	1	0.064	0.053	0.049	0.048	0.051	–	0.025	0.018	–	0.031	–
500	500	1	0.045	0.055	0.053	0.054	0.052	–	0.032	0.028	–	0.039	–
200	100	2	0.081	0.057	0.056	0.052	0.052	–	0.036	0.022	–	0.011	–
400	200	2	0.047	0.061	0.052	0.055	0.047	–	0.045	0.028	–	0.039	–
600	300	2	0.043	0.046	0.047	0.042	0.040	–	0.037	0.023	–	0.030	–
For Distribution Model (II)													
100	200	0.5	0.057	0.055	0.044	0.052	0.040	–	0.022	0.017	–	0.036	–
200	400	0.5	0.047	0.048	0.048	0.048	0.047	–	0.022	0.017	–	0.039	–
300	600	0.5	0.055	0.050	0.057	0.049	0.054	–	0.023	0.017	–	0.050	–
100	100	1	0.065	0.054	0.055	0.048	0.050	–	0.019	0.013	–	0.018	–
300	300	1	0.059	0.056	0.040	0.056	0.039	–	0.035	0.024	–	0.038	–
500	500	1	0.062	0.051	0.056	0.048	0.053	–	0.026	0.021	–	0.039	–
200	100	2	0.076	0.059	0.056	0.052	0.047	–	0.052	0.037	–	0.027	–
400	200	2	0.052	0.053	0.050	0.053	0.048	–	0.040	0.023	–	0.020	–
600	300	2	0.044	0.066	0.061	0.063	0.059	–	0.044	0.033	–	0.050	–
For Distribution Model (III)													

4. Proof of Theorem 2.1. Generally speaking, the proof of our main result follows similar routine as establishing the CLT for LSS of a large dimensional sample covariance matrix given in Bai and Silverstein (2004). However, the model structure we considered here is much more complicated than the well-studied SCM model as has already been discussed in the Introduction. Extra new techniques are needed to overcome such difficulties.

(1) Similar to Bai and Silverstein (2004), concrete results for limiting mean and variance of LSS for $\mathbf{K}_n$ are obtainable due to the explicit expression of the expectation of certain quadratic form like $\mathbb{E}\mathbf{v}_k^\top\mathbf{A}\mathbf{v}_k\mathbf{v}_k^\top\mathbf{B}\mathbf{v}_k$. Inequalities related to higher order moment of $\mathbf{v}_k\mathbf{A}\mathbf{v}_k^\top$ and its centralized version are also needed to filter out the leading order terms. The challenge originates from the special dependence structure of $\mathbf{v}_k$. Therefore, techniques like Hoeffding's decomposition has been adopted to reestablish corresponding results described in Lemmas S1.3, S1.4 and S1.5. Subsequently, we have several extra terms contributed by the expectation of the quadratic forms in Lemma S1.5 regarding the so-called off-diagonal terms, especially the derivations of the limits of (4.19)∼(4.23) and (4.32)∼(4.33).

(2) Notice that our model is structured as $\mathbf{K}_n = \mathbf{\Theta}\mathbf{\Theta}^\top$ where $\mathbf{\Theta}$ is a $p \times M$ ($M = n(n-1)/2$) matrix with $p/M \to 0$ while in the commonly studied ICM structure, the data matrix

TABLE 2
Empirical power for all test statistics under Toeplitz population matrix Model (IV) *with different distributions*

p	n	c	$Q_{\tau,\log}$	$Q_{\tau,4}$	$Q_{\tau,2}$	$Q_{S,4}$	$Q_{S,2}$	$Q_{R,2}$	$Q_{\tau,1}$	$Q_{S,1}$	$Q_{R,1}$	$Q_{S,\max}$	$Q_{R,\max}$
100	200	0.5	0.871	0.863	0.898	0.844	0.894	0.880	0.190	0.163	0.168	0.246	0.290
200	400	0.5	1	1	1	1	1	1	0.494	0.452	0.405	0.885	0.850
300	600	0.5	1	1	1	1	1	1	0.768	0.732	0.652	1	1
100	100	1	0.424	0.410	0.405	0.373	0.391	0.368	0.108	0.066	0.069	0.058	0.075
300	300	1	0.991	0.995	0.994	0.995	0.992	0.996	0.332	0.283	0.265	0.488	0.511
500	500	1	1	1	1	1	1	1	0.584	0.534	0.504	0.987	0.981
200	100	2	0.419	0.401	0.470	0.366	0.437	0.419	0.112	0.055	0.056	0.038	0.097
400	200	2	0.865	0.915	0.908	0.892	0.909	0.870	0.245	0.182	0.155	0.178	0.237
600	300	2	0.991	0.993	0.996	0.991	0.995	0.994	0.369	0.287	0.241	0.437	0.484
$\rho_s = 0.06$ *for Distribution Model* (I)													
100	200	0.5	0.889	0.892	0.899	0.868	0.882	–	0.225	0.187	–	0.466	–
200	400	0.5	1	1	1	1	1	–	0.583	0.528	–	0.999	–
300	600	0.5	1	1	1	1	1	–	0.838	0.804	–	1	–
100	100	1	0.464	0.447	0.473	0.405	0.443	–	0.095	0.072	–	0.086	–
300	300	1	1	0.998	0.999	0.997	0.996	–	0.391	0.336	–	0.936	–
500	500	1	1	1	1	1	1	–	0.683	0.631	–	1	–
200	100	2	0.508	0.450	0.494	0.405	0.448	–	0.137	0.083	–	0.099	–
400	200	2	0.901	0.922	0.925	0.900	0.912	–	0.239	0.165	–	0.457	–
600	300	2	0.995	0.999	0.999	0.998	0.999	–	0.399	0.305	–	0.924	–
$\rho_s = 0.03$ *for Distribution Model* (II)													
100	200	0.5	0.843	0.818	0.833	0.786	0.806	–	0.181	0.154	–	0.233	–
200	400	0.5	1	1	1	1	1	–	0.461	0.427	–	0.823	–
300	600	0.5	1	1	1	1	1	–	0.743	0.708	–	0.999	–
100	100	1	0.410	0.397	0.390	0.358	0.353	–	0.071	0.049	–	0.053	–
300	300	1	0.985	0.993	0.993	0.990	0.992	–	0.337	0.274	–	0.508	–
500	500	1	1	1	1	1	1	–	0.592	0.524	–	0.978	–
200	100	2	0.501	0.460	0.474	0.411	0.425	–	0.110	0.075	–	0.034	–
400	200	2	0.873	0.874	0.899	0.848	0.882	–	0.204	0.154	–	0.146	–
600	300	2	0.984	0.990	0.994	0.988	0.992	–	0.340	0.277	–	0.421	–
$\rho_s = 0.02$ *for Distribution Model* (III)													

is $p \times n$ $(p/n = O(1))$, this difference gives an extra term in our limiting mean term. More specifically, when dealing with the difference between $\operatorname{Tr} \mathbb{E} \mathbf{D}^{-1}(z)$ and $\operatorname{Tr}(-\widehat{\mathbf{T}}_N(z))^{-1}$ in (4.25), one extra smaller order term (denoted by $D_n(z)$ in (4.26)) should be taken extra care of. As far as we know, such phenomenon has never occurred in previous works where the appearance of such term is irrelevant with the results on quadratic forms in Lemmas S1.3, S1.4 and S1.5.

(3) Since the main proof relies on expressing the LSS as a contour integral, bounds for both the smallest and largest eigenvalues of $\mathbf{K}_n$ are needed to ensure that its entire spectrum is incorporated in the contour with high probability. However, we cannot directly apply the spectral norm bound for SCM established in Bai and Silverstein (1998) because the nonlinear dependence structure in $\mathbf{K}_n$ is incompatible with that in SCM model. Thus ad hoc bounds for both the smallest and largest eigenvalues of $\mathbf{K}_n$ have been rebuilt in Proposition 4.1.

As for the main proof, let us first define a rectangular contour $\mathcal{C}$ that encloses the interval $[d_{c_n,-}, d_{c_n,+}]$, which is the support of the LSDs $\{F_{c_n}\}$. If $c_n > 1$, let x_l be a positive number smaller than $\frac{1}{3}$, otherwise choose two numbers $x_l < x_r$ such that $[d_{c_n,-}, d_{c_n,+}] \subset (x_l, x_r)$ and letting $v_0 > 0$ be arbitrary, then the contour can be described as $\mathcal{C} = \{x \pm i v_0 : x \in [x_l, x_r]\} \cup$

TABLE 3
Empirical power for all feasible test statistics under nonlinear correlation Model (V) *with heavy-tailed distributions* (II) *and* (III)

p	n	c	$Q_{\tau,\log}$	$Q_{\tau,4}$	$Q_{\tau,2}$	$Q_{S,4}$	$Q_{S,2}$	$Q_{\tau,1}$	$Q_{S,1}$	$Q_{S,\max}$
100	200	0.5	0.949	0.931	0.947	0.899	0.929	0.235	0.179	0.765
200	400	0.5	1	1	1	1	1	0.551	0.491	1
300	600	0.5	1	1	1	1	1	0.804	0.750	1
100	100	1	0.573	0.495	0.561	0.433	0.478	0.091	0.065	0.180
300	300	1	1	0.994	0.999	0.992	0.996	0.374	0.313	0.997
500	500	1	1	1	1	1	1	0.654	0.584	1
200	100	2	0.641	0.485	0.545	0.424	0.456	0.123	0.071	0.132
400	200	2	0.962	0.934	0.963	0.896	0.934	0.268	0.187	0.725
600	300	2	0.998	0.999	0.999	0.999	0.999	0.350	0.263	0.998
$r_1 = 0.01, r_2 = 0.02, r_3 = 0.006, r_4 = 0.5$ *for Distribution Model* (II)										
100	200	0.5	0.699	0.667	0.708	0.620	0.670	0.119	0.092	0.159
200	400	0.5	0.994	0.992	0.995	0.987	0.992	0.260	0.219	0.580
300	600	0.5	1	1	1	1	1	0.482	0.437	0.953
100	100	1	0.383	0.318	0.349	0.271	0.310	0.066	0.034	0.041
300	300	1	0.935	0.926	0.946	0.910	0.932	0.192	0.146	0.279
500	500	1	1	1	1	1	1	0.335	0.286	0.801
200	100	2	0.407	0.310	0.342	0.266	0.284	0.095	0.062	0.031
400	200	2	0.730	0.665	0.717	0.611	0.656	0.130	0.083	0.097
600	300	2	0.944	0.930	0.954	0.908	0.940	0.210	0.161	0.219
$r_1 = 0.002, r_2 = 0.005, r_3 = 0.0015, r_4 = 0.5$ *for Distribution Model* (III)										

$\{x + iv : x \in \{x_l, x_r\}, v \in [-v_0, v_0]\}$. Then we can define a two-dimensional random process

$$M_n(z) = p\big(m_{F^{\mathbf{K}_n}}(z) - m_{F^{c_n}}(z)\big)$$

on $\mathcal{C}$ of the complex plane. From Cauchy's integral formula, for any analytic function f, we have when z moves along the contour $\mathcal{C}$ in the positive direction, $\int f(x)dG_n(x) = -\oint_{\mathcal{C}} f(z)M_n(z)\,\mathrm{d}z$ when all sample eigenvalues of $\mathbf{K}_n$ fall in the interval (x_l, x_r), which holds asymptotically with probability one. However, it could happen that in finite dimensional situations, some extreme eigenvalues of $\mathbf{K}_n$ may locate outside the interval (x_l, x_r), though such event would happen with very small probability. To deal with this, Bai and Silverstein (2004) suggested truncating $M_n(z)$ as, for $z = x + \mathrm{i}v \in \mathcal{C}$,

$$\widehat{M}_n(z) = \begin{cases} M_n(z), & z \in \mathcal{C}_n, \\ M_n(x + \mathrm{i}n^{-1}\varepsilon_n), & v \in [0, n^{-1}\varepsilon_n], x = x_r \text{ or } \{x = x_l \text{ and } c_n \le 1\}, \\ M_n(x - \mathrm{i}n^{-1}\varepsilon_n), & v \in [-n^{-1}\varepsilon_n, 0], x = x_r \text{ or } \{x = x_l \text{ and } c_n \le 1\}, \end{cases} \tag{4.1}$$

where if $c_n \le 1$, $\mathcal{C}_n = \{x \pm \mathrm{i}v_0 : x \in [x_l, x_r]\} \cup \{x \pm \mathrm{i}v : x \in \{x_l, x_r\}, v \in [n^{-1}\varepsilon_n, v_0]\}$, and for $c_n > 1$, $\mathcal{C}_n = \{x \pm \mathrm{i}v_0 : x \in [x_l, x_r]\} \cup \{x_l \pm \mathrm{i}v : v \in [0, v_0]\} \cup \{x_r \pm \mathrm{i}v : v \in [n^{-1}\varepsilon_n, v_0]\}$, and the positive sequence (ε_n) decreases to zero satisfying

$$\{\varepsilon_n > n^{-\alpha} \text{ for some } \alpha \in (0, 1)\}. \tag{4.2}$$

Such construction of $\widehat{M}_n(z)$ ensures that it matches $M_n(z)$ except for the two small segments near the real line, that is, $\{x + iv, x \in \{x_l, x_r\}, v \in [-n^{-1}\varepsilon_n, n^{-1}\varepsilon_n]\}$. Therefore, with probability one, for n large,

$$\left|\oint_{\mathcal{C}} f(z)\big(M_n(z) - \widehat{M}_n(z)\big)\,\mathrm{d}z\right| \le K\varepsilon_n\big(|\lambda_p \wedge d_{c_n,-} - x_\ell|^{-1} + |\lambda_1 \vee d_{c_n,+} - x_r|^{-1}\big),$$

which converges to zero as $n \to \infty$ due to the bound for extreme eigenvalues of $\mathbf{K}_n$ established in the following Proposition 4.1, whose proof is relegated to the Supplementary Material (Li, Wang and Li (2021)).

PROPOSITION 4.1. *Assuming that conditions in Theorem 2.1 hold, for any $\eta_l < d_{c_n,-}$, $\eta_r > d_{c_n,+}$ and any positive ℓ,*

$$\mathbb{P}(\lambda_1 \geq \eta_r) = o(n^{-\ell}), \qquad \mathbb{P}(\lambda_p \leq \eta_l) = o(n^{-\ell}). \tag{4.3}$$

Hence, one may get that $\oint_{\mathcal{C}} f(z) M_n(z)\,\mathrm{d}z = \oint_{\mathcal{C}} f(z)\widehat{M}_n(z)\,\mathrm{d}z + o_p(1)$, and the proof of Theorem 2.1 can be completed by the convergence of $\widehat{M}_n(z)$ on $\mathcal{C}$ as stated in the following lemma.

LEMMA 4.1. *Under the same assumptions as in Theorem 2.1, $\{\widehat{M}_n(\cdot)\}$ forms a tight sequence on $\mathcal{C}$ and converges weakly to a two-dimensional Gaussian process $M(\cdot)$ satisfying, for $z \in \mathcal{C}$,*

$$\begin{aligned} \mathbb{E}M(z) = {}& \frac{36cm_{F_c}^3(z)(1+\frac{2}{3}cm_{F_c}(z))}{[-9(1+\frac{2}{3}cm_{F_c}(z))^2 + 4cm_{F_c}^2(z)]^2} \\ & - \frac{2c^2m_{F_c}^3(z)[(1+\frac{2}{3}cm_{F_c}(z))^2 + 6 + \frac{4}{3}cm_{F_c}(z)]}{-9(1+\frac{2}{3}cm_{F_c}(z))^2 + 4cm_{F_c}^2(z)} \\ & + \frac{8cm_{F_c}^3(z)}{(1+\frac{2}{3}cm_{F_c}(z))[-9(1+\frac{2}{3}cm_{F_c}(z))^2 + 4cm_{F_c}^2(z)]} \end{aligned} \tag{4.4}$$

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

$$\begin{aligned} \operatorname{Cov}\bigl(M(z_1), M(z_2)\bigr) = {}& \frac{2m'_{F_c}(z_1)m'_{F_c}(z_2)}{(m_{F_c}(z_1) - m_{F_c}(z_2))^2} - \frac{2}{(z_1 - z_2)^2} \\ & - \frac{8cm'_{F_c}(z_1)m'_{F_c}(z_2)}{9(1+\frac{2}{3}cm_{F_c}(z_1))^2(1+\frac{2}{3}cm_{F_c}(z_2))^2}. \end{aligned} \tag{4.5}$$

4.1. *Proof of Lemma 4.1.* First, we decompose $\widehat{M}_n(z)$ into the summation of a random part $M_n^1(z)$ and a deterministic part $M_n^2(z)$, where for $z \in \mathcal{C}_n$,

$$M_n^1(z) = p\bigl(m_{F^{\mathbf{K}_n}}(z) - \mathbb{E}m_{F^{\mathbf{K}_n}}(z)\bigr) \quad \text{and} \quad M_n^2(z) = p\bigl(\mathbb{E}m_{F^{\mathbf{K}_n}}(z) - m_{F^{c_n}}(z)\bigr),$$

for $z \in \mathcal{C} \setminus \mathcal{C}_n$, define $M_n^1(z)$ and $M_n^2(z)$ in the similar way as (4.1). The proof of Lemma 4.1 is then complete if we can verify the following three steps:

Step 1: Finite dimensional convergence of $M_n^1(z)$ in distribution to a centered multivariate Gaussian random vector with covariance function given by (4.5);
Step 2: Tightness of $M_n^1(z)$ for $z \in \mathcal{C}$;
Step 3: Convergence of $M_n^2(z)$ to the mean function given by (4.4).

In the following, we will demonstrate Step 1 and Step 3 while the proof of Step 2 is relegated to the Supplementary Material (Li, Wang and Li (2021)).

4.1.1. *Preliminary*: *Hoeffding decomposition.* For the $p \times n$ data matrix $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ with $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})^\mathsf{T}$, $i = 1, \dots, n$, define

$$v_{k,(ij)} = \operatorname{sign}(x_{ki} - x_{kj}), \qquad \boldsymbol{\theta}_{(ij)} = \frac{1}{\sqrt{M}}(v_{1,(ij)}, v_{2,(ij)}, \dots, v_{p,(ij)})^\mathsf{T},$$

$$\boldsymbol{\Theta} = (\boldsymbol{\theta}_{(12)}, \dots, \boldsymbol{\theta}_{(1n)}, \boldsymbol{\theta}_{(23)}, \dots, \boldsymbol{\theta}_{(2n)}, \dots, \boldsymbol{\theta}_{(n-1,n)}),$$

where $M := M(n) = n(n-1)/2$. We can represent $\mathbf{K}_n$ in (2.1) as

$$\mathbf{K}_n = \sum_{1 \le i < j \le n} \boldsymbol{\theta}_{(ij)} \boldsymbol{\theta}_{(ij)}^\mathsf{T} = \boldsymbol{\Theta}\boldsymbol{\Theta}^\mathsf{T}. \tag{4.6}$$

Notice that the kth row of $\boldsymbol{\Theta}$ contains information only related to the kth component of the original data (kth row of $\mathbf{X}_n$), thus rows of $\boldsymbol{\Theta}$ are independent. We denote the kth row of $\boldsymbol{\Theta}$ by

$$\mathbf{v}_k^\mathsf{T} = \frac{1}{\sqrt{M}}(v_{k,(12)}, \dots, v_{k,(1n)}, v_{k,(23)}, \dots, v_{k,(2n)}, \dots, v_{k,(n-1,n)}). \tag{4.7}$$

Further, if we look into the components of $\mathbf{v}_k^\mathsf{T}$ carefully, we will find out that not all of them are independent, for example, $v_{k,(ij)}$ and $v_{k,(i\ell)}$ are correlated when $j \neq \ell$. To deal with such dependence structure within $\mathbf{v}_k^\mathsf{T}$, a variation of Hoeffding decomposition is first introduced in Bandeira, Lodhia and Rigollet (2017) and further refined by Bao (2019a). Specifically, let $v_{k,(i\cdot)} = \mathbb{E}(\operatorname{sign}(x_{ki} - x_{kj}) \mid x_{ki})$, $v_{k,(\cdot j)} = \mathbb{E}(\operatorname{sign}(x_{ki} - x_{kj}) \mid x_{kj})$, then $v_{k,(ij)}$ can be decomposed into three parts $v_{k,(ij)} := v_{k,(i\cdot)} + v_{k,(\cdot j)} + \bar{v}_{k,(ij)}$, where $\{v_{k,(i\cdot)}, k = 1, \dots, p; i = 1, \dots, n\}$ are i.i.d. uniformly distributed on $[-1, 1]$ and the three terms $v_{k,(i\cdot)}$, $v_{k,(\cdot j)}$ and $\bar{v}_{k,(ij)}$ are pairwisely uncorrelated. If we further set $u_{k,(ij)} := v_{k,(i\cdot)} + v_{k,(\cdot j)}$ and define

$$\begin{aligned} \mathbf{u}_k^\mathsf{T} &= \frac{1}{\sqrt{M}}(u_{k,(12)}, \dots, u_{k,(1n)}, u_{k,(23)}, \dots, u_{k,(2n)}, \dots, u_{k,(n-1n)}), \\ \bar{\mathbf{v}}_k^\mathsf{T} &= \frac{1}{\sqrt{M}}(\bar{v}_{k,(12)}, \dots, \bar{v}_{k,(1n)}, \bar{v}_{k,(23)}, \dots, \bar{v}_{k,(2n)}, \dots, \bar{v}_{k,(n-1n)}), \end{aligned} \tag{4.8}$$

then the kth row of $\boldsymbol{\Theta}$ can be expressed as the summation of two terms $\mathbf{v}_k^\mathsf{T} = \mathbf{u}_k^\mathsf{T} + \bar{\mathbf{v}}_k^\mathsf{T}$, with the following covariance structure:

$$\mathbb{E}\mathbf{u}_k\mathbf{u}_k^\mathsf{T} = \frac{1}{3M}\mathbf{T}^\mathsf{T}\mathbf{T}, \qquad \mathbb{E}\bar{\mathbf{v}}_k\bar{\mathbf{v}}_k^\mathsf{T} = \frac{1}{3M}\mathbf{I}_M \quad \text{and} \quad \mathbb{E}\mathbf{v}_k\mathbf{v}_k^\mathsf{T} = \frac{1}{3M}(\mathbf{T}^\mathsf{T}\mathbf{T} + \mathbf{I}_M).$$

Here, $\mathbf{T}$ is a $n \times M$ matrix with entries $\mathbf{T} = (t_{\ell,(ij)})_{\ell,i<j}$, $t_{\ell,(ij)} := \delta_{\ell i} - \delta_{\ell j}$, $1 \le \ell \le n$, $1 \le i < j \le n$. and $(\mathbf{T}\mathbf{T}^\mathsf{T})^2 = n(\mathbf{T}\mathbf{T}^\mathsf{T})$, $(\mathbf{T}^\mathsf{T}\mathbf{T})^2 = n(\mathbf{T}^\mathsf{T}\mathbf{T})$ (see more details in Bao (2019a)).

4.1.2. *Proof of Step* 1: *Finite dimensional convergence of $M_n^1(z)$ in distribution.* In the first step, we will show that, for any positive integer r, the r dimensional random vector $(M_n^1(z_1) \cdots M_n^1(z_r))^\mathsf{T}$ is jointly Gaussian. To this end, we will show that the sum $\sum_{i=1}^r \alpha_i M_n^1(z_i)$, $\Im z_i \neq 0$ forms a tight sequence of random functions for z_i and will converge in distribution to a Gaussian random variable. The main strategy of the proof is based on the martingale CLT given in Lemma S1.2, together with our newly established Lemmas S1.3$\sim$S1.5.

Let $v = \Im z$ and for the following analysis we will assume $v > 0$. Define

$$\boldsymbol{\Theta}_k = (\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_p)^\mathsf{T}, \qquad \mathbf{K}_{n,k} = \boldsymbol{\Theta}_k\boldsymbol{\Theta}_k^\mathsf{T},$$

$$\mathbf{A}_k = \boldsymbol{\Theta}_k^\mathsf{T}(\mathbf{K}_{n,k} - z\mathbf{I}_{p-1})^{-2}\boldsymbol{\Theta}_k, \qquad \mathbf{B}_k = \boldsymbol{\Theta}_k^\mathsf{T}(\mathbf{K}_{n,k} - z\mathbf{I}_{p-1})^{-1}\boldsymbol{\Theta}_k,$$

$$\beta_k = \frac{1}{-\mathbf{v}_k^\top \mathbf{v}_k + z + \mathbf{v}_k^\top \mathbf{B}_k \mathbf{v}_k}, \qquad \widetilde{b}_k = \frac{1}{-1 + z + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}(\mathbf{B}_k)},$$

$$b_n = \frac{1}{-1 + z + \frac{1}{3M}\mathbb{E}\operatorname{Tr}(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top) + \frac{1}{3M}\mathbb{E}\operatorname{Tr}(\mathbf{B}_k)},$$

$$g_k = \mathbf{v}_k^\top \mathbf{v}_k - 1 - \mathbf{v}_k^\top \mathbf{B}_k \mathbf{v}_k + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}\mathbf{B}_k,$$

$$h_k = \mathbf{v}_k^\top \mathbf{A}_k \mathbf{v}_k - \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top) - \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k,$$

$$a_k = -g_k\beta_k\widetilde{b}_k(1 + \mathbf{v}_k^\top \mathbf{A}_k \mathbf{v}_k), \qquad d_k = h_k\widetilde{b}_k.$$

Let $\mathbb{E}_k(\cdot)$ denote the conditional expectation with respect to the σ-field generated by $\mathbf{v}_1, \dots, \mathbf{v}_k$, we have

$$\begin{aligned}
M_n^1(z) &= p\big(m_{F^{\mathbf{K}_n}}(z) - \mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big) \\
&= \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})\big\{\operatorname{Tr}(\mathbf{K}_n - z\mathbf{I}_p)^{-1} - \operatorname{Tr}(\mathbf{K}_{n,k} - z\mathbf{I}_{p-1})^{-1}\big\} \\
&= \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})\frac{1 + \mathbf{v}_k^\top \mathbf{\Theta}_k^\top (\mathbf{K}_{n,k} - z\mathbf{I}_{p-1})^{-2}\mathbf{\Theta}_k \mathbf{v}_k}{\mathbf{v}_k^\top \mathbf{v}_k - z - \mathbf{v}_k^\top \mathbf{\Theta}_k^\top (\mathbf{K}_{n,k} - z\mathbf{I}_{p-1})^{-1}\mathbf{\Theta}_k \mathbf{v}_k} \\
&= \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})\left\{a_k - d_k + \frac{1 + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k}{1 - z - \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top) - \frac{1}{3M}\operatorname{Tr}\mathbf{B}_k}\right\} \\
&= \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})(a_k - d_k) = \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})a_k - \mathbb{E}_k d_k.
\end{aligned} \tag{4.9}$$

By applying the equality,

$$\beta_k = \widetilde{b}_k + \beta_k\widetilde{b}_k g_k,$$

we have

$$\begin{aligned}
a_k &= -g_k\beta_k\widetilde{b}_k(1 + \mathbf{v}_k^\top \mathbf{A}_k \mathbf{v}_k) \\
&= -\widetilde{b}_k^2 g_k\left(1 + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k\right) - h_k g_k \widetilde{b}_k^2 - \beta_k\widetilde{b}_k^2(1 + \mathbf{v}_k^\top \mathbf{A}_k \mathbf{v}_k)g_k^2 \\
&:= a_{k1} + a_{k2} + a_{k3},
\end{aligned}$$

which together with (4.9) implies that

$$M_n^1(z) = \sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})(a_{k1} + a_{k2} + a_{k3}) - \mathbb{E}_k d_k.$$

Next, we will show that the contribution of a_{k2} and a_{k3} to $M_n^1(z)$ can be negligible as $n \to \infty$. For the contribution of a_{k3}, we have

$$\begin{aligned}
\mathbb{E}\left|\sum_{k=1}^p (\mathbb{E}_k - \mathbb{E}_{k-1})a_{k3}\right|^2 &\le C_0 \sum_{k=1}^p \mathbb{E}(1 + \mathbf{v}_k^\top \mathbf{A}_k \mathbf{v}_k)^2 g_k^4 \\
&\le C_0 \sum_{k=1}^p \left\{\mathbb{E}\left[\left(1 + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k\right)^2 g_k^4\right] + \mathbb{E}(h_k^2 g_k^4)\right\}.
\end{aligned}$$

Using the Hoeffding decomposition for $\mathbf{v}_k$ which is denoted as $\mathbf{v}_k = \mathbf{u}_k + \overline{\mathbf{v}}_k$ (see (4.8)) together with Propositions 2.2 and 2.4 in Bao (2019a), we have

$$
\begin{aligned}
\mathbb{E}g_k^4 &= \mathbb{E}\bigg[\mathbf{v}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\mathbf{v}_k + \frac{1}{3M}\operatorname{Tr}\big(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top\big) + \frac{1}{3M}\operatorname{Tr}\mathbf{B}_k - 1\bigg]^4 \\
&= \mathbb{E}\bigg[\mathbf{u}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\mathbf{u}_k - \frac{1}{3M}\operatorname{Tr}\big(\mathbf{T}(\mathbf{I}_M - \mathbf{B}_k)\mathbf{T}^\top\big) + 2\overline{\mathbf{v}}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\mathbf{u}_k \\
&\quad + \overline{\mathbf{v}}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\overline{\mathbf{v}}_k - \frac{1}{3M}\operatorname{Tr}(\mathbf{I}_M - \mathbf{B}_k)\bigg]^4 \\
&\le C_0\bigg\{\mathbb{E}\bigg(\mathbf{u}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\mathbf{u}_k - \frac{1}{3M}\operatorname{Tr}\big(\mathbf{T}(\mathbf{I}_M - \mathbf{B}_k)\mathbf{T}^\top\big)\bigg)^4 + \mathbb{E}\big(\overline{\mathbf{v}}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\mathbf{u}_k\big)^4 \\
&\quad + \mathbb{E}\bigg(\overline{\mathbf{v}}_k^\top(\mathbf{I}_M - \mathbf{B}_k)\overline{\mathbf{v}}_k - \frac{1}{3M}\operatorname{Tr}(\mathbf{I}_M - \mathbf{B}_k)\bigg)^4\bigg\} \\
&\prec \bigg(\frac{\operatorname{Tr}|\mathbf{T}(\mathbf{I}_M - \mathbf{B}_k)\mathbf{T}^\top|^2}{9M^2}\bigg)^2 + \bigg(\frac{n}{M^2}\operatorname{Tr}|\mathbf{I}_M - \mathbf{B}_k|^2\bigg)^2 \\
&\quad + \bigg(\frac{1}{M^2}\sum_{l=1}^{n}\bigg|\sum_{j=l+1}^{n}\big(\mathbf{T}(\mathbf{I}_M - \mathbf{B}_k)\big)_{j,(lj)}\bigg|^2\bigg)^2 \\
&= O\big(n^{-2}\big),
\end{aligned} \tag{4.10}
$$

which implies that

$$
\mathbb{E}\bigg[\bigg(1 + \frac{1}{3M}\operatorname{Tr}\big(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top\big) + \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k\bigg)^2 g_k^4\bigg] = O\big(n^{-2}\big). \tag{4.11}
$$

According to Lemma S1.4,

$$
\mathbb{E}h_k^2 g_k^4 \le \sqrt{\mathbb{E}h_k^4 \cdot \mathbb{E}g_k^8} \le \sqrt{O\big(n^{-2}\big)\cdot O\big(n^{-4}\big)} = O\big(n^{-3}\big). \tag{4.12}
$$

Collecting (4.11) and (4.12), we have

$$
\mathbb{E}\bigg|\sum_{k=1}^{p}(\mathbb{E}_k - \mathbb{E}_{k-1})a_{k3}\bigg|^2 \le C_0\sum_{k=1}^{p} O\bigg(\frac{1}{n^2}\bigg) = O\big(n^{-1}\big).
$$

Similarly,

$$
\begin{aligned}
\mathbb{E}\bigg|\sum_{k=1}^{p}(\mathbb{E}_k - \mathbb{E}_{k-1})a_{k2}\bigg|^2 &= \sum_{k=1}^{p}\mathbb{E}\big|(\mathbb{E}_k - \mathbb{E}_{k-1})a_{k2}\big|^2 \\
&\le \sum_{k=1}^{p}\mathbb{E}\big|h_k g_k \widetilde{b}_k^2\big|^2 \le C_0\sum_{k=1}^{p}\mathbb{E}h_k^2 g_k^2 \le C_0\sum_{k=1}^{p}\sqrt{\mathbb{E}h_k^4\cdot\mathbb{E}g_k^4} \\
&= C_0\sum_{k=1}^{p}\sqrt{O\big(n^{-2}\big)\cdot O\big(n^{-2}\big)} = O\big(n^{-1}\big),
\end{aligned}
$$

which implies that

$$\begin{aligned}
M_n^1(z) &= \sum_{k=1}^{p}(\mathbb{E}_k - \mathbb{E}_{k-1})(a_{k1} + a_{k2} + a_{k3}) - \mathbb{E}_k d_k = \sum_{k=1}^{p} \mathbb{E}_k(a_{k1} - d_k) + o(1) \\
&= \sum_{k=1}^{p} \mathbb{E}_k\left\{-\widetilde{b}_k^2 g_k\left(1 + \frac{1}{3M}\operatorname{Tr}(\mathbf{T}\mathbf{A}_k\mathbf{T}^\top) + \frac{1}{3M}\operatorname{Tr}\mathbf{A}_k\right) - h_k\widetilde{b}_k\right\} + o(1) \\
&= \sum_{k=1}^{p} \mathbb{E}_k\left(\frac{\mathrm{d}}{\mathrm{d}z}\widetilde{b}_k(z)g_k(z)\right) + o(1),
\end{aligned}$$

where $\mathbb{E}_k \frac{\mathrm{d}}{\mathrm{d}z}(\widetilde{b}_k(z)g_k(z))$ is a sequence of martingale difference. Since

$$\sum_{i=1}^{r}\alpha_i M_n^1(z_i) = \sum_{k=1}^{p}\sum_{i=1}^{r}\alpha_i\mathbb{E}_k\frac{\mathrm{d}}{\mathrm{d}z_i}(\widetilde{b}_k(z_i)g_k(z_i)) + o(1) := \sum_{k=1}^{p}\left(\sum_{i=1}^{r}\alpha_i Y_k(z_i)\right) + o(1),$$

by applying the martingale central limit theorem (Billingsley (2008)) as stated in Lemma S1.2, it is enough to verify

$$\sum_{k=1}^{p}\mathbb{E}\left(\left(\sum_{i=1}^{r}\alpha_i Y_k(z_i)\right)^2 \mathbb{1}_{\{|\sum_{i=1}^{r}\alpha_i Y_k(z_i)|\ge\varepsilon\}}\right) \xrightarrow{p} 0, \tag{4.13}$$

and prove that under the assumptions of our Theorem 2.1, for $z_1, z_2 \in \mathbb{C}^+$, the term

$$\sum_{k=1}^{p}\mathbb{E}_{k-1}\big(Y_k(z_1)Y_k(z_2)\big) \tag{4.14}$$

converges in probability to a constant, or to determine the limit of

$$\sum_{k=1}^{p}\mathbb{E}_{k-1}\big[\mathbb{E}_k\big(\widetilde{b}_k(z_1)g_k(z_1)\big)\mathbb{E}_k\big(\widetilde{b}_k(z_2)g_k(z_2)\big)\big]. \tag{4.15}$$

Actually, we have

$$\frac{\mathrm{d}}{\mathrm{d}z_1}\frac{\mathrm{d}}{\mathrm{d}z_2}(4.15) = (4.14).$$

As for the condition (4.13), by taking (4.10) and (4.12) into account, we have

$$\mathbb{E}\left|\mathbb{E}_k\frac{\mathrm{d}}{\mathrm{d}z}\big(\widetilde{b}_k(z)g_k(z)\big)\right|^4 = \mathbb{E}|\mathbb{E}_k a_{k1} - \mathbb{E}_k d_k|^4 \le C_0\big(\mathbb{E}g_k^4 + \mathbb{E}h_k^4\big) = O\big(n^{-2}\big),$$

which implies that for any $\varepsilon > 0$,

$$\begin{aligned}
\sum_{k=1}^{p}\mathbb{E}\left(\left(\sum_{i=1}^{r}\alpha_i Y_k(z_i)\right)^2 \mathbb{1}_{\{|\sum_{i=1}^{r}\alpha_i Y_k(z_i)|\ge\varepsilon\}}\right) &\le \frac{1}{\varepsilon^2}\sum_{k=1}^{p}\mathbb{E}\left|\sum_{i=1}^{r}\alpha_i Y_k(z_i)\right|^4 \\
&\le \frac{pC_0}{\varepsilon^2}\mathbb{E}\left|\mathbb{E}_k\frac{\mathrm{d}}{\mathrm{d}z_i}\big(\widetilde{b}_k(z_i)g_k(z_i)\big)\right|^4 = O\big(n^{-1}\big),
\end{aligned}$$

then condition (4.13) is verified.

Therefore, the remaining is devoted to find the limit of (4.14). Denote $\mathbf{D}_k^{-1}(z) = (\mathbf{\Theta}_k^\top\mathbf{\Theta}_k - z\mathbf{I}_M)^{-1}$, then $\mathbf{B}_k = \mathbf{I}_M + z\mathbf{D}_k^{-1}(z)$, so we have

$$\begin{aligned}
\mathbb{E}\big|\widetilde{b}_k(z) - b_n(z)\big|^2 &= \mathbb{E}\widetilde{b}_k^2(z)b_n^2(z)\left[\frac{1}{3M}\operatorname{Tr}\big(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top + \mathbf{B}_k\big) - \frac{1}{3M}\mathbb{E}\operatorname{Tr}\big(\mathbf{T}\mathbf{B}_k\mathbf{T}^\top + \mathbf{B}_k\big)\right]^2 \\
&= O\big(n^{-2}\big),
\end{aligned}$$

which implies

$$\sum_{k=1}^{p} \mathbb{E}_{k-1}\big[\mathbb{E}_k\big(\widetilde{b}_k(z_1)g_k(z_1)\big)\mathbb{E}_k\big(\widetilde{b}_k(z_2)g_k(z_2)\big)\big] - b_n(z_1)b_n(z_2)\sum_{k=1}^{p} \mathbb{E}_{k-1}\big[\mathbb{E}_k g_k(z_1)\mathbb{E}_k g_k(z_2)\big] \xrightarrow{p} 0.$$

Therefore, it remains to find the limit of

$$\frac{\mathrm{d}}{\mathrm{d}z_1}\frac{\mathrm{d}}{\mathrm{d}z_2}\left\{b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\mathbb{E}_{k-1}\big[\mathbb{E}_k g_k(z_1)\mathbb{E}_k g_k(z_2)\big]\right\}, \tag{4.16}$$

which in turn gives the limit of (4.14).

In fact, denote $\mathbf{H}_k(z) = \mathbf{I}_M - \mathbb{E}_k\mathbf{B}_k(z) = -z\mathbb{E}_k\mathbf{D}_k^{-1}(z) \in \sigma(\mathcal{F}_{k-1})$, we can write

$$\begin{aligned}&\mathbb{E}_{k-1}\big[\mathbb{E}_k g_k(z_1)\mathbb{E}_k g_k(z_2)\big]\\ &\quad= \mathbb{E}\big[\big(\mathbf{v}_k^\top\mathbf{H}_k(z_1)\mathbf{v}_k - \mathbb{E}\mathbf{v}_k^\top\mathbf{H}_k(z_1)\mathbf{v}_k\big)\big(\mathbf{v}_k^\top\mathbf{H}_k(z_2)\mathbf{v}_k - \mathbb{E}\mathbf{v}_k^\top\mathbf{H}_k(z_2)\mathbf{v}_k\big)\big].\end{aligned}$$

By applying Lemma S1.5, we have

$$b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\mathbb{E}_{k-1}\big[\mathbb{E}_k g_k(z_1)\mathbb{E}_k g_k(z_2)\big]$$

$$= b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{2}{9M^2}\operatorname{Tr}\big(\mathbf{T}\mathbf{H}_k(z_1)\mathbf{T}^\top\mathbf{T}\mathbf{H}_k(z_2)\mathbf{T}^\top\big) \tag{4.17}$$

$$- b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{2}{15M^2}\sum_{i=1}^{n}\big(\mathbf{T}\mathbf{H}_k(z_1)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{H}_k(z_2)\mathbf{T}^\top\big)_{ii} \tag{4.18}$$

$$\begin{aligned}&+ b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{4}{45M^2}\bigg[\sum_{i<\ell}\big(\mathbf{T}\mathbf{H}_k(z_1)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{\ell,(i\ell)}\\ &- \sum_{\ell<i}\big(\mathbf{T}\mathbf{H}_k(z_1)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{\ell,(\ell i)}\bigg]\end{aligned} \tag{4.19}$$

$$\begin{aligned}&+ b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{4}{45M^2}\bigg[\sum_{i<\ell}\big(\mathbf{T}\mathbf{H}_k(z_2)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{\ell,(i\ell)}\\ &- \sum_{\ell<i}\big(\mathbf{T}\mathbf{H}_k(z_2)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{\ell,(\ell i)}\bigg]\end{aligned} \tag{4.20}$$

$$\begin{aligned}&+ b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{4}{45M^2}\bigg[\sum_{j<i<t}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{i,(it)}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{j,(jt)}\\ &+ \sum_{t<j<i}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{i,(ti)}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{j,(tj)}\bigg]\end{aligned} \tag{4.21}$$

$$\begin{aligned}&+ b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{4}{45M^2}\bigg[\sum_{s<i<j}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{i,(si)}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{j,(sj)}\\ &+ \sum_{i<j<s}\big(\mathbf{T}\mathbf{H}_k(z_1)\big)_{i,(is)}\big(\mathbf{T}\mathbf{H}_k(z_2)\big)_{j,(js)}\bigg]\end{aligned} \tag{4.22}$$

$$\begin{aligned}
&- b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\frac{4}{45M^2}\bigg[\sum_{j<t<i}(\mathbf{TH}_k(z_1))_{i,(ti)}(\mathbf{TH}_k(z_2))_{j,(jt)} \\
&+ \sum_{i<s<j}(\mathbf{TH}_k(z_1))_{i,(is)}(\mathbf{TH}_k(z_2))_{j,(sj)}\bigg] \\
&+ O(n^{-1}).
\end{aligned} \tag{4.23}$$

Therefore, in order to obtain the limiting variance of $M_n^1(z)$, we need to find out the limit of each terms in the expansion of $b_n(z_1)b_n(z_2)\sum_{k=1}^{p}\mathbb{E}_{k-1}[\mathbb{E}_k g_k(z_1)\mathbb{E}_k g_k(z_2)]$, i.e, the limit of (4.17)∼(4.23). The main technical point of derivation for the limit of these terms is to replace $\mathbf{H}_k(z)$ with $z(-\widehat{\mathbf{T}}_N(z))^{-1}$ and figure out the orders of the remainder. Specifically, by making use of the results in Lemmas S1.3∼S1.5, we have the limit for (4.17)∼(4.23) given in the following lemma, whose proof is relegated to the Supplementary Material (Li, Wang and Li (2021)).

LEMMA 4.2. *With the same notations as in the previous context*, (4.17) *can be approximated by*

$$\frac{2}{p}\sum_{k=1}^{p}\frac{A(z_1,z_2)}{1-\frac{k-1}{p}A(z_1,z_2)}+O(n^{-1/2}),$$

where

$$A(z_1,z_2)=\frac{m_{F^{c_n}}(z_1)m_{F^{c_n}}(z_2)pn^3}{9M^2(1+\frac{2}{3}c_nm_{F^{c_n}}(z_1))(1+\frac{2}{3}c_nm_{F^{c_n}}(z_2))}=\frac{4c_nm_{F^{c_n}}(z_1)m_{F^{c_n}}(z_2)}{m_{F^{c_n}}(z_1)-m_{F^{c_n}}(z_2)};$$

$$(4.18)=-\frac{8c_nm_{F^{c_n}}(z_1)m_{F^{c_n}}(z_2)}{15(1+\frac{2}{3}c_nm_{F^{c_n}}(z_1))(1+\frac{2}{3}c_nm_{F^{c_n}}(z_2))}+o_p(1);$$

(4.19) *and* (4.20) *share the same limit*, *which is*

$$-\frac{16c_nm_{F^{c_n}}(z_1)m_{F^{c_n}}(z_2)}{45(1+\frac{2}{3}c_nm_{F^{c_n}}(z_1))(1+\frac{2}{3}c_nm_{F^{c_n}}(z_2))};$$

all the remaining terms (4.21), (4.22) *and* (4.23) *share the same limit*, *which is*

$$\frac{16c_nm_{F^{c_n}}(z_1)m_{F^{c_n}}(z_2)}{135(1+\frac{2}{3}c_nm_{F^{c_n}}(z_1))(1+\frac{2}{3}c_nm_{F^{c_n}}(z_2))}.$$

Combining the limits of (4.17)∼(4.23) using Lemma 4.2 and then taking derivatives with respect to z_1 and z_2 gives the limit of (4.16), which is exactly given by (4.5). Therefore, we have proved that $(M_n^1(z_1)\cdots M_n^1(z_r))^{\mathsf{T}}$ is jointly Gaussian with covariance function $\mathrm{Cov}(M_n^1(z_1),M_n^1(z_2))$ given by (4.5).

4.1.3. *Proof of Step* 3: *Convergence of* $M_n^2(z)$. The proof of Lemma 4.1 will be complete if we could show that $\{M_n^2(z)\}$ for $z\in\mathcal{C}$ is bounded and form an equicontinuous family and converges to some fixed constant, which is given by (4.4). As for the boundedness and equicontinuity, it is easy to verify following the steps in Bai and Silverstein (2004) so the main task in this part is to derive the limit of $\{M_n^2(z)\}$.

Notice that for $m_{F^{c_n}}(z)$, it satisfies $\frac{1}{1+\frac{2}{3}c_nm_{F^{c_n}}(z)}=1-c_n-c_n(z-\frac{1}{3})m_{F^{c_n}}(z)$, so we define

$$A_n=\frac{1}{1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}-1+c_n+c_n\Big(z-\frac{1}{3}\Big)\mathbb{E}m_{F^{\mathbf{K}_n}}(z).$$

Consider the following:

$$
\begin{aligned}
m_{F^{c_n}}(z)A_n &= c_n\big(m_{F^{c_n}}(z) - \mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big) + c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\left[1 + \left(z - \frac{1}{3}\right)m_{F^{c_n}}(z)\right] \\
&\quad - \frac{2c_n m_{F^{c_n}}(z)\cdot \mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3 + 2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z)} \\
&= c_n\big(m_{F^{c_n}}(z) - \mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big) - \frac{4c_n^2 m_{F^{c_n}}(z)\mathbb{E}m_{F^{\mathbf{K}_n}}(z)(m_{F^{c_n}}(z) - \mathbb{E}m_{F^{\mathbf{K}_n}}(z))}{(3 + 2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))(3 + 2c_n m_{F^{c_n}}(z))},
\end{aligned}
$$

so we have

$$
\begin{aligned}
&\mathbb{E}m_{F^{\mathbf{K}_n}}(z) - m_{F^{c_n}}(z) \\
&\quad = m_{F^{c_n}}(z)A_n\left[-c_n + \frac{4c_n^2 m_{F^{c_n}}(z)\cdot\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{(3 + 2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))(3 + 2c_n m_{F^{c_n}}(z))}\right]^{-1} \\
&\quad = \frac{m_{F^{c_n}}(z)[1 + (z - \frac{1}{3})\mathbb{E}m_{F^{\mathbf{K}_n}}(z) - \frac{2\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3+2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}]}{-1 + \frac{4c_n m_{F^{c_n}}(z)\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{(3+2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))(3+2c_n m_{F^{c_n}}(z))}}.
\end{aligned}
\tag{4.24}
$$

On the other hand, with some nontrivial calculations,

$$
\begin{aligned}
&\mathrm{Tr}\big(\mathbb{E}\mathbf{D}^{-1}(z)\big) - \mathrm{Tr}\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1} \\
&\quad = p\mathbb{E}m_{F^{\mathbf{K}_n}}(z) + \frac{p - M}{z} + \frac{1}{z}\mathrm{Tr}\left(\mathbf{I}_M + \frac{p-1}{3M}\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big(\mathbf{T}^\top\mathbf{T} + \mathbf{I}_M\big)\right)^{-1} \\
&\quad = \frac{p}{z}\left[1 + \left(z - \frac{1}{3}\right)\mathbb{E}m_{F^{\mathbf{K}_n}}(z) - \frac{2\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3 + 2c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}\right] + D_n(z) + o(1),
\end{aligned}
\tag{4.25}
$$

where $D_n(z)$ is a notation to denote the constant part, that is,

$$
\begin{aligned}
D_n(z) &= \frac{2c_n^2(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}{9z} + \frac{\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3z} \\
&\quad + \frac{\frac{8}{9}c_n^3(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3 + 4c_n^2(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2 + 2\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3z(1 + \frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}.
\end{aligned}
\tag{4.26}
$$

Therefore, combining (4.24) and (4.25), we have

$$
\begin{aligned}
&p\big(\mathbb{E}m_{F^{\mathbf{K}_n}}(z) - m_{F^{c_n}}(z)\big) \\
&\quad = \frac{zm_{F^{c_n}}(z)(\mathrm{Tr}(\mathbb{E}\mathbf{D}^{-1}(z)) - \mathrm{Tr}(-\widehat{\mathbf{T}}_N(z))^{-1} - D_n(z))}{-1 + \frac{\frac{4}{9}c_n m_{F^{c_n}}(z)\cdot\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))(1+\frac{2}{3}c_n m_{F^{c_n}}(z))}} + o(1).
\end{aligned}
\tag{4.27}
$$

Thus, what remains is to find the limit of the term on the right-hand side of equation (4.27). First, considering the part involving $\mathrm{Tr}(-\widehat{\mathbf{T}}_N(z))^{-1} - \mathrm{Tr}(\mathbb{E}\mathbf{D}^{-1}(z))$, we have

$$
\begin{aligned}
&\mathrm{Tr}\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1} - \mathrm{Tr}\big(\mathbb{E}\mathbf{D}^{-1}(z)\big) \\
&\quad = \mathrm{Tr}\left[\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbb{E}\left(\sum_{k=1}^{p}\mathbf{v}_k\mathbf{v}_k^\top + \frac{p-1}{3M}z\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big(\mathbf{T}^\top\mathbf{T} + \mathbf{I}_M\big)\right)\mathbf{D}^{-1}(z)\right] \\
&\quad = \sum_{k=1}^{p}\mathbb{E}z\beta_k\mathbf{v}_k^\top\mathbf{D}_k^{-1}(z)\left(-\frac{p-1}{3M}z\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\big(\mathbf{T}^\top\mathbf{T} + \mathbf{I}_M\big) - z\mathbf{I}_M\right)^{-1}\mathbf{v}_k
\end{aligned}
$$

$$
\begin{aligned}
&-\sum_{k=1}^{p} z\mathbb{E}\beta_k \\
&\cdot \operatorname{Tr}\left[\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbb{E}\mathbf{D}^{-1}(z)\left(-\frac{p-1}{3M}z\mathbb{E}m_{F^{\mathbf{K}_n}}(z)(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)-z\mathbf{I}_M\right)^{-1}\right] \\
&+z\mathbb{E}\beta_k\cdot\operatorname{Tr}\left[\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbb{E}\mathbf{D}^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\right] \\
=&\sum_{k=1}^{p}\mathbb{E}z\beta_k\left[\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k-\frac{1}{3M}\mathbb{E}\operatorname{Tr}\mathbf{D}^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\right] \\
&-z\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\cdot\operatorname{Tr}\left[\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbb{E}\mathbf{D}^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\right].
\end{aligned}
$$

Define $r_k = z[\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\mathbf{v}_k - \frac{1}{3M}\mathbb{E}\operatorname{Tr}\mathbf{D}_k^{-1}(z)(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)]$, we have

$$
\beta_k = b_n - b_n\beta_k r_k = b_n - b_n^2 r_k + \beta_k b_n^2 r_k^2.
$$

Then

$$
\begin{aligned}
&\frac{1}{3M}\mathbb{E}\operatorname{Tr}\big(\mathbf{D}^{-1}(z)-\mathbf{D}_k^{-1}(z)\big)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M) \\
&\quad=-\frac{1}{3M}\mathbb{E}z\beta_k\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbf{D}_k^{-1}(z)\mathbf{v}_k \\
&\quad=-zb_n\mathbb{E}\operatorname{Tr}\left[\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbf{D}_k^{-1}(z)\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\right] \\
&\qquad+o\left(\frac{1}{n}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\operatorname{Tr}\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}-\operatorname{Tr}\big(\mathbb{E}\mathbf{D}^{-1}(z)\big) \\
&\quad=\sum_{k=1}^{p}\mathbb{E}z\beta_k\Big[\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k \\
&\qquad-\frac{1}{3M}\mathbb{E}\operatorname{Tr}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\Big] \\
&\qquad-zb_n^2\sum_{k=1}^{p}\mathbb{E}\operatorname{Tr}\Big[\mathbf{D}_k^{-1}(z)\left(\mathbf{I}_M+\frac{p-1}{3M}\mathbb{E}m_{F^{\mathbf{K}_n}}(z)(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\right)^{-1} \\
&\qquad\cdot\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbf{D}_k^{-1}(z)\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\Big] \\
&\qquad-z\mathbb{E}m_{F^{\mathbf{K}_n}}(z)\cdot\operatorname{Tr}\left[\frac{1}{3M}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\mathbb{E}\mathbf{D}^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\right].
\end{aligned}
\tag{4.28}
$$

As for the first term in (4.28), we have

$$
\begin{aligned}
&\sum_{k=1}^{p}\mathbb{E}z\beta_k\left[\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k-\frac{1}{3M}\mathbb{E}\operatorname{Tr}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}(\mathbf{T}^{\top}\mathbf{T}+\mathbf{I}_M)\right] \\
&\quad=-b_n^2\sum_{k=1}^{p}\mathbb{E}zr_k\mathbf{v}_k^{\top}\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k
\end{aligned}
$$

$$
\begin{aligned}
&+ b_n^2 \sum_{k=1}^{p} \mathbb{E} z\beta_k r_k^2 \Big[\mathbf{v}_k^\top \mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1} \mathbf{v}_k \\
&- \frac{1}{3M} \mathbb{E} \operatorname{Tr} \mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\big(\mathbf{T}^\top\mathbf{T} + \mathbf{I}_M\big)\Big] \\
= &-b_n^2 \sum_{k=1}^{p} \mathbb{E} z r_k \mathbf{v}_k^\top \mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k + b_n^2 \sum_{k=1}^{p} \Big\{ \mathbb{E} z\beta_k r_k^2 \mathbf{v}_k^\top \mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\mathbf{v}_k \\
&- \mathbb{E} z\beta_k r_k^2 \operatorname{Tr}\Big(\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\frac{1}{3M}\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big) \\
&+ \operatorname{Cov}\Big(z\beta_k r_k^2, \operatorname{Tr}\Big[\mathbf{D}_k^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\frac{1}{3M}\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big]\Big)\Big\} \\
= &\, z b_n^2 \sum_{k=1}^{p} \mathbb{E}\Big[\mathbf{v}_k^\top \mathbf{D}_k^{-1}(z)\mathbf{v}_k - \frac{1}{3M}\mathbb{E}\operatorname{Tr}\mathbf{D}_k^{-1}(z)\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big] \\
&\cdot \Big[\mathbf{v}_k^\top \mathbf{D}_k^{-1}(z)\Big(\mathbf{I}_M + \frac{p-1}{3M}\mathbb{E} m_{F^{\mathbf{K}_n}}(z)\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big)^{-1}\mathbf{v}_k \\
&- \frac{1}{3M}\operatorname{Tr}\Big(\mathbf{D}_k^{-1}(z)\Big(\mathbf{I}_M + \frac{p-1}{3M}\mathbb{E} m_{F^{\mathbf{K}_n}}(z)\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big)^{-1}\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\Big)\Big] + o(1).
\end{aligned}
$$

Denote $\mathcal{W}^{-1}(z) = (\mathbf{I}_M + \frac{p-1}{3M}\mathbb{E} m_{F^{\mathbf{K}_n}}(z)(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M))^{-1}$ for short, then using Lemma S1.5, we have

$$
\operatorname{Tr}\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1} - \operatorname{Tr}\big(\mathbb{E}\mathbf{D}^{-1}(z)\big)
$$

$$
= -z\mathbb{E} m_{F^{\mathbf{K}_n}}(z) \operatorname{Tr}\Big[\frac{1}{3M}\big(\mathbf{T}^\top\mathbf{T}+\mathbf{I}_M\big)\mathbb{E}\mathbf{D}^{-1}(z)\big(-\widehat{\mathbf{T}}_N(z)\big)^{-1}\Big] \tag{4.29}
$$

$$
+ z b_n^2 \frac{1}{9M^2}\sum_{k=1}^{p}\mathbb{E}\operatorname{Tr}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathbf{T}^\top\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\mathbf{T}^\top\big) \tag{4.30}
$$

$$
- z b_n^2 \frac{2}{15M^2}\sum_{k=1}^{p}\sum_{i=1}^{n}\mathbb{E}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\mathbf{T}^\top\big)_{ii} \tag{4.31}
$$

$$
\begin{aligned}
&+ z b_n^2 \frac{4}{45M^2}\sum_{k=1}^{p}\sum_{i<\ell}\mathbb{E}\big[\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\big)_{\ell,(i\ell)} \\
&- \big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\mathbf{T}^\top\big)_{\ell\ell}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\big)_{i,(i\ell)}\big]
\end{aligned} \tag{4.32}
$$

$$
\begin{aligned}
&+ z b_n^2 \frac{4}{45M^2}\sum_{k=1}^{p}\sum_{i<\ell}\mathbb{E}\big[\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\mathbf{T}^\top\big)_{ii}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\big)_{\ell,(i\ell)} \\
&- \big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathbf{T}^\top\big)_{\ell\ell}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\big)_{i,(i\ell)}\big]
\end{aligned}
$$

$$
\begin{aligned}
&+ z b_n^2 \sum_{k=1}^{p}\frac{4}{45M^2}\mathbb{E}\Big[\sum_{j<i<t}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\big)_{i,(it)}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\big)_{j,(jt)} \\
&+ \sum_{j<i<t}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z)\big)_{i,(it)}\big(\mathbf{T}\mathbf{D}_k^{-1}(z)\big)_{j,(jt)}\Big]
\end{aligned} \tag{4.33}
$$

$$+ z b_n^2 \sum_{k=1}^{p} \frac{4}{45 M^2} \mathbb{E}\Big[\sum_{t<j<i} (\mathbf{T}\mathbf{D}_k^{-1}(z))_{i,(ti)} (\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z))_{j,(tj)}$$
$$+ \sum_{t<j<i} (\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z))_{i,(ti)} (\mathbf{T}\mathbf{D}_k^{-1}(z))_{j,(tj)} \Big]$$
$$- z b_n^2 \sum_{k=1}^{p} \frac{4}{45 M^2} \mathbb{E}\Big[\sum_{j<t<i} (\mathbf{T}\mathbf{D}_k^{-1}(z))_{i,(ti)} (\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z))_{j,(jt)}$$
$$+ \sum_{j<t<i} (\mathbf{T}\mathbf{D}_k^{-1}(z)\mathcal{W}^{-1}(z))_{i,(ti)} (\mathbf{T}\mathbf{D}_k^{-1}(z))_{j,(jt)} \Big]$$
$$+ o(1).$$

Next, we look into each terms above by replacing $\mathbf{D}^{-1}(z)$ with $(-\widehat{\mathbf{T}}_N(z))^{-1}$ and figure out the orders of the remainder. Through some nontrivial calculations, we have the following lemma, whose proof is relegated to the supplement file.

LEMMA 4.3. *With the same notation as in the previous context, we have*

$$(4.29) = -\frac{\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3z}\left(1 + \frac{2}{(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}\right) + o(1),$$
$$(4.30) = \frac{4c_n b_n^2}{9z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3}\left[1 - \frac{4c_n(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}{9(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}\right]^{-1} + o(1),$$
$$(4.31) = -\frac{8c_n b_n^2}{15z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3} + o(1).$$

Moreover, the four terms in (4.32) *share the same limit, which is*

$$-\frac{8c_n b_n^2}{45z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3}$$

and all the six terms in (4.33) *share the same limit, which is*

$$\frac{8c_n b_n^2}{135z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3}.$$

Collecting all the limits in Lemma 4.3, we have

$$\begin{aligned}
&\operatorname{Tr}(-\widehat{\mathbf{T}}_N(z))^{-1} - \operatorname{Tr}(\mathbb{E}\mathbf{D}^{-1}(z)) \\
&\quad = -\frac{\mathbb{E}m_{F^{\mathbf{K}_n}}(z)}{3z}\left(1 + \frac{2}{(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}\right) \\
&\qquad + \frac{4c_n(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}{9z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3}\left[1 - \frac{4c_n(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}{9(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}\right]^{-1} \\
&\qquad - \frac{8c_n(\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^2}{9z(1+\frac{2}{3}c_n\mathbb{E}m_{F^{\mathbf{K}_n}}(z))^3} + o(1).
\end{aligned} \tag{4.34}$$

Finally, combing (4.26), (4.27) and (4.34), the limit of $\{M_n^2(z)\}$ is found to be (4.4). The proof of Lemma 4.1 is complete.

Acknowledgments. We thank the Editor, Associate Editor and referees for insightful comments that have significantly improved the paper. We are most grateful to Dr. Zhigang Bao for helpful discussions.

Qinwen Wang is the corresponding author.

Funding. The research of Zeng Li is supported by NSFC (National Natural Science Foundation of China) Grant No. 12031005.

The research of Qinwen Wang is supported by NSFC (No. 11801085) and Shanghai Sailing Program (No. 18YF1401500).

The research of Runze Li is supported by NSF Grants DMS 1820702, 1953196 and 2015539.

SUPPLEMENTARY MATERIAL

Supplement to "Central limit theorem for linear spectral statistics of large dimensional Kendall's rank correlation matrices and its applications" (DOI: 10.1214/20-AOS2013SUPP; .pdf). Additional lemmas and proofs in this paper can be found in the Supplementary Material.

REFERENCES

ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. 2nd ed. *Wiley Publications in Statistics*. Wiley, New York. MR0091588

BAI, Z. D. and SILVERSTEIN, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051 https://doi.org/10.1214/aop/1022855421

BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845

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

BAI, Z. D. and YAO, J. (2005). On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli* **11** 1059–1092. MR2189081 https://doi.org/10.3150/bj/1137421640

BANDEIRA, A. S., LODHIA, A. and RIGOLLET, P. (2017). Marčenko–Pastur law for Kendall's tau. *Electron. Commun. Probab.* **22** Paper No. 32, 7. MR3663103 https://doi.org/10.1214/17-ECP59

BAO, Z. (2019a). Tracy–Widom limit for Kendall's tau. *Ann. Statist.* **47** 3504–3532. MR4025750 https://doi.org/10.1214/18-AOS1786

BAO, Z. (2019b). Tracy–Widom limit for spearman's rho. Preprint.

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

BAO, Z., LIN, L.-C., PAN, G. and ZHOU, W. (2015). Spectral statistics of large dimensional Spearman's rank correlation matrix and its application. *Ann. Statist.* **43** 2588–2623. MR3405605 https://doi.org/10.1214/15-AOS1353

BILLINGSLEY, P. (2008). *Probability and Measure*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR0534323

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

COUILLET, R. and DEBBAH, M. (2011). *Random Matrix Methods for Wireless Communications*. Cambridge Univ. Press, Cambridge. MR2884783 https://doi.org/10.1017/CBO9780511994746

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

HAN, F., CHEN, S. and LIU, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306 https://doi.org/10.1093/biomet/asx050

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 361–379. Univ. California Press, Berkeley, CA. MR0133191

JIANG, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880. MR2052906 https://doi.org/10.1214/105051604000000143

JIANG, D., BAI, Z. and ZHENG, S. (2013). Testing the independence of sets of large-dimensional variables. *Sci. China Math.* **56** 135–147. MR3016588 https://doi.org/10.1007/s11425-012-4501-0

JIANG, T. and QI, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. Stat.* **42** 988–1009. MR3426306 https://doi.org/10.1111/sjos.12147

JOHNSTONE, I. M. (2006). High dimensional statistical inference and random matrices. Preprint. Available at arXiv:math/0611589.

LEUNG, D. and DRTON, M. (2018). Testing independence in high dimensions with sums of rank correlations. *Ann. Statist.* **46** 280–307. MR3766953 https://doi.org/10.1214/17-AOS1550

LI, Z., WANG, Q. and LI, R. (2021). Supplement to "Central limit theorem for linear spectral statistics of large dimensional Kendall's rank correlation matrices and its applications." https://doi.org/10.1214/20-AOS2013SUPP

MARČENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb.* (*N.S.*) **72 (114)** 507–536. MR0208649

MUIRHEAD, R. J. (2009). *Aspects of Multivariate Statistical Theory* **97**. Wiley, New York.

PAN, G. M. and ZHOU, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* **18** 1232–1270. MR2418244 https://doi.org/10.1214/07-AAP477

PAUL, D. and AUE, A. (2014). Random matrix theory in statistics: A review. *J. Statist. Plann. Inference* **150** 1–29. MR3206718 https://doi.org/10.1016/j.jspi.2013.09.005

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

TULINO, A. M. and VERDÚ, S. (2004). Random matrix theory and wireless communications. *Found. Trends Commun. Inf. Theory* **1** 1–182.

YANG, Y. and PAN, G. (2015). Independence test for high dimensional data based on regularized canonical correlation coefficients. *Ann. Statist.* **43** 467–500. MR3316187 https://doi.org/10.1214/14-AOS1284

YAO, J., ZHENG, S. and BAI, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis. Cambridge Series in Statistical and Probabilistic Mathematics* **39**. Cambridge Univ. Press, New York. MR3468554 https://doi.org/10.1017/CBO9781107588080

ZHENG, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 444–476. MR2954263 https://doi.org/10.1214/11-AIHP414

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

ZHENG, S., CHEN, Z., CUI, H. and LI, R. (2019). Hypothesis testing on linear structures of high-dimensional covariance matrix. *Ann. Statist.* **47** 3300–3334. MR4025743 https://doi.org/10.1214/18-AOS1779

ZHOU, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363. MR2327033 https://doi.org/10.1090/S0002-9947-07-04192-X