Statistica Sinica Preprint No: SS-2024-0078	
Title	Structural Testing of High-dimensional Correlation Matrices
Manuscript ID	SS-2024-0078
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0078
Complete List of Authors	Tingting Zou, Guangren Yang, Ruitao Lin, Guoliang Tian and Shurong Zheng
Corresponding Authors	Tingting Zou
E-mails	zoutt260@jlu.edu.cn
Notice: Accepted author version.	

Structural Testing of High-dimensional Correlation Matrices

Tingting Zou, Guangren Yang, Ruitao Lin, Guo-Liang Tian and Shurong Zheng*

Jilin University, Jinan University,

The University of Texas MD Anderson Cancer Center,

Southern University of Science and Technology and Northeast Normal University

Abstract: Due to scale invariance, correlation matrices play a critical role in multivariate statistical analysis. Statistical inference about correlation matrices encounter enormous challenges and is fundamentally different from inference about covariance matrices in both low- and high-dimensional settings. This paper studies the test of general linear structures of high-dimensional correlation matrices, which include commonly-used banded matrices and compound symmetry matrices as special cases. We first propose a procedure using the quadratic loss function to estimate the unknown parameters associated with the assumed linear structure. We then develop test statistics, based on the quadratic and infinite norms, which are suitable for dense and sparse alternatives, respectively. The limiting distributions of our proposed test statistics are derived under the null and alternative hypotheses. Extensive simulation studies are conducted to demonstrate the finite sample performance of our proposed tests. Moreover, a real data example is provided to show the applicability and the practical utility of the tests.

Key words and phrases: Correlation matrix, structural testing, high-dimensional, random matrix theory

1. Introduction

Due to scale invariance, correlation matrices play a critical role in multivariate statistical analysis, and the structural analysis of correlation matrices has a wide range of applications in genomics, psychology, finance, and other fields. For example, correlation networks are

*Corresponding author.

important in bioinformatics studies, and the identification of dependencies in microarray data is one of the key issues in gene co-expression network analysis (Oldham et al., 2006; Opgen-Rhein and Strimmer, 2007). As another example, psychological studies usually need to investigate the structure in correlated psychological variables based on factor analysis, where the inference on correlation matrices is always required (Yong and Pearce, 2013).

In the low-dimensional setting, the problem of testing a correlation matrix $\mathbf{R} = (r_{ij})_{i,j=1}^{p}$ has been well studied. For example, Bartlett and Rajalakshman (1953), Kullback (1967) and Aitkin (1969) considered the problem of testing $H_{01} : \mathbf{R} = \mathbf{R}_{*}$, where $\mathbf{R}_{*}$ is a prespecified correlation matrix; McDonald (1975), Larzelere and Mulaik (1977), Jöreskog (1978) and Steiger (1980) investigated the pattern structures of correlation matrices. High-dimensional data, for which the dimension p can be much larger than the sample size n, are increasingly encountered in contemporary scientific applications. The high dimensionality brings new challenges to the analysis of correlation matrices. In recent years, research efforts have been devoted to the independence test, that is, testing $H_{02} : \mathbf{R} = \mathbf{I}_p$, where $\mathbf{I}_p$ is the $p \times p$ dimensional identity matrix. Jiang (2004), Li and Rosalsky (2006), Zhou (2007), Liu et al. (2008), Li et al. (2010), Li et al. (2012), Cai and Jiang (2011), and Cai and Jiang (2012) studied the asymptotic distribution of $\max_{1 \le i < j \le p} n(\hat{r}_{ij})^2$ with $\widehat{\mathbf{R}}_n = (\hat{r}_{ij})_{i,j=1}^{p}$ denoting the sample correlation matrix. Schott (2005) derived the asymptotic distribution of the statistic $\mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{I}_p)^2]$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. Mestre and Vallet (2017) derived the asymptotic distributions for both $\mathrm{tr}(\widehat{\mathbf{R}}_n^2)$ and the generalized likelihood ratio statistic by establishing the central limit theorem (CLT) of linear spectral statistics (LSSs) of $\widehat{\mathbf{R}}_n$. Meanwhile, Gao et al. (2017) established the CLT of LSSs of $\widehat{\mathbf{R}}_n$ with $\mathbf{R} = \mathbf{I}_p$ and applied the statistic $\mathrm{tr}(\widehat{\mathbf{R}}_n^2)$ for testing H_{02}. Besides, Leung and Drton (2018) proposed a statistic consisting of the sum of squared sample rank correlations to test H_{02}. For testing a more general matrix $\mathbf{R}_*$ under H_{01}, Zheng et al. (2019) derived the asymptotic distribution of the statistic $\mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_*)^2]$ using random matrix theory. Yin et al. (2022) proposed three tests by establishing the CLT for LSSs of the rescaled sample correlation matrix $\widehat{\mathbf{R}}_n\mathbf{R}^{-1}$.

The objective of this paper is to develop a unified approach under the convergence regime $p/n \to y \in (0, \infty)$ for testing the general linear structures of $\mathbf{R}$ as follows:

$$H_0 : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K \quad \text{versus} \quad H_1 : \mathbf{R} \neq \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K, \tag{1.1}$$

where $\mathbf{J}_0, \ldots, \mathbf{J}_K$ $(K \geq 0)$ are $p \times p$ dimensional prespecified symmetric matrices assumed to be linearly independent, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ is an unknown parameter vector making $\mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K$ positive definite. Structural analysis of the correlation matrix is crucial in many applications, whereas the existing literature related to this topic mainly focuses on a special case of our considered problem by testing whether the correlation matrix is equal to the identity matrix $\mathbf{I}_p$ (i.e., $\mathbf{J}_0 = \mathbf{I}_p$ and $K = 0$) or a prespecified matrix $\mathbf{R}_*$ (i.e., $\mathbf{J}_0 = \mathbf{R}_*$ and $K = 0$). Through testing a more general structure under the null hypothesis H_0, one can determine whether or not the population correlation matrix $\mathbf{R}$ of the observed sample has one or multiple particular structures characterized by the basis matrices $\mathbf{J}_0, \ldots, \mathbf{J}_K$.

In fact, a pattern hypothesis on correlation matrix has many direct applications in various research fields, such as psychology and social sciences (McDonald, 1975; Jöreskog, 1978; Steiger, 1980), and it is also of interest from perspective of theoretical development. A correlational pattern hypothesis specifies that certain groups of elements in a correlation matrix are equal to a specified value or to each other, and some of special cases can be transformed into testing H_0. For example, testing a compound symmetry correlation, which has numerous applications in the analysis of multivariate repeated measures, can be done by specifying $\mathbf{J}_0 = \mathbf{I}_p$, $\mathbf{J}_1 = \mathbf{1}_p\mathbf{1}_p^T - \mathbf{I}_p$, where $\mathbf{1}_p$ is a p-dimensional vector with all entries being one. As another example, in the circumplex model, when the points on the circle are equidistant, the correlation matrix of variables forms a circular symmetric pattern that can be represented as a linear structure (see (35) in Jöreskog (1978)).

Zheng et al. (2019) studied a similar structural testing problem of the high-dimensional covariance matrix $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^p$, under the null hypothesis $H_0' : \mathbf{\Sigma} = \mathbf{A}_0 + \theta_1\mathbf{A}_1 + \cdots + \theta_K\mathbf{A}_K$, where $\mathbf{A}_0, \ldots, \mathbf{A}_K$ are a set of basic matrices. Since the correlation matrix is a standardized version of the covariance matrix and the two hypotheses H_0' and H_0 are analogous, one might easily perceive that the methodological and theoretical developments of Zheng et al. (2019) can be directly applied to test H_0. In contrast, testing correlation matrices is essentially different from testing covariance matrices in high-dimensional setting. In order to derive the asymptotic distribution of the test statistic associated with the sample correlation matrix, we decompose it into two parts: one part involves the sample covariance matrix, and the other part is related to the diagonal matrix formed by the diagonal elements of the sample covariance matrix. The contribution of the second part is non-negligible for the asymptotic distribution. Consequently, the asymptotic distributions of the statistics for testing correlation matrices and covariance matrices are completely different. We use the following example to demonstrate the difference between the two testing problems.

Example 1: Consider the Gaussian population and the high-dimensional setting where p increases proportionally with the sample size n. Under the null hypothesis $H_{02}' : \mathbf{\Sigma} = \mathbf{I}_p$ and when $y_{n-1} = p/(n-1) \to y \in (0, \infty)$, the test statistic $L_{1n} = \mathrm{tr}[(\widehat{\mathbf{\Sigma}}_n - \mathbf{I}_p)^2]$ has an asymptotic normal distribution (Chen et al., 2020),

$$\frac{L_{1n} - (py_{n-1} + y_{n-1})}{2y_{n-1}\sqrt{1 + y_{n-1}}} \xrightarrow{d} N(0, 1), \tag{1.2}$$

where $\widehat{\mathbf{\Sigma}}_n$ is the sample covariance matrix and "$\xrightarrow{d}$" stands for convergence in distribution. By analogy, Zheng et al. (2019) proposed the statistic $L_{2n} = \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{I}_p)^2]$ for testing $H_{02} : \mathbf{R} = \mathbf{I}_p$, and showed that its limiting null distribution is given by

$$\frac{L_{2n} - [n^{-1}(n^2 - n - 1)y_{n-1}^2 - p^{-2}(n^2 + 4n + 1)y_{n-1}^3]}{2y_{n-1}} \xrightarrow{d} N(0, 1). \tag{1.3}$$

Contrasting (1.2) and (1.3) indicates that L_{1n} for testing $\mathbf{\Sigma}$ and L_{2n} for testing $\mathbf{R}$ have fundamentally different behaviors under the high-dimensional setting.

In our paper, we propose several unified tests for testing the general linear structures of high-dimensional correlation matrices (1.1) based on different matrix norms. In particular, we consider two quadratic norms and the infinite norm to characterize the similarities between the unstructured and structured correlation matrices. We use the random matrix theory to study the limiting behaviors of the proposed tests under both null and alternative scenarios, and also show that the asymptotic results for testing H_0 are different from those derived by Zheng et al. (2019) for testing H_0'. Actually, the aforementioned Example 1 can be treated as a special case of H_0 with $\mathbf{J}_0 = \mathbf{I}_p$ and $K = 0$, the existing results given in (1.2)-(1.3) preliminarily showcase the differences.

The arrangement of this paper is as follows: In Section 2, we introduce the structural testing problem and develop a useful central limit theorem as preparatory work. We then propose three novel statistics for testing the hypothesis (1.1), with two constructed using the quadratic norm as given in Section 3 and one based on the infinite norm in Section 4. We report a simulation study to examine the performance of the three new tests in Section 5. We also use a real data analysis to illustrate our methods in Section 6. Finally, in Section 7, we provide some concluding remarks and also discuss the combination tests. All technical details and additional simulation results are presented in the Supplementary Material.

2. Notations and preliminaries

Let $\{\mathbf{x}_k = (x_{1k}, ..., x_{pk})^T, k = 1, \ldots, n\}$ be n independent and identically distributed (i.i.d.) observations from a population with the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^p$. The population correlation matrix is defined as $\mathbf{R} = (r_{ij})_{i,j=1}^p = [\mathrm{diag}(\boldsymbol{\Sigma})]^{-1/2}\boldsymbol{\Sigma}[\mathrm{diag}(\boldsymbol{\Sigma})]^{-1/2}$, where $\mathrm{diag}(\boldsymbol{\Sigma})$ denotes the diagonal matrix formed by the diagonal elements of $\boldsymbol{\Sigma}$. The sample covariance matrix and sample correlation matrix are

$$\widehat{\boldsymbol{\Sigma}}_n = (\hat{\sigma}_{ij})_{i,j=1}^p = n^{-1}\sum_{k=1}^n (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})^T, \tag{2.4}$$

$$\widehat{\mathbf{R}}_n = (\hat{r}_{ij})_{i,j=1}^p = [\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)]^{-1/2}\widehat{\boldsymbol{\Sigma}}_n[\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)]^{-1/2}, \tag{2.5}$$

respectively, where $\bar{\mathbf{x}} = n^{-1}\sum_{k=1}^n \mathbf{x}_k = (\bar{x}_1, ..., \bar{x}_p)^T$ is the sample mean. For ease of reference, Table S1 in the Supplementary Material provides the definitions and descriptions of the key notations used in the paper.

2.1 Central limit theorem

The theoretical results of our proposed tests rely on the CLT for two-dimensional random vector $\left(\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_1\big), \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\right)$, where $\mathbf{D}_1$ and $\mathbf{D}_2$ are two non-random symmetric matrices. As a preliminary step, we first establish the CLT for $\left(\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1), \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\right)$ under the following assumptions:

- **Assumption A.** Let the sample $\mathbf{x}_k = (x_{1k}, ..., x_{pk})^T$ for $k = 1, ..., n$ be i.i.d. from the model $\mathbf{x}_k = \boldsymbol{\Sigma}^{1/2}\mathbf{w}_k + \boldsymbol{\mu}$, where $\mathbf{w}_k = (w_{1k}, \ldots, w_{pk})^T$, $\{w_{\ell k}, \ell = 1, \ldots, p, k = 1, \ldots, n\}$ are i.i.d. with $\mathrm{E}(w_{\ell k}) = 0$, $\mathrm{E}(w_{\ell k}^2) = 1$, $\mathrm{E}(w_{\ell k}^4(\log(|w_{\ell k}|))^{2+2\epsilon}) < \infty$ for a small positive constant ϵ, and the kurtosis $\beta_w = \mathrm{E}(w_{\ell k}^4) - 3$.

- **Assumption B.** The spectral norm of the population correlation matrix $\mathbf{R}$ is bounded uniformly for all p.

- **Assumption C.** The data dimension p increases proportionally with the sample size n, i.e., $y_n = p/n \to y \in (0, \infty)$.

Assumptions A-C are commonly used in random matrix theory. Assumption A imposes the independent component structure on $\mathbf{x}_k$ and the same moment condition on $w_{\ell k}$ as in Noureddine (2009). Assumption B requires that the spectral norm of $\mathbf{R}$ is uniformly bounded. Assumption C specifies the convergence regime of the data dimension and sample size.

Lemma 1. *Under Assumptions A-B-C, if the spectral norms of $\mathbf{D}_1$ and $\mathbf{D}_2$ are uniformly bounded in p, then we have*

$$[\Lambda_n(\mathbf{D}_1, \mathbf{D}_2)]^{-1/2} \begin{pmatrix} \mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_1) - \nu_1(\mathbf{D}_1) \\ \mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_2 \widehat{\mathbf{R}}_n \mathbf{D}_2) - \nu_2(\mathbf{D}_2) \end{pmatrix} \xrightarrow{d} N_2\left((0,0)^T, \mathbf{I}_2\right),$$

where $\Lambda_n(\mathbf{D}_1, \mathbf{D}_2) = \begin{pmatrix} \sigma_{11}(\mathbf{D}_1) & \sigma_{12}(\mathbf{D}_1, \mathbf{D}_2) \\ \sigma_{21}(\mathbf{D}_1, \mathbf{D}_2) & \sigma_{22}(\mathbf{D}_2) \end{pmatrix}$*, the expressions for $\nu_1(\mathbf{D}_1)$, $\nu_2(\mathbf{D}_2)$, $\sigma_{11}(\mathbf{D}_1)$, $\sigma_{12}(\mathbf{D}_1, \mathbf{D}_2)$, and $\sigma_{22}(\mathbf{D}_2)$ are presented in the Supplementary Material.*

The proof of Lemma 1 is deferred to the Supplementary Material, we only sketch the main idea. We start by introducing some notations. Let $\mathbf{y}_k = \mathbf{\Gamma}\mathbf{w}_k$ with $\mathbf{\Gamma} = [\mathrm{diag}(\mathbf{\Sigma})]^{-1/2}\mathbf{\Sigma}^{1/2}$, and $\widehat{\mathbf{\Sigma}}_n^* = n^{-1}\sum_{k=1}^{n}(\mathbf{y}_k - \bar{\mathbf{y}})(\mathbf{y}_k - \bar{\mathbf{y}})^T$, then $\widehat{\mathbf{R}}_n$ can be written as

$$\widehat{\mathbf{R}}_n = [\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*)]^{-1/2}\widehat{\mathbf{\Sigma}}_n^*[\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*)]^{-1/2}.$$

The proof of Lemma 1 is divided into four steps.

Step 1: Truncation, centralization and rescaling. Denote $\tilde{\mathbf{w}}_k = (\tilde{w}_{1k}, \ldots, \tilde{w}_{pk})^T$ with $\tilde{w}_{\ell k} = (\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})/\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}$, $\breve{w}_{\ell k} = w_{\ell k} I\{|w_{\ell k}| < \eta_n \sqrt{n}\}$ and $\eta_n = (\log n)^{-(1+\epsilon)/2}$. Let $\tilde{\mathbf{y}}_k = \mathbf{\Gamma} \tilde{\mathbf{w}}_k$ and $\tilde{\mathbf{R}}_n$ be the corresponding sample correlation matrix. In this step, we prove that $|\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_1) - \mathrm{tr}(\tilde{\mathbf{R}}_n \mathbf{D}_1)| = o_p(1)$ and $|\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_2 \widehat{\mathbf{R}}_n \mathbf{D}_2) - \mathrm{tr}(\tilde{\mathbf{R}}_n \mathbf{D}_2 \tilde{\mathbf{R}}_n \mathbf{D}_2)| = o_p(1)$. Therefore, in the subsequent proof, we assume that $|w_{\ell k}| \le \eta_n \sqrt{n}$, $\mathrm{E}w_{\ell k} = 0$ and $\mathrm{E}w_{\ell k}^2 = 1$.

Step 2: Expansions for $\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_1)$ *and* $\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_2 \widehat{\mathbf{R}}_n \mathbf{D}_2)$. By the Taylor expansion

$$(\hat{\sigma}^*_{\ell\ell})^{-1/2} = 1 - \frac{1}{2}(\hat{\sigma}^*_{\ell\ell} - 1) + \frac{3}{8}(\hat{\sigma}^*_{\ell\ell} - 1)^2 + o_p(n^{-1})$$

for $\ell = 1, \ldots, p$, where $\hat{\sigma}^*_{\ell\ell}$ denotes the ℓth diagonal element of $\widehat{\mathbf{\Sigma}}^*_n$, we decompose $\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_1)$ and $\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_2 \widehat{\mathbf{R}}_n \mathbf{D}_2)$ into

$$\begin{aligned}
\mathrm{tr}(\widehat{\mathbf{R}}_n \mathbf{D}_1) &= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_1\big) - \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_1\big] + \frac{3}{4}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)^2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_1\big] \\
&\quad + \frac{1}{4}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_1\big] + o_p(1), \qquad (2.6)
\end{aligned}$$

$$\begin{aligned}
\mathrm{tr}\big(\widehat{\mathbf{R}}_n \mathbf{D}_2 \widehat{\mathbf{R}}_n \mathbf{D}_2\big) &= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big) - 2\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big] \qquad (2.7) \\
&\quad + \frac{3}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)^2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big] \\
&\quad + \frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big] \\
&\quad + \frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_2 \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n\big] \\
&\quad + \frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big] + o_p(1).
\end{aligned}$$

Step 3: Analysis of the constant order terms. We show that the following terms

$$\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)^2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_1\big], \quad \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_1\big],$$

$$\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)^2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big], \quad \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big],$$

$$\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\mathbf{D}_2 \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \widehat{\mathbf{\Sigma}}^*_n\big],$$

$$\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2 \big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}^*_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}^*_n \mathbf{D}_2\big]$$

converge in probability to their expectations.

Step 4: Analysis of the main terms. Using the martingale central limit theorem, we derive the CLT for the following random vector

$$\begin{pmatrix} \operatorname{tr}\big(\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_1\big) - \operatorname{tr}\big[\big(\operatorname{diag}(\widehat{\boldsymbol{\Sigma}}_n^*) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_1\big] \\ \operatorname{tr}\big(\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2\big) - 2\operatorname{tr}\big[\big(\operatorname{diag}(\widehat{\boldsymbol{\Sigma}}_n^*) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2\big] \end{pmatrix}.$$

Then based on the Slutsky's theorem, we can obtain the CLT for $\Big(\operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1), \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\Big)$.

Remark 1. This newly developed CLT in Lemma 1 is of independent interest. In other words, this lemma is not only useful in testing H_0, but also can be applied to other inferences on high-dimensional correlation matrices, especially when $\mathbf{D}_1 \neq \mathbf{D}_2$. For example,

1. Yin et al. (2022) and Zheng et al. (2019) considered the problem of testing $H_{01} : \mathbf{R} = \mathbf{R}_*$ and proposed the statistics $T_Y = \operatorname{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_*^{-1} - \mathbf{I}_p)^2]$ and $T_Z = \operatorname{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_*)^2]$ respectively. Let $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{R}_*^{-1}$ and $\mathbf{D}_1 = \mathbf{R}_*$, $\mathbf{D}_2 = \mathbf{I}_p$, we can use Lemma 1 to derive the asymptotic distributions of T_Y and T_Z simultaneously.

2. Mimicking the proof of Lemma 1, we can obtain the CLT for

$$\Big(\operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1), \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_3), \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2), \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_4\widehat{\mathbf{R}}_n\mathbf{D}_4)\Big),$$

   where $\mathbf{D}_3$ and $\mathbf{D}_4$ are $p\times p$ dimensional non-random symmetric matrices with uniformly bounded spectral norms. From this new CLT, we can derive the joint asymptotic distribution of T_Z and T_Y, which can be used to construct the test based on the statistic $T_M = \max\{|T_Z - \mu_Z|/\sigma_Z, |T_Y - \mu_Y|/\sigma_Y\}$ for testing H_{01}, where μ_Z, μ_Y and σ_Z, σ_Y denotes the asymptotic means and standard deviations respectively.

Remark 2. From the expansions (2.6) and (2.7) for $\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1)$ and $\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)$ in Step 2, we get

$$\begin{aligned}\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1) &= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_1\big) + R_1(\widehat{\mathbf{\Sigma}}_n^*, \mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*), \mathbf{D}_1) + o_p(1),\\ \mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2\big) &= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2\big) + R_2(\widehat{\mathbf{\Sigma}}_n^*, \mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*), \mathbf{D}_2) + o_p(1).\end{aligned}$$

The analyses in Steps 3-4 demonstrate that $R_1(\widehat{\mathbf{\Sigma}}_n^*, \mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*), \mathbf{D}_1)$ and $R_2(\widehat{\mathbf{\Sigma}}_n^*, \mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*), \mathbf{D}_2)$ are not asymptotically negligible, it follows that the CLTs for random vectors

$$\Big(\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1), \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\Big) \quad \text{and} \quad \Big(\mathrm{tr}(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_1), \mathrm{tr}(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2)\Big)$$

are essentially different in high-dimensional setting. Since the sample covariance matrix of the standardized data is identical to the sample correlation matrix of the raw data, if we standardize the data and employ the tests developed from $\Big(\mathrm{tr}(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_1), \mathrm{tr}(\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n^*\mathbf{D}_2)\Big)$ instead of $\Big(\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1), \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\Big)$, we may obtain incorrect test results.

2.2 Structural testing of correlation matrices

We are interested in studying the linear structure of the population correlation matrix $\mathbf{R}$ by testing the following hypothesis:

$$H_0 : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K \quad \text{versus} \quad H_1 : \mathbf{R} \neq \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K.$$

Without the linear structure, a natural estimator of $\mathbf{R}$ is the sample correlation matrix $\widehat{\mathbf{R}}_n$ given in (2.5). Under H_0, the specific representation of $\mathbf{R}$ is still unknown because the parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ is unspecified. We propose a structured estimator of $\mathbf{R}$ under H_0, denoted by $\widehat{\mathbf{R}}_0$, based on the least squares method. Specifically,

$$\widehat{\mathbf{R}}_0 = (\hat{r}_{0ij})_{i,j=1}^p = \mathbf{J}_0 + \hat{\theta}_1\mathbf{J}_1 + \cdots + \hat{\theta}_K\mathbf{J}_K,$$

where

$$\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_K)^T = \text{Arg}\left\{ \min_{(\theta_1,\ldots,\theta_K)} \text{tr}\left[\left(\widehat{\mathbf{R}}_n - \mathbf{J}_0 - \sum_{k=1}^{K} \theta_k \mathbf{J}_k\right)^2\right]\right\} = \mathbf{A}^{-1}\hat{\mathbf{a}}, \tag{2.8}$$

with $\mathbf{A} = \big(\text{tr}(\mathbf{J}_i\mathbf{J}_j)\big)_{i,j=1}^K$ and $\hat{\mathbf{a}} = (\text{tr}[(\widehat{\mathbf{R}}_n - \mathbf{J}_0)\mathbf{J}_1], \ldots, \text{tr}[(\widehat{\mathbf{R}}_n - \mathbf{J}_0)\mathbf{J}_K])^T$. The invertibility of $\mathbf{A}$ follows from the linear independence of the basis matrices $\mathbf{J}_1, \ldots, \mathbf{J}_K$.

- **Assumption D.** For $k = 0, 1, \ldots, K$, the spectral norm of $\mathbf{J}_k$ is bounded uniformly for all p. In addition, the basis matrices $\mathbf{J}_0, \mathbf{J}_1, \ldots, \mathbf{J}_K$ are linearly independent.

The following Theorem 1 shows the estimation consistency for the unknown parameter vector $\boldsymbol{\theta}$, and its proof is given in the Supplementary Material.

Theorem 1. *Under Assumptions A-B-C-D and the null hypothesis H_0, we have $\hat{\theta}_k$ converges in probability to θ_k for $k = 1, \ldots, K$.*

3. Tests based on the quadratic norm

Motivated by the quadratic norm, two test statistics are constructed based on the ratio and difference between $\widehat{\mathbf{R}}_n$ and $\widehat{\mathbf{R}}_0$, respectively. Formally, when $\widehat{\mathbf{R}}_0$ is invertible, we consider the ratio-based test statistic

$$T_{1n} = \text{tr}[(\widehat{\mathbf{R}}_n\widehat{\mathbf{R}}_0^{-1} - \mathbf{I}_p)^2].$$

In addition, we can also consider the difference-based test statistic

$$T_{2n} = \text{tr}[(\widehat{\mathbf{R}}_n - \widehat{\mathbf{R}}_0)^2].$$

Intuitively, if H_0 correctly specifies the structure of the population matrix $\mathbf{R}$, the discrepancy between the sample estimator $\widehat{\mathbf{R}}_n$ and the structured estimator $\widehat{\mathbf{R}}_0$ should be small. On the other side, if the structure of $\mathbf{R}$ is misspecified, then we should expect that the ratio-based T_{1n} or the difference-based T_{2n} is large enough so that the null hypothesis H_0 is rejected.

3.1 Limiting distributions of T_{1n} and T_{2n}

Denote

$$\mathbf{R}_P = \mathbf{J}_0 + \theta_{1P}\mathbf{J}_1 + \cdots + \theta_{KP}\mathbf{J}_K,$$

where $\boldsymbol{\theta}_P = (\theta_{1P}, ..., \theta_{KP})^T = \mathbf{A}^{-1}\mathbf{a}_P$ and $\mathbf{a}_P = (a_{1P}, ..., a_{KP})^T$ with $a_{kP} = \mathrm{tr}[(\mathbf{R} - \mathbf{J}_0)\mathbf{J}_k]$ for $k = 1, \ldots, K$. The subscript P indicates "projection" because $\mathbf{R}_P - \mathbf{J}_0$ is the L_2 projection of $\mathbf{R} - \mathbf{J}_0$ onto the space spanned by $\mathbf{J}_1, \ldots, \mathbf{J}_K$. As a result, both $\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p$ and $\mathbf{R} - \mathbf{R}_P$ characterize the departure of $\mathbf{R}$ from the null hypothesis. Define

$$\mathbf{H} = n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{R}_P^{-1})\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1} + \mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p),$$
$$\mathbf{B}_P = \sum_{k=1}^{K} h_{kP}\mathbf{J}_k, \quad h_{kP} = (\mathrm{tr}(\mathbf{J}_1\mathbf{H}), \ldots, \mathrm{tr}(\mathbf{J}_K\mathbf{H}))\mathbf{A}^{-1}\mathbf{e}_k,$$
$$\tilde{\mathbf{B}}_P = \sum_{k=1}^{K} \tilde{h}_{kP}\mathbf{J}_k, \quad \tilde{h}_{kP} = \big(\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_1], \ldots, \mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_K]\big)\mathbf{A}^{-1}\mathbf{e}_k,$$

where $\mathbf{e}_k$ denotes the kth column of $\mathbf{I}_K$.

Theorem 2. *Under Assumptions A-B-C-D, it holds that*

(a) if the spectral norm of $\mathbf{R}_P^{-1}$ is uniformly bounded in p, then $\sigma_{1n}^{-1}(T_{1n} - \mu_{1n}) \xrightarrow{d} N(0, 1)$;

(b) $\sigma_{2n}^{-1}(T_{2n} - \mu_{2n}) \xrightarrow{d} N(0, 1)$,

where the mean and variance terms are given by

$$\begin{aligned}
\mu_{1n} &= \nu_2(\mathbf{R}_P^{-1}) - 2\nu_1(\mathbf{R}_P^{-1} + \mathbf{B}_P) + 2\mathrm{tr}(\mathbf{R}\mathbf{B}_P) + p,\\
\sigma_{1n}^2 &= \sigma_{22}(\mathbf{R}_P^{-1}) + 4\sigma_{11}(\mathbf{R}_P^{-1} + \mathbf{B}_P) - 4\sigma_{12}(\mathbf{R}_P^{-1} + \mathbf{B}_P, \mathbf{R}_P^{-1}),\\
\mu_{2n} &= \nu_2(\mathbf{I}_p) - 2\nu_1(\mathbf{R}_P + \tilde{\mathbf{B}}_P) + \mathrm{tr}(\mathbf{R}_P^2) + 2\mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P),\\
\sigma_{2n}^2 &= \sigma_{22}(\mathbf{I}_p) + 4\sigma_{11}(\mathbf{R}_P + \tilde{\mathbf{B}}_P) - 4\sigma_{12}(\mathbf{R}_P + \tilde{\mathbf{B}}_P, \mathbf{I}_p),
\end{aligned}$$

and $\nu_1(\cdot), \nu_2(\cdot), \sigma_{11}(\cdot), \sigma_{12}(\cdot, \cdot), \sigma_{22}(\cdot)$ are defined in Lemma 1.

The proof of Theorem 2 is deferred to the Supplementary Material, we give a brief description of the proof idea. We begin by considering T_{1n}. Based on the equation

$$\widehat{\mathbf{R}}_0^{-1} = \mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1} + \sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(\hat{\theta}_{k_1} - \theta_{k_1P})(\hat{\theta}_{k_2} - \theta_{k_2P})\widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1},$$

and the conclusion that $\hat{\theta}_k = \theta_{kP} + O_p(n^{-1})$ for $k = 1, \ldots, K$, we have

$$\begin{aligned} T_{1n} &= \mathrm{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) \\ &\quad -2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) + o_p(1). \end{aligned}$$

For $p \times p$ dimensional non-random symmetric matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ with uniformly bounded spectral norms, we have $p^{-1}\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{M}_1) = p^{-1}\mathrm{tr}(\mathbf{R}\mathbf{M}_1) + o_p(1)$, and

$$p^{-1}\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{M}_1\widehat{\mathbf{R}}_n\mathbf{M}_2) = y_n p^{-1}\mathrm{tr}(\mathbf{R}\mathbf{M}_1)p^{-1}\mathrm{tr}(\mathbf{R}\mathbf{M}_2) + p^{-1}\mathrm{tr}(\mathbf{R}\mathbf{M}_1\mathbf{R}\mathbf{M}_2) + o_p(1).$$

It follows that

$$\begin{aligned} T_{1n} &= \mathrm{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathrm{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) \\ &\quad -2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\big[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{R}_P^{-1})\mathrm{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) + \mathrm{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1})\big] + o_p(1). \end{aligned}$$

After calculation and simplification, we obtain that

$$T_{1n} = \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1})^2 - 2\mathrm{tr}[\widehat{\mathbf{R}}_n(\mathbf{R}_P^{-1} + \mathbf{B}_P)] + 2\mathrm{tr}(\mathbf{R}\mathbf{B}_P) + p + o_p(1).$$

Let $\mathbf{D}_1 = \mathbf{R}_P^{-1} + \mathbf{B}_P$ and $\mathbf{D}_2 = \mathbf{R}_P^{-1}$, then from Lemma 1 and the Delta method, we get $\sigma_{1n}^{-1}(T_{1n} - \mu_{1n}) \xrightarrow{d} N(0,1)$. Similarly, based on the equation $\widehat{\mathbf{R}}_0 - \mathbf{R}_P = \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathbf{J}_k$ and the conclusion that $\hat{\theta}_k = \theta_{kP} + O_p(n^{-1})$ for $k = 1, \ldots, K$, we have

$$T_{2n} = \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_k] + o_p(1).$$

After calculation, we get

$$T_{2n} = \mathrm{tr}(\widehat{\mathbf{R}}_n^2) - 2\mathrm{tr}[\widehat{\mathbf{R}}_n(\mathbf{R}_P + \tilde{\mathbf{B}}_P)] + \mathrm{tr}(\mathbf{R}_P^2) + 2\mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P) + o_p(1).$$

Let $\mathbf{D}_1 = \mathbf{R}_P + \tilde{\mathbf{B}}_P$ and $\mathbf{D}_2 = \mathbf{I}_p$, still based on Lemma 1 and the Delta method, we get $\sigma_{2n}^{-1}(T_{2n} - \mu_{2n}) \xrightarrow{d} N(0,1)$.

3.2 Testing methods

In this section, we first derive the asymptotic distributions of T_{1n} and T_{2n} under the null hypothesis H_0, and then construct the corresponding testing methods. Let $\mathbf{R}_0$ represent the structured population correlation matrix under H_0, that is,

$$\mathbf{R}_0 = (r_{0ij})_{i,j=1}^p = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \cdots + \theta_K\mathbf{J}_K,$$

$\mathbf{B} = \sum_{k=1}^{K} h_k\mathbf{J}_k$, $h_k = (\mathrm{tr}(\mathbf{J}_1\mathbf{R}_0^{-1}), \ldots, \mathrm{tr}(\mathbf{J}_K\mathbf{R}_0^{-1}))\mathbf{A}^{-1}\mathbf{e}_k$, and $\mathbf{C}_1 = \mathbf{R}_0^{-1} + y_n\mathbf{B}$. Denote $\mathbf{C}_0 = (c_{0ij})_{i,j=1}^p$ as a $p \times p$ dimensional matrix with $c_{0ij} = 2r_{0ij}^3 + \beta_w r_{0ij}\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2$, where β_w is the kurtosis defined in Assumption A and $\boldsymbol{\Gamma} = [\mathrm{diag}(\boldsymbol{\Sigma})]^{-1/2}\boldsymbol{\Sigma}^{1/2}$. Moreover, $\mathbf{e}_i$ denotes the ith column of the identity matrix, and its dimension is determined by the matrix in the product.

The following theorem provides the limiting null distributions of T_{1n} and T_{2n}, which is essentially a corollary of Theorem 2, and its proof is deferred to the Supplementary Material.

Theorem 3. *Under Assumptions A-B-C-D and the null hypothesis H_0, we have*

(a) if the spectral norm of $\mathbf{R}_0^{-1}$ is uniformly bounded in p, then $\sigma_{10}^{-1}(T_{1n} - \mu_{10}) \xrightarrow{d} N(0,1)$;

(b) $\sigma_{20}^{-1}(T_{2n} - \mu_{20}) \xrightarrow{d} N(0,1)$,

where the expressions for μ_{10}, μ_{20}, σ_{10}^2, *and* σ_{20}^2 *are given by*

$$\begin{aligned}
\mu_{10} &= \nu_2(\mathbf{R}_0^{-1}) - 2\nu_1(\mathbf{C}_1) + 2y_n \mathrm{tr}(\mathbf{R}_0\mathbf{B}) + p \qquad (3.9)\\
&= py_n - 3y_n^2 - 7y_n + \beta_w y_n - 0.5n^{-1}\mathrm{tr}(\mathbf{C}_0\mathbf{C}_1)\\
&\quad -(2y_n+4)\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{R}_0^{-1}\boldsymbol{\Gamma}\mathbf{e}_\ell\\
&\quad +\Big(1.5y_n+2\Big)n^{-1}\Big[2p + \beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big]\\
&\quad +\Big(0.5y_n+1\Big)n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}_0^{-1}\mathbf{e}_j\\
&\qquad \times\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1) + \beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{C}_1\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]\\
&\quad -1.5n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1) + \beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big],\\
\mu_{20} &= \nu_2(\mathbf{I}_p) - 2\nu_1(\mathbf{R}_0) + \mathrm{tr}(\mathbf{R}_0^2) \qquad (3.10)\\
&= py_n + y_n^2 + n^{-1}\mathrm{tr}(\mathbf{R}_0^2) + \beta_w n^{-1}\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2 - 0.5n^{-1}\mathrm{tr}(\mathbf{C}_0\mathbf{R}_0)\\
&\quad -2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2) + \beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{R}_0\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]\\
&\quad +0.5n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2) + \beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}_0^2\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big]\\
&\quad +n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big],\\
\sigma_{10}^2 &= \sigma_{22}(\mathbf{R}_0^{-1}) + 4\sigma_{11}(\mathbf{C}_1) - 4\sigma_{12}(\mathbf{C}_1,\mathbf{R}_0^{-1}) \qquad (3.11)\\
&= 4y_n^2 - 4(2+\beta_w)y_n(1+y_n)^2\\
&\quad +4(1+y_n)^2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2) + \beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +4n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{C}_1\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]
\end{aligned}$$

$$+4n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\mathbf{e}_j^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_j$$
$$\times\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]$$
$$-8n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\Big[2\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{R}_0\mathbf{e}_i+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{C}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big]$$
$$+8(1+y_n)n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2\mathbf{C}_1)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{e}_k\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{C}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big]$$
$$-8(1+y_n)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\Big[2\mathbf{e}_i^T\mathbf{R}_0^2\mathbf{e}_i+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{e}_k\Big],$$

and

$$\sigma_{20}^2=\sigma_{22}(\mathbf{I}_p)+4\sigma_{11}(\mathbf{R}_0)-4\sigma_{12}(\mathbf{R}_0,\mathbf{I}_p)=4[n^{-1}\mathrm{tr}(\mathbf{R}_0^2)]^2. \tag{3.12}$$

Remark 3. From the expansion (S3.26) for T_{1n} in the Supplementary Material, we have $T_{1n}=\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_0^{-1}-\mathbf{I}_p)^2-2y_n\big[\mathrm{tr}(\widehat{\mathbf{R}}_0\mathbf{R}_0^{-1})-p\big]+o_p(1)$, which implies that the effect of estimating the unknown parameter vector $\boldsymbol{\theta}$ on the asymptotic distribution of T_{1n} is non-negligible. When $\mathbf{R}_0$ is known, the second term in the expansion becomes 0, then T_{1n} will be reduced to the commonly used statistic $\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_0^{-1}-\mathbf{I}_p)^2$ for testing H_{01}. As for T_{2n}, based on the expansion (S3.27) about T_{2n} in the Supplementary Material, the asymptotic distribution of T_{2n} is the same as that of $\mathrm{tr}[(\widehat{\mathbf{R}}_n-\mathbf{R}_0)^2]$ whether R_0 is known or not.

Remark 4. From Theorem 3, we find that the expressions for μ_{10} and σ_{10}^2 are more complicated than the asymptotic mean and variance of T_{n2} given in Zheng et al. (2019). This is because we derive the asymptotic distribution of T_{1n} based on the CLT for $\Big(\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1),\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2)\Big)$ rather than $\Big(\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_1),\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n^*\mathbf{D}_2)\Big)$. Therefore, if we just standardize the data and employ the test proposed by Zheng et al. (2019), we may obtain an incorrect test result.

Expressions for $\mu_{10},\mu_{20},\sigma_{10}^2,\sigma_{20}^2$ given in equations (3.9)-(3.10)-(3.11)-(3.12) involve five unknown quantities, i.e., $\mathbf{R}_0$, $\mathbf{\Gamma}$, β_w, $\mathbf{C}_0$, and $\mathbf{C}_1$. Under the conditions of Theorem 3,

$$\|\widehat{\mathbf{R}}_0 - \mathbf{R}_0\| = \Big\| \sum_{k=1}^{K} (\hat{\theta}_k - \theta_k)\mathbf{J}_k \Big\| \leq \sum_{k=1}^{K} |\hat{\theta}_k - \theta_k| \|\mathbf{J}_k\| = o_p(1),$$

where $\|\cdot\|$ denotes the spectral norm of a matrix, then we replace $\mathbf{R}_0$ by $\widehat{\mathbf{R}}_0$ in the expressions. If $\max_{1\leq\ell\leq p} \sigma_{\ell\ell} < M$, similar to the proof of Lemma 5.7 in Yin and Ma (2022), we have

$$\max_{1\leq\ell\leq p} |\hat{\sigma}_{\ell\ell} - \sigma_{\ell\ell}| = O_p(n^{-1/2}). \tag{3.13}$$

Let $\widehat{\mathbf{\Sigma}}_0 = [\text{diag}(\widehat{\mathbf{\Sigma}}_n)]^{1/2}\widehat{\mathbf{R}}_0[\text{diag}(\widehat{\mathbf{\Sigma}}_n)]^{1/2}$, from (3.13), if $\min_{1\leq\ell\leq p} \sigma_{\ell\ell} > m$, we have

$$\big\|\widehat{\mathbf{\Sigma}}_0 - \mathbf{\Sigma}\big\| = o_p(1), \quad \big\|[\text{diag}(\widehat{\mathbf{\Sigma}}_n)]^{-1/2} - [\text{diag}(\mathbf{\Sigma})]^{-1/2}\big\| = o_p(1).$$

Thus, we substitute $\widehat{\mathbf{\Gamma}} = [\text{diag}(\widehat{\mathbf{\Sigma}}_n)]^{-1/2}\widehat{\mathbf{\Sigma}}_0^{1/2}$ for $\mathbf{\Gamma}$. According to Theorem 2.2 in Zheng et al. (2019), we obtain that

$$\hat{\beta}_w = \frac{\hat{V} - 2\big[\text{tr}(\widehat{\mathbf{\Sigma}}_n^2) - n^{-1}\text{tr}^2(\widehat{\mathbf{\Sigma}}_n)\big]}{\sum_{\ell=1}^{p} \hat{\sigma}_{\ell\ell}^2}$$

is a consistent estimate of β_w, where

$$\hat{V} = (n-1)^{-1} \sum_{k=1}^{n} \Big\{ (\mathbf{x}_k - \bar{\mathbf{x}})^T(\mathbf{x}_k - \bar{\mathbf{x}}) - n^{-1} \sum_{k=1}^{n} \big[(\mathbf{x}_k - \bar{\mathbf{x}})^T(\mathbf{x}_k - \bar{\mathbf{x}})\big] \Big\}^2.$$

The matrices $\mathbf{C}_0$ and $\mathbf{C}_1$ can be estimated by $\widehat{\mathbf{C}}_0$ and $\widehat{\mathbf{C}}_1$, where the (i,j)th entry of $\widehat{\mathbf{C}}_0$ is

$$\hat{c}_{0ij} = 2\hat{r}_{0ij}^3 + \hat{\beta}_w \hat{r}_{0ij} \sum_{k=1}^{p} (\mathbf{e}_i^T \widehat{\mathbf{\Gamma}} \mathbf{e}_k)^2 (\mathbf{e}_j^T \widehat{\mathbf{\Gamma}} \mathbf{e}_k)^2,$$

and $\widehat{\mathbf{C}}_1 = \widehat{\mathbf{R}}_0^{-1} + y_n \widehat{\mathbf{B}}$ with $\widehat{\mathbf{B}} = \sum_{k=1}^{K} \hat{h}_k \mathbf{J}_k$ and $\hat{h}_k = (\text{tr}(\mathbf{J}_1\widehat{\mathbf{R}}_0^{-1}), \ldots, \text{tr}(\mathbf{J}_K\widehat{\mathbf{R}}_0^{-1}))\mathbf{A}^{-1}\mathbf{e}_k$. As a result, the estimators of $\mu_{10}, \mu_{20}, \sigma_{10}^2, \sigma_{20}^2$, denoted by $\hat{\mu}_{10}, \hat{\mu}_{20}, \hat{\sigma}_{10}^2, \hat{\sigma}_{20}^2$, are obtained by replacing $\mathbf{R}_0$, $\mathbf{\Gamma}$, β_w, $\mathbf{C}_0$, and $\mathbf{C}_1$ by $\widehat{\mathbf{R}}_0$, $\widehat{\mathbf{\Gamma}}$, $\hat{\beta}_w$, $\widehat{\mathbf{C}}_0$, and $\widehat{\mathbf{C}}_1$, respectively.

Theorem 4. *Under the conditions of Theorem 3, if $m < \min_{1\leq\ell\leq p} \sigma_{\ell\ell} \leq \max_{1\leq\ell\leq p} \sigma_{\ell\ell} < M$ with m and M being positive constants, we have*

$$\hat{\sigma}_{10}^{-1}(T_{1n} - \hat{\mu}_{10}) \xrightarrow{d} N(0,1) \quad \textit{and} \quad \hat{\sigma}_{20}^{-1}(T_{2n} - \hat{\mu}_{20}) \xrightarrow{d} N(0,1).$$

Based on Theorem 4, for a given significance level α, the rejection regions of the tests based on T_{1n} and T_{2n} can be derived as follows:

$$\big\{\mathbf{x}_1, \ldots, \mathbf{x}_n : \hat{\sigma}_{10}^{-1}|T_{1n} - \hat{\mu}_{10}| > q_{1-\alpha/2}\big\},$$

$$\big\{\mathbf{x}_1, \ldots, \mathbf{x}_n : \hat{\sigma}_{20}^{-1}|T_{2n} - \hat{\mu}_{20}| > q_{1-\alpha/2}\big\},$$

where $q_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

As a special case, it is important to note that testing $H_{01} : \mathbf{R} = \mathbf{R}_*$ with $\mathbf{R}_*$ prespecified is equivalent to testing H_0 by fixing $K = 0$ and $\mathbf{J}_0 = \mathbf{R}_*$ in (1.1). This leads to $\mathbf{R}_0 = \widehat{\mathbf{R}}_0 = \mathbf{J}_0$, $\mathbf{B} = \mathbf{0}_{p\times p}$ and $\mathbf{C}_1 = \mathbf{J}_0^{-1}$, where $\mathbf{0}_{p\times p}$ denotes the zero matrix. Therefore, the test statistic $T_{1n} = \mathrm{tr}[(\widehat{\mathbf{R}}_n\mathbf{J}_0^{-1} - \mathbf{I}_p)^2]$ becomes T_Y in Yin et al. (2022), and $T_{2n} = \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{J}_0)^2]$ is identical to T_Z in Zheng et al. (2019). As a result, the limiting null distributions of T_{1n} and T_{2n} can be simplified as follows.

Corollary 1. *Under the conditions of Theorem 4 and $H_{01} : \mathbf{R} = \mathbf{J}_0$, it holds that*

$$\tilde{\sigma}_{10}^{-1}(T_{1n} - \tilde{\mu}_{10}) \xrightarrow{d} N(0,1) \quad \text{and} \quad \tilde{\sigma}_{20}^{-1}(T_{2n} - \tilde{\mu}_{20}) \xrightarrow{d} N(0,1),$$

where the expressions for $\tilde{\mu}_{10}$, $\tilde{\mu}_{20}$, $\tilde{\sigma}_{10}^2$, $\tilde{\sigma}_{20}^2$ are

$$\begin{aligned}
\tilde{\mu}_{10} &= py_n - 3y_n^2 - 3y_n + \hat{\beta}_w y_n - 0.5n^{-1}\mathrm{tr}(\widehat{\mathbf{C}}_0\mathbf{J}_0^{-1}) \\
&\quad -(2y_n+2)\hat{\beta}_w n^{-1}\sum_{k=1}^p\sum_{\ell=1}^p(\mathbf{e}_k^T\widehat{\boldsymbol{\Gamma}}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{J}_0^{-1}\widehat{\boldsymbol{\Gamma}}\mathbf{e}_\ell \\
&\quad +(1.5y_n+0.5)n^{-1}\Big[2p + \hat{\beta}_w\sum_{k=1}^p\sum_{\ell=1}^p(\mathbf{e}_k^T\widehat{\boldsymbol{\Gamma}}\mathbf{e}_\ell)^4\Big] \\
&\quad +(0.5y_n+1)n^{-1}\sum_{i=1}^p\sum_{j=1}^p\mathbf{e}_i^T\mathbf{J}_0\mathbf{e}_j\mathbf{e}_i^T\mathbf{J}_0^{-1}\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{J}_0\mathbf{e}_j)^2 + \hat{\beta}_w\sum_{k=1}^p(\mathbf{e}_i^T\widehat{\boldsymbol{\Gamma}}\mathbf{e}_k)^2(\mathbf{e}_j^T\widehat{\boldsymbol{\Gamma}}\mathbf{e}_k)^2\Big], \\
\tilde{\mu}_{20} &= py_n + y_n^2 + n^{-1}\mathrm{tr}(\mathbf{J}_0^2) + \hat{\beta}_w n^{-1}\sum_{k=1}^p(\mathbf{e}_k^T\widehat{\boldsymbol{\Gamma}}^T\widehat{\boldsymbol{\Gamma}}\mathbf{e}_k)^2 - 0.5n^{-1}\mathrm{tr}(\widehat{\mathbf{C}}_0\mathbf{J}_0)
\end{aligned}$$

$$
\begin{aligned}
&-2n^{-1}\Big[2\mathrm{tr}(\mathbf{J}_0^2)+\hat{\beta}_w\sum_{k=1}^p\sum_{\ell=1}^p(\mathbf{e}_k^T\widehat{\mathbf{\Gamma}}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{J}_0\widehat{\mathbf{\Gamma}}\mathbf{e}_\ell\Big]\\
&+0.5n^{-1}\Big[2\mathrm{tr}(\mathbf{J}_0^2)+\hat{\beta}_w\sum_{k=1}^p\mathbf{e}_k^T\mathbf{J}_0^2\mathbf{e}_k\sum_{\ell=1}^p(\mathbf{e}_k^T\widehat{\mathbf{\Gamma}}\mathbf{e}_\ell)^4\Big]\\
&+n^{-1}\sum_{i=1}^p\sum_{j=1}^p(\mathbf{e}_i^T\mathbf{J}_0\mathbf{e}_j)^2\Big[2(\mathbf{e}_i^T\mathbf{J}_0\mathbf{e}_j)^2+\hat{\beta}_w\sum_{k=1}^p(\mathbf{e}_i^T\widehat{\mathbf{\Gamma}}\mathbf{e}_k)^2(\mathbf{e}_j^T\widehat{\mathbf{\Gamma}}\mathbf{e}_k)^2\Big],
\end{aligned}
$$

$$
\tilde{\sigma}_{10}^2 = 4y_n^2-4(2+\hat{\beta}_w)y_n^3+4y_n^2n^{-1}\Big[2\mathrm{tr}(\mathbf{J}_0^2)+\hat{\beta}_w\sum_{k=1}^p(\mathbf{e}_k^T\widehat{\mathbf{\Gamma}}^T\widehat{\mathbf{\Gamma}}\mathbf{e}_k)^2\Big],
$$

and $\tilde{\sigma}_{20}^2 = 4[n^{-1}\mathrm{tr}(\mathbf{J}_0^2)]^2$.

Remark 5. Although in this special case, the statistics T_{1n} and T_{2n} have the same form as T_Y and T_Z, their CLTs are slightly different. This is because both Yin et al. (2022) and Zheng et al. (2019) assume that $\mathbf{x}_k = [\mathrm{diag}(\mathbf{\Sigma})]^{1/2}\mathbf{R}^{1/2}\mathbf{w}_k + \boldsymbol{\mu}$, which is different from Assumption A. Nevertheless, we can still refer to the simulation studies in Yin et al. (2022) and Zheng et al. (2019) to gain insights into the performance of the tests based on T_{1n} and T_{2n}.

3.3 Power analysis

The limiting distributions in Theorem 2 can also be used to analyze the power functions of the tests based on T_{1n} and T_{2n}, denoted as $g_{1n}(\mathbf{R})$ and $g_{2n}(\mathbf{R})$, respectively, and given by

$$
g_{1n}(\mathbf{R}) = P\left(\hat{\sigma}_{10}^{-1}|T_{1n}-\hat{\mu}_{10}| > q_{1-\alpha/2}\right),
$$

$$
g_{2n}(\mathbf{R}) = P\left(\hat{\sigma}_{20}^{-1}|T_{2n}-\hat{\mu}_{20}| > q_{1-\alpha/2}\right).
$$

Let $\mathbf{\Delta}_1 = \mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p$, that is, $\mathbf{\Delta}_1$ represents the difference between $\mathbf{R}\mathbf{R}_P^{-1}$ and $\mathbf{I}_p$. When $\beta_w = 0$ and $\|\mathbf{\Delta}_1\| = o(1)$, under the conditions of Theorem 2, the differences between the corresponding constant order terms in the expressions for $\hat{\mu}_{10}$ and μ_{1n} as well as those between $\hat{\sigma}_{10}^2$ and σ_{1n}^2 are $o_p(1)$, then we have

$$
\mu_{1n} - \hat{\mu}_{10} = \mathrm{tr}(\mathbf{\Delta}_1^2) + 2y_n\mathrm{tr}\mathbf{\Delta}_1 + o_p(1), \quad \sigma_{1n} - \hat{\sigma}_{10} = o_p(1).
$$

Similarly, denote the difference between $\mathbf{R}$ and $\mathbf{R}_P$ as $\boldsymbol{\Delta}_2$, that is, $\boldsymbol{\Delta}_2 = \mathbf{R} - \mathbf{R}_P$. When $\beta_w = 0$ and $\|\boldsymbol{\Delta}_2\| = o(1)$, under the conditions of Theorem 2, we have

$$\mu_{2n} - \hat{\mu}_{20} = \mathrm{tr}(\boldsymbol{\Delta}_2^2) + o_p(1), \quad \sigma_{2n} - \hat{\sigma}_{20} = o_p(1).$$

If the limits of σ_{1n}, σ_{2n}, $\mathrm{tr}(\boldsymbol{\Delta}_1^2) + 2y_n\mathrm{tr}\boldsymbol{\Delta}_1$, and $\mathrm{tr}(\boldsymbol{\Delta}_2^2)$ exist, denoted as σ_1, σ_2, d_1, and d_2, respectively, then by Slutsky's theorem, we have

$$\hat{\sigma}_{10}^{-1}(T_{1n} - \hat{\mu}_{10}) \xrightarrow{d} N(d_1/\sigma_1, 1) \quad \text{and} \quad \hat{\sigma}_{20}^{-1}(T_{2n} - \hat{\mu}_{20}) \xrightarrow{d} N(d_2/\sigma_2, 1).$$

Corollary 2. *Under the conditions of Theorem 2, and given that the kurtosis satisfies* $\beta_w = 0$ *and the limits of* σ_{1n} *and* σ_{2n} *exist, we have*

(a) if $\|\boldsymbol{\Delta}_1\| = o(1)$ *and* $\lim_{n\to\infty} [\mathrm{tr}(\boldsymbol{\Delta}_1^2) + 2y_n\mathrm{tr}\boldsymbol{\Delta}_1] = d_1 \neq 0$*, then* $\lim_{n\to\infty} g_{1n}(\mathbf{R}) > \alpha$*;*

(b) if $\|\boldsymbol{\Delta}_2\| = o(1)$ *and* $\lim_{n\to\infty} \mathrm{tr}(\boldsymbol{\Delta}_2^2) = d_2 \neq 0$*, then* $\lim_{n\to\infty} g_{2n}(\mathbf{R}) > \alpha$*.*

Corollary 2 indicates that the tests based on T_{1n} and T_{2n} are asymptotically unbiased under the local alternatives.

4. Extreme-value test based on the infinite norm

Since T_{1n} (or T_{2n}) is constructed based on the quadratic norm, the corresponding test generally possesses high power when there are many small differences between $\mathbf{R}$ and $\mathbf{R}_0$, i.e., when $\mathbf{R}\mathbf{R}_0^{-1} - \mathbf{I}_p$ (or $\mathbf{R} - \mathbf{R}_0$) is dense. To detect large disturbances when $\mathbf{R} - \mathbf{R}_0$ is sparse, we consider the extreme-value statistic,

$$M_n = \max_{1\le i<j\le p} |\hat{r}_{ij} - \hat{r}_{0ij}| \,.$$

Before establishing the limiting null distribution of M_n, we first provide some intuition. Under some suitable moment conditions, we have $\max_{1\le i\le j\le p} |\hat{\sigma}_{ij} - \sigma_{ij}| = O_p\left(\sqrt{\log p/n}\right)$. Based on the first order Taylor expansion of the 3-variate function $x(yz)^{-1/2}$ for $x \in \mathbb{R}$ and $y, z > 0$ (see (5) in Cai and Zhang (2016))

$$\frac{\hat{x}}{(\hat{y}\hat{z})^{1/2}} = \frac{x}{(yz)^{1/2}} + \frac{\hat{x}-x}{(yz)^{1/2}} - \frac{x}{(yz)^{1/2}}\left(\frac{\hat{y}-y}{2y} + \frac{\hat{z}-z}{2z}\right) + o(\hat{x}-x) + o(\hat{y}-y) + o(\hat{z}-z), \quad (4.14)$$

we have

$$\begin{aligned}\hat{r}_{ij} &= \frac{\hat{\sigma}_{ij}}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{\sigma_{ij}}{(\sigma_{ii}\sigma_{jj})^{1/2}}\left(\frac{\hat{\sigma}_{ii}-\sigma_{ii}}{2\sigma_{ii}} + \frac{\hat{\sigma}_{jj}-\sigma_{jj}}{2\sigma_{jj}}\right) + o_p(n^{-1/2}) \\ &= r_{ij} + \frac{\hat{\sigma}_{ij}}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{ij}}{2}\left(\frac{\hat{\sigma}_{ii}}{\sigma_{ii}} + \frac{\hat{\sigma}_{jj}}{\sigma_{jj}}\right) + o_p(n^{-1/2}).\end{aligned}$$

It follows from (2.4) that

$$\hat{r}_{ij} - r_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left[\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{ij}}{2}\left(\frac{(x_{ik}-\bar{x}_i)^2}{\sigma_{ii}} + \frac{(x_{jk}-\bar{x}_j)^2}{\sigma_{jj}}\right)\right] + o_p(n^{-1/2}).$$

Note that $(\hat{r}_{ij} - r_{ij})$'s are in general on different scales, so we consider the standardized version $\sqrt{n}(\hat{r}_{ij} - r_{ij})/\sqrt{\eta_{ij}}$, where

$$\eta_{ij} = \mathrm{Var}\left(\frac{(x_{i1}-\mu_i)(x_{j1}-\mu_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{ij}}{2}\left(\frac{(x_{i1}-\mu_i)^2}{\sigma_{ii}} + \frac{(x_{j1}-\mu_j)^2}{\sigma_{jj}}\right)\right).$$

In practice, η_{ij}'s are usually unknown, and can be estimated by

$$\hat{\eta}_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left[\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\hat{\sigma}_{ii}\hat{\sigma}_{jj})^{1/2}} - \frac{\hat{r}_{ij}}{2}\left(\frac{(x_{ik}-\bar{x}_i)^2}{\hat{\sigma}_{ii}} + \frac{(x_{jk}-\bar{x}_j)^2}{\hat{\sigma}_{jj}}\right)\right]^2.$$

Thus, we consider the standardized statistic

$$\widetilde{M}_n = \max_{1\le i<j\le p}\frac{n(\hat{r}_{ij}-r_{ij})^2}{\hat{\eta}_{ij}}.$$

Intuitively, $n(\hat{r}_{ij}-r_{ij})^2/\hat{\eta}_{ij}$ are approximately square of standard normal variables and weakly dependent under some suitable conditions. The statistic $\widetilde{M}_n$ is the maximum of $p(p-1)/2$ such variables. Next, we show that $\widetilde{M}_n - 4\log p + \log\log p$ converges to the Type I extreme-value distribution under certain regularity conditions. Based on this result, we subsequently derive the limiting null distribution of the test statistic M_n.

Let the index set be $\mathcal{I}_n := \{(i,j) : 1 \le i < j \le p\}$, for $\alpha = (i,j)$, $\beta = (k,\ell) \in \mathcal{I}_n$, define

$$\begin{aligned}
W_{ij} &= \frac{(x_{i1}-\mu_i)(x_{j1}-\mu_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}}, \\
V_{ij} &= V_\alpha := W_{ij} - \frac{r_{ij}}{2}(W_{ii}+W_{jj}), \\
\gamma_n &= \sup_{\alpha,\beta\in\mathcal{I}_n \text{ and } \alpha\neq\beta} |\mathrm{Cor}(V_\alpha, V_\beta)|, \\
\gamma_n(b) &= \sup_{\alpha\in\mathcal{I}_n} \sup_{\mathcal{A}\subset\mathcal{I}_n, |\mathcal{A}|=b} \inf_{\beta\in\mathcal{A}} |\mathrm{Cor}(V_\alpha, V_\beta)|.
\end{aligned}$$

We consider the following sparse settings and moment conditions.

- **Assumption E.** For any sequence $\{b_n\}$ such that $b_n \to \infty$, $\gamma_n(b_n)\log(b_n) = o(1)$, and $\limsup_{n\to\infty} \gamma_n < 1$.

- **Assumption E*.** For any sequence $\{b_n\}$ such that $b_n \to \infty$, $\gamma_n(b_n) = o(1)$, and for some constant $\epsilon > 0$,

$$\sum_{\alpha,\beta\in\mathcal{I}_n} \big[\mathrm{Cov}(V_\alpha, V_\beta)\big]^2 = O(p^{4-\epsilon}).$$

  In addition, $\limsup_{n\to\infty} \gamma_n < 1$.

- **Assumption F.** Assume that $\log p = o(n^{1/5})$, there exist constants $\eta > 0$ and $K > 0$ satisfying the following moment conditions:

$$\mathrm{E}[\exp(\eta(x_{i1}-\mu_i)^2/\sigma_{ii})] \le K \quad \text{for } i = 1,\ldots,p.$$

  Furthermore, assume that for some constant $\tau > 0$,

$$\min_{1\le i<j\le p} \eta_{ij} > \tau \quad \text{and} \quad \min_{1\le i\le j\le p} \vartheta_{ij} > \tau, \tag{4.15}$$

  where $\eta_{ij} = \mathrm{Var}(V_{ij})$ and $\vartheta_{ij} = \mathrm{Var}(W_{ij})$ respectively.

- **Assumption F*.** Assume that for some constants $\gamma_0 > 0$, $c_1 > 0$, $p \leq c_1 n^{\gamma_0}$, and for some constants $\epsilon > 0$, $K > 0$, the following condition holds,

$$\mathrm{E}|(x_{i1} - \mu_i)/\sigma_{ii}^{1/2}|^{4\gamma_0+4+\epsilon} \leq K \quad \text{for } i = 1, \ldots, p.$$

  Furthermore, there exists some constant $\tau > 0$ that makes (4.15) hold.

Assumptions E (or E*) is similar to (A3) (or (A3′)) in Xiao and Wu (2013) that requires that the dependence among $\{V_\alpha, \alpha \in \mathcal{I}_n\}$ are not too strong. The condition $\limsup_{n\to\infty} \gamma_n < 1$ excludes the case that there may be many pairs $(\alpha, \beta) \in \mathcal{I}_n$ such that V_α and V_β are completely linear correlated. Assumptions F and F* are proposed in Cai et al. (2013), which indicate that the growth speed of p relative to n is exponential or polynomial for the distributions with sub-gaussian-type or polynomial-type tails respectively. In addition, the minimum values of η_{ij} and ϑ_{ij} greater than τ implies that V_{ij} and W_{ij} are not constants.

Theorem 5. *Under Assumptions E(or E*)-F(or F*), we have for any $t \in \mathbb{R}$,*

$$P(\widetilde{M}_n - 4\log p + \log\log p \leq t) \to \exp\left(-\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{t}{2}\right)\right).$$

The detailed proof of Theorem 5 is deferred to the Supplementary Material. The main idea of the proof is similar to the method established in Cai et al. (2013). Specifically, we divide the proof into two steps. In the first step we show that plugging in the estimated mean and variance parameters doesn't change the limiting distribution. The second one is a truncation step, we prove Theorem 5 under the assumption that all the involved mean and variance parameters are known.

Step 1: Effects of estimated variances and means. Let

$$\widetilde{\mathbf{\Sigma}}_n = (\tilde{\sigma}_{ij})_{i,j=1}^p = n^{-1} \sum_{k=1}^n (\mathbf{x}_k - \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu})^T,$$

and $\widetilde{\mathbf{R}}_n = (\tilde{r}_{ij})_{i,j=1}^p$ be the sample correlation matrix corresponding to $\widetilde{\mathbf{\Sigma}}_n$. Denote

$$\widetilde{M}_{n,1} = \max_{1\le i<j\le p} \frac{n(\hat{r}_{ij} - r_{ij})^2}{\eta_{ij}} \quad \text{and} \quad \widetilde{M}_{n,2} = \max_{1\le i<j\le p} \frac{n(\tilde{r}_{ij} - r_{ij})^2}{\eta_{ij}}.$$

In this step, we prove that both $|\widetilde{M}_n - \widetilde{M}_{n,1}|$ and $|\widetilde{M}_{n,1} - \widetilde{M}_{n,2}|$ converge to 0 in probability.

Step 2: Truncation. From (4.14), we can get $\widetilde{M}_{n,2} = \max_{\alpha\in\mathcal{I}_n} Q_\alpha^2 + o_p(1)$, where

$$Q_\alpha = \frac{1}{\sqrt{n\eta_{ij}}} \sum_{k=1}^n V_{k\alpha},$$

and

$$V_{k\alpha} = \frac{(x_{ik}-\mu_i)(x_{jk}-\mu_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{ij}}{2}\left(\frac{(x_{ik}-\mu_i)^2}{\sigma_{ii}} + \frac{(x_{jk}-\mu_j)^2}{\sigma_{jj}}\right).$$

Let $\hat{V}_{k\alpha} = V_{k\alpha} I\{|V_{k\alpha}| \le \tau_n\} - \mathrm{E} V_{k\alpha} I\{|V_{k\alpha}| \le \tau_n\}$, where $\tau_n = \eta^{-1} 8\log(p+n)$ if assumption F holds and $\tau_n = \sqrt{n}/(\log p)^8$ if assumption F* holds. Denote

$$\hat{Q}_\alpha = \frac{1}{\sqrt{n\eta_{ij}}} \sum_{k=1}^n \hat{V}_{k\alpha}.$$

In this step, we prove that $\left|\max_{\alpha\in\mathcal{I}_n} Q_\alpha^2 - \max_{\alpha\in\mathcal{I}_n} \hat{Q}_\alpha^2\right| = o_p(1)$, and for any $t \in \mathbb{R}$,

$$P\left(\max_{\alpha\in\mathcal{I}_n} \hat{Q}_\alpha^2 - 4\log p + \log\log p \le t\right) \to \exp\left(-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right).$$

Then based on the Slutsky's theorem, we prove that Theorem 5 holds.

Note that under the null hypothesis H_0,

$$\eta_{ij} = \mathrm{Var}\left(\frac{(x_{i1}-\mu_i)(x_{j1}-\mu_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{0ij}}{2}\left(\frac{(x_{i1}-\mu_i)^2}{\sigma_{ii}} + \frac{(x_{j1}-\mu_j)^2}{\sigma_{jj}}\right)\right),$$

where r_{0ij} denotes the (i,j)th element of $\mathbf{R}_0$. Then under H_0, η_{ij} can be estimated by

$$\hat{\eta}_{0ij} = n^{-1}\sum_{k=1}^n \left\{\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\hat{\sigma}_{ii}\hat{\sigma}_{jj})^{1/2}} - \hat{r}_{ij} - \frac{\hat{r}_{0ij}}{2}\left[\frac{(x_{ik}-\bar{x}_i)^2}{\hat{\sigma}_{ii}} + \frac{(x_{jk}-\bar{x}_j)^2}{\hat{\sigma}_{jj}} - 2\right]\right\}^2.$$

Thus, we consider the test statistic

$$M_n = \max_{1 \le i < j \le p} \frac{n(\hat{r}_{ij} - \hat{r}_{0ij})^2}{\hat{\eta}_{0ij}}.$$

The following Corollary 3 gives the limiting null distribution of M_n.

Corollary 3. *Under Assumptions A-B-C-D-E(or E*)-F(or F*) and the null hypothesis* H_0*, for any* $t \in \mathbb{R}$*,*

$$P(M_n - 4\log p + \log\log p \le t) \to \exp\left(-\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{t}{2}\right)\right).$$

The proof of Corollary 3 is given in the Supplementary Material. According to Corollary 3, M_n converges to the Type I extreme-value distribution under the null hypothesis H_0. Thus, the rejection region of the test based on M_n can be expressed as

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n : M_n > t_\alpha + 4\log p - \log\log p\},$$

where $t_\alpha = -\log(8\pi) - 2\log\log(1-\alpha)^{-1}$ is the $(1-\alpha)$th quantile of the Type I extreme-value distribution.

5. Simulation studies

We conduct simulation studies to evaluate the finite-sample performance of the proposed testing procedures under various scenarios. Specifically, we generate $n = 100$ or 300 i.i.d. observations from $\mathbf{x}_k = \mathbf{\Sigma}^{1/2}\mathbf{w}_k$ for $k = 1, \ldots, n$, where $\mathbf{w}_k = (w_{1k}, \ldots, w_{pk})^T$ is a random vector from (a) the Gaussian $N(0,1)$ population or (b) the Gamma$(4,2) - 2$ population. Throughout the simulation studies, the jth element of the diagonal vector of the population covariance matrix $\mathbf{\Sigma}$ is fixed as $\sqrt{1 + 2j/p}$, $j = 1, \ldots, p$. The dimension p is taken to be 50, 100, 300, 500, or 1000. We configure eight scenarios for the population correlation matrix $\mathbf{R}$.

- Scenarios 1–3 are designed for testing $H_0 : \mathbf{R} = \mathbf{J}_0$ with $\mathbf{J}_0 = (a_{0ij})_{i,j=1}^p$ being a prespecified correlation matrix. To evaluate the power of the proposed tests, we take $\mathbf{R} = \mathbf{J}_0 + (b_{ij})_{i,j=1}^p$ under the alternative hypothesis H_1. The values of $(a_{0ij})_{i,j=1}^p$ and $(b_{ij})_{i,j=1}^p$ are given below.

  - Scenario 1: $a_{0ij} = \delta_{\{i=j\}} + 0.4\delta_{\{|i-j|=1\}}$ and $b_{ij} = 0.1\delta_{\{|i-j|=2\}}$, such that $\mathbf{R}$ is a banded matrix with a bandwidth of one under H_0 and two under H_1, respectively.

  - Scenario 2: $a_{0ij} = \delta_{\{i=j\}} + 0.4\delta_{\{|i-j|=1\}} + 0.3\delta_{\{|i-j|=2\}} + 0.2\delta_{\{|i-j|=3\}}$ and $b_{ij} = 0.04\delta_{\{i\neq j\}}$, i.e., $\mathbf{R}$ is a banded matrix under H_0, whereas $\mathbf{R}$ under H_1 is a combination of a banded matrix and a compound symmetric matrix.

  - Scenario 3: $a_{ij} = 0.3^{|i-j|}$, $b_{1j} = b_{j1} = 0.4^{\sqrt{j-1}}$ for $2 \leq j \leq (p/25 + 1)$, $b_{2j} = b_{j2} = 0.3^{\sqrt{j-2}}$ for $3 \leq j \leq (p/25 + 2)$, and $b_{ij} = 0$ elsewhere. By this configuration, $\mathbf{R}$ under H_0 is an autoregressive correlation matrix from the $AR(1)$ model. Under H_1, $\mathbf{R}$ deviates from the $AR(1)$ matrix in $4p/25$ entries, corresponding to a sparse alternative.

- Scenarios 4–6 test $H_0 : \mathbf{R} = \mathbf{I}_p + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2$ with $\mathbf{J}_k = (a_{kij})_{i,j=1}^p, k = 1, 2$, leading to a three-component population correlation matrix. For power analysis, we fix the alternative hypothesis as $H_1 : \mathbf{R} = \mathbf{I}_p + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2 + (b_{ij})_{i,j=1}^p$.

  - Scenario 4: $a_{1ij} = \delta_{\{|i-j|=1\}}$, $a_{2ij} = \delta_{\{|i-j|=2\}}$, and $b_{ij} = 0.1\delta_{\{|i-j|=3\}}$, $\theta_1 = \theta_2 = 0.1$. The null hypothesis tests whether the bandwidth of $\mathbf{R}$ exceeds two.

  - Scenario 5: $a_{1ij} = 0.4^{|i-j|}\delta_{\{i\neq j\}}$, $a_{2ij} = \delta_{\{|i-j|=2\}}$, $b_{12} = b_{21} = b_{13} = b_{31} = 0.5$, and $b_{ij} = 0.025\delta_{\{i\neq j\}}$ elsewhere. Under H_0, $\theta_1 = 1$ and $\theta_2 = 0.1$, thus $\mathbf{R}$ is a combination of an $AR(1)$ matrix and a banded matrix. Under H_1, $\theta_1 = \theta_2 = 0.1$, the alternative hypothesis has a mixture of sparse and dense signals.

– Scenario 6: $a_{1ij} = (-0.5)^{|i-j|}\delta_{\{i\neq j\}}$, $a_{2ij} = 0.5^{|i-j|}\delta_{\{i\neq j\}}$, and $\theta_1 = \theta_2 = 0.1$, i.e., $\mathbf{R}$ is the weighted average of two $AR(1)$ matrices under H_0. Under H_1, we take $b_{ij} = 0.125\delta_{\{|i-j|=3\}}$ and $b_{12} = b_{21} = 0.25\sqrt{\log p}$, making $\mathbf{R}$ a combination of two $AR(1)$ matrices, a banded matrix, and a highly sparse matrix.

- Scenarios 7 and 8 examines the performance of the proposed tests under the linear structure of five components by specifying $H_0 : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2 + \theta_3\mathbf{J}_3 + \theta_4\mathbf{J}_4$ and $H_1 : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2 + \theta_3\mathbf{J}_3 + \theta_4\mathbf{J}_4 + (b_{ij})_{i,j=1}^p$, with $\mathbf{J}_k = (a_{kij})_{i,j=1}^p, k = 0, \dots, 4$.

  – Scenario 7: $a_{0ij} = 0.25^{|i-j|}$, $a_{kij} = \delta_{\{|i-j|=k\}}$ for $k = 1, 2, 3, 4$, $b_{ij} = 0.05\delta_{\{|i-j|=5\}}$, and $\theta_1 = \theta_2 = 0.2$, and $\theta_3 = \theta_4 = 0.1$. Thus, both the correlation matrices under H_0 and H_1 have the banded structure but with different bandwidths.

  – Scenario 8: $a_{0ij} = 0.25^{|i-j|}$, $a_{1ij} = 0.3\delta_{\{|i-j|=1\}}$, $a_{2ij} = 0.2\delta_{\{i\neq j\}}$, $a_{3ij} = 0.2^{1/|i-j|}\delta_{\{1\le|i-j|\le2\}}$, $a_{4ij} = 0.1^{1/|i-j|}\delta_{\{3\le|i-j|\le5\}}$, $b_{12} = b_{21} = 0.15\sqrt{\log p}$, $b_{13} = b_{31} = 0.7/\sqrt{\log p}$, and $b_{ij} = 0.05^{|i-j|}\delta_{\{|i-j|=3\}}$ elsewhere. Under H_0, $\theta_k = k/10, k = 1, 2, 3, 4$, and under H_1, $\theta_1 = 0.1$, and $\theta_2 = \theta_3 = \theta_4 = 0.2$. As a result, there are a mixture of sparse and dense differences between the correlation matrices specified under H_0 and H_1. We also note that this scenario corresponds to an unbounded spectral norm case where the maximum eigenvalue of $\mathbf{R}$ tends to infinity as the dimension p increases.

The simulation results summarized based on 10,000 replications under the Gaussian population are provided in Table 1 (for scenarios 1–4) and Table 2 (for scenarios 5–8). Those under the Gamma population are given in Tables S2-S3 of the Supplementary Material.

Overall, the simulation results provide sufficient empirical evidence about the consistency of the proposed tests. According to the simulation results for scenarios 1–7 where the maximum eigenvalue of $\mathbf{R}$ is finite, we find that the sizes of our three tests are particularly close to the nominal level of 0.05, including both low and high dimensions as well as Gaussian and non-Gaussian populations. When the spectral norm of $\mathbf{R}$ is unbounded under scenario 8, the difference-based test T_{2n} breaks down as its size deviates significantly from the level of 0.05. By contrast, the other two tests, including M_n and T_{1n}, can still preserve reasonable type I error rates. Under H_1, the empirical power of each of the three new tests increases as the sample size n increases. As expected, the extreme-value test M_n possesses better power under the sparse alternative such as in scenarios 3, 5, and 8. On the other hand, when there are many small differences between the two matrices specified under H_0 and H_1, the two tests T_{1n} and T_{2n}, constructed based on the quadratic norm, are able to detect the dense signals with high power.

6. Application

We investigate the correlation structure of Canadian weather data, which include daily temperature and precipitation (in log10 scale) at 35 different weather stations in Canada averaged over the period from 1960 to 1994. The dataset can be obtained from the `fda` package on `https://cran.r-project.org/`. We denote the precipitation and temperature on day j (a total of 365 days) at the ith station by y_{ij} and v_{ij}, respectively, and use $z_i \in \{\text{Atlantic, Pacific, Continental, Arctic}\}$ to indicate the climate zone of the ith station. The relationship between temperature and precipitation can be quantified by the functional linear model and the Fourier series expansion (Ramsay and Silverman, 2002; Zhong et al.,

Table 1: Empirical size and power for the tests based on the infinite norm (M_n), ratio-based quadratic norm (T_{1n}) and difference-based quadratic norm (T_{2n}) under scenarios 1–4, where n observations with dimension p are generated from the Gaussian population.

		Scenario 1			Scenario 2			Scenario 3			Scenario 4		
n	p	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}
		Empirical size (%)											
100	50	4.6	5.9	4.3	4.8	5.7	3.9	4.7	5.3	4.4	4.4	5.0	4.7
	100	4.5	5.6	4.6	4.8	5.5	4.2	4.6	5.1	4.5	4.5	4.9	4.9
	300	4.5	5.6	4.8	4.6	5.9	4.5	4.8	5.2	4.7	4.3	4.5	4.4
	500	4.8	5.5	4.5	4.5	5.9	4.4	4.8	5.3	4.4	4.7	5.0	4.5
	1000	4.9	5.7	4.5	5.1	5.6	4.6	4.8	5.6	4.8	4.9	4.8	4.7
300	50	4.5	5.5	4.3	4.4	5.2	4.2	4.2	5.1	4.5	4.5	4.7	4.5
	100	4.2	5.4	4.7	4.1	5.2	4.7	4.3	5.0	4.6	4.3	4.7	5.0
	300	4.5	5.7	5.0	4.3	5.5	4.6	4.5	5.2	4.9	4.1	5.0	4.9
	500	4.0	5.3	4.8	4.1	5.6	4.6	4.0	4.6	5.0	4.0	5.1	5.3
	1000	4.1	5.5	4.5	4.6	5.0	4.8	4.5	5.0	4.8	4.9	5.4	5.2
		Empirical power (%)											
100	50	7.1	100.0	11.3	10.2	19.5	54.2	79.9	5.7	9.3	7.0	22.0	13.7
	100	6.2	100.0	11.8	11.6	62.7	89.8	65.3	5.2	6.6	6.2	28.8	15.6
	300	5.5	100.0	11.9	12.2	99.9	100.0	41.6	5.4	5.1	5.2	56.9	16.0
	500	5.5	100.0	12.0	12.8	100.0	100.0	31.1	5.8	4.7	5.3	80.0	16.0
	1000	4.9	100.0	12.1	13.8	100.0	100.0	22.1	6.0	4.7	5.3	99.5	16.7
300	50	15.4	100.0	30.2	17.4	7.9	90.9	100.0	14.8	21.7	14.7	56.9	40.8
	100	12.7	100.0	31.0	20.1	35.7	99.9	99.9	5.8	11.7	11.8	67.2	43.9
	300	8.6	100.0	31.4	22.9	98.8	100.0	99.1	5.6	5.9	9.2	88.8	45.7
	500	7.1	100.0	32.3	23.5	100.0	100.0	98.0	5.6	5.5	6.9	97.0	47.0
	1000	6.4	100.0	32.5	25.8	100.0	100.0	95.3	5.8	4.7	6.6	100.0	47.3

Table 2: Empirical size and power for the tests based on the infinite norm (M_n), ratio-based quadratic norm (T_{1n}) and difference-based quadratic norm (T_{2n}) under scenarios 5–8, where n observations with dimension p are generated from the Gaussian population.

		Scenario 5			Scenario 6			Scenario 7			Scenario 8		
n	p	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}
		Empirical size (%)											
100	50	4.7	5.1	3.9	4.5	4.8	4.9	4.7	5.2	3.7	4.2	6.7	2.6
	100	4.6	5.1	4.2	4.7	4.6	4.7	4.5	4.8	4.2	4.2	6.0	3.1
	300	4.5	4.9	4.6	4.6	4.4	4.4	4.7	4.8	4.6	4.5	5.5	3.4
	500	4.7	5.4	4.6	4.2	5.1	5.0	4.6	5.2	5.0	4.5	5.5	3.3
	1000	5.1	5.4	4.9	5.1	4.7	4.7	4.7	5.3	4.3	4.8	5.5	2.7
300	50	4.1	5.1	4.0	4.4	4.8	4.7	4.2	5.2	3.8	3.7	5.9	2.4
	100	4.4	4.9	5.0	4.6	5.0	5.1	4.5	5.1	4.6	4.0	5.8	3.4
	300	4.2	4.9	4.7	4.2	5.1	5.0	4.2	5.1	4.6	3.9	5.5	3.8
	500	4.1	5.0	4.8	4.1	5.1	5.1	4.1	5.1	4.7	4.0	5.3	3.5
	1000	4.7	5.2	4.7	4.4	5.3	5.5	4.4	5.2	4.7	3.4	5.2	3.1
		Empirical power (%)											
100	50	92.9	34.7	50.4	62.6	45.2	39.6	5.0	26.1	4.2	49.5	26.5	3.2
	100	83.9	47.2	76.3	64.1	42.4	35.9	4.8	44.6	4.7	38.7	24.9	3.1
	300	60.7	96.1	99.8	63.8	44.5	33.2	4.8	95.3	5.3	32.4	42.0	2.9
	500	49.4	99.9	100.0	64.1	47.4	32.1	5.0	99.9	5.0	31.1	61.3	2.5
	1000	36.5	100.0	100.0	62.5	58.9	32.3	4.9	100.0	5.0	29.9	92.0	2.2
300	50	100.0	83.6	94.5	99.6	94.4	91.2	5.3	45.9	5.1	98.1	60.4	5.2
	100	100.0	91.8	99.5	99.8	92.9	88.4	4.9	69.1	5.7	97.2	50.5	3.8
	300	100.0	100.0	100.0	99.9	92.5	85.5	4.6	99.3	5.9	96.9	66.0	2.8
	500	99.9	100.0	100.0	100.0	93.5	84.0	3.8	100.0	5.9	97.8	83.1	2.3
	1000	99.7	100.0	100.0	100.0	95.9	83.4	4.5	100.0	6.0	98.1	99.0	2.1

2017),

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i, \tag{6.16}$$

where $\mathbf{y}_i = (y_{i1}, \ldots, y_{i365})^T$, the design matrix $\mathbf{X}_i$ contains the information of climate zone effect and the Fourier bases of zone-adjusted temperature, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{365})^T$ are unknown coefficients, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{i365})^T$ are i.i.d. random residuals with mean zero and covariance $\boldsymbol{\Sigma}$. We assume that the variances of the random residuals are homogeneous, i.e., $\boldsymbol{\Sigma} = \sigma^2\mathbf{R}$, where $\mathbf{R} = (r_{ij})_{i,j=1}^{365}$ is the correlation matrix. More details about the functional model can be found in Zhong et al. (2017).

Under model (6.16), we are interested in predicting future precipitations using the best linear unbiased predictor (BLUP), which has a prediction accuracy that highly depends on the structure of the correlation (or covariance) matrix. Zhong et al. (2017) examined four types of covariance structures: the sphericity, moving average with lag one, AR(1), and compound symmetry. They found that the compound symmetry yields the most accurate forecasting performance.

Continuing the analysis of Zhong et al. (2017), we further investigate the correlation structure of the residuals $\boldsymbol{\epsilon}_i$ based on our proposed testing procedures. Specifically, let $\mathbf{J}_k = (a_{ijk})_{i,j=1}^{p}$, $k = 0, 1, 2, 3$, where $a_{ij0} = \delta_{\{i=j\}} + \hat{\rho}\delta_{\{i\neq j\}}$ with $\hat{\rho}$ obtained by equation (2.8), $a_{ij1} = \delta_{\{i\neq j\}}$, $a_{ij2} = -|i-j|\delta_{\{i\neq j\}}$, and $a_{ij3} = |i-j|^{0.05}\delta_{\{i\neq j\}}$, Three correlation structures are specified as follows:

$$H_{01} : \mathbf{R} = \mathbf{J}_0,$$

$$H_{02} : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2,$$

$$H_{03} : \mathbf{R} = \mathbf{J}_0 + \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_3.$$

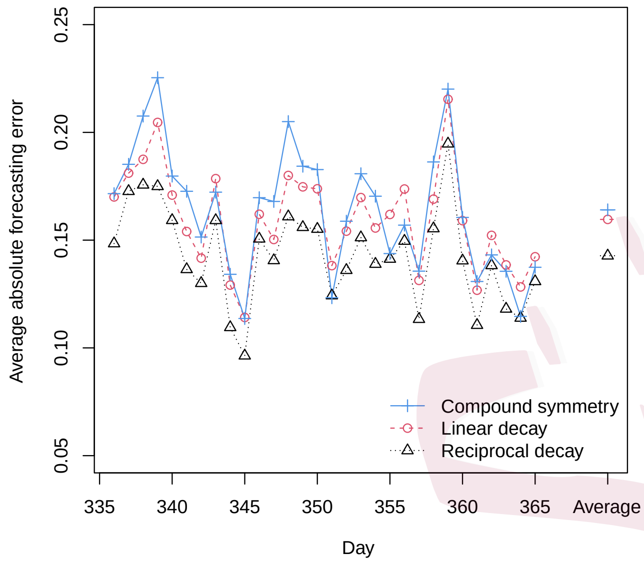
Here, H_{01} double-checks if the correlation matrix is compound symmetric, H_{20} tests whether the correlation matrix has an additional linear decay structure, and H_{03} examines whether a reciprocal decay structure exists in the correlation matrix.

Since the three null hypotheses specify unbounded spectral norm structures for $\mathbf{R}$, and also for illustrative purposes, we only consider the testing procedure based on T_{1n}. The resulting test statistics and the associated p-values (in parentheses) based on T_{1n} for assessing H_{01}–H_{03} are 17.7 (< 0.0001), 55.0 (< 0.0001), and 0.66 (0.51), respectively. As a result, the ratio-based T_{1n} test indicates that the compound symmetry is insufficient to characterize the weather data and the reciprocal decay structure might be more suitable.

To further validate the above conclusion, we randomly split the dataset into a training dataset of 30 stations and a validation dataset of 5 stations. Based on the first 335 observations of each training dataset, we then compare the out-of-sample forecasting performance based on the aforementioned three correlation structures for the last 30 days. According to 100 replications, we summarize the absolute prediction error and the standard error of the predicated precipitation in Figure 1. As expected, the prediction based on the reciprocal decay structure on average yields the smallest prediction bias and standard error.

7. Concluding remarks and combination tests

In this paper, we have developed several new testing procedures to examine the linear structure of the correlation matrix. Without imposing Gaussian assumptions on the random sample, our tests are applicable to various low-dimensional and high-dimensional cases. The asymptotic distributions of the proposed statistics possess explicit expressions, which can be further simplified under some special situations. As aforementioned and demonstrated in the



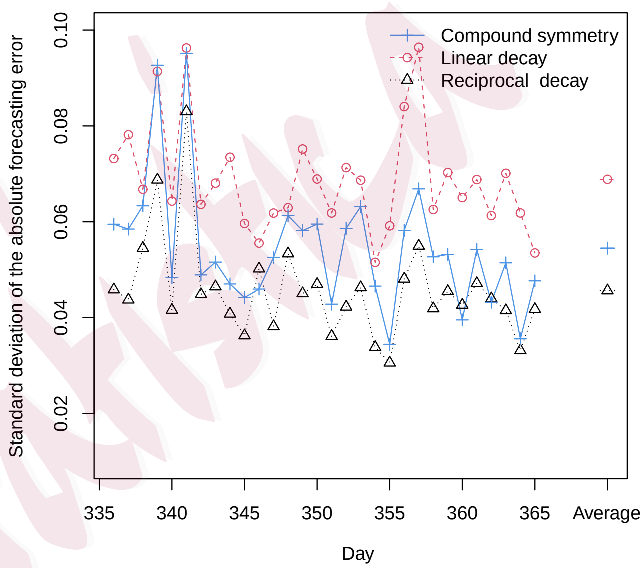


Figure 1: Average (top panel) and standard errors (bottom panel) of absolute prediction errors obtained from three structures of the correlation matrix of the Canadian weather data: compound symmetry, linear decay, and reciprocal decay.

numerical study, the tests based on M_n, T_{1n}, and T_{2n} deal with different scenarios, and they have high power in their respective target regions. It may occur that one test has inferior performance in the target region of another test. For example, when the differences between the population correlation matrix $\mathbf{R}$ and the structured correlation matrix $\mathbf{R}_0$ are sparse (as shown in scenario 3), it is difficult for T_{1n} and T_{2n} to detect such sparse signals. Similarly, when there are dense but very small disturbances between $\mathbf{R}$ and $\mathbf{R}_0$ (as shown in scenario 2), the test based on M_n may be underpowered. Even under the dense scenarios, the power of the ratio-based T_{1n} may be significantly different from that of the difference-based T_{2n}, as shown in scenarios 4 and 5.

In real applications, it is unknown *a priori* whether the differences between $\mathbf{R}$ and $\mathbf{R}_0$ are sparse and dense, it would be practical and desirable to propose a testing procedure that can possess robust performance under various scenarios. Let p_1, p_2, and p_3 be the p-values of the tests based on M_n, T_{1n}, and T_{2n}, respectively. We consider the following two combination tests to borrow strengths from the three testing procedures.

(1) Tippett's minimum p-value test T_{tn}: We reject the null hypothesis if $\min\{p_1, p_2, p_3\} < 1-(1-\alpha)^{1/3}$.

(2) Cauchy combination test T_{cn}: According to Liu and Xie (2020), define the Cauchy combination test statistic as $T_{cn} = \tan\{(0.5-p_1)\}/3+\tan\{(0.5-p_2)\}/3+\tan\{(0.5-p_3)\}/3$, and the null hypothesis is rejected if $T_{cn} > t_\alpha$, where t_α is the upper αth quantile of the standard Cauchy distribution.

We have performed an additional simulation study to investigate the finite-sample performance of T_{tn} and T_{cn}. The simulation results under the above eight scenarios are exhibited in Table 3. We find that the type I error rates of the two tests are close to 0.05. Due to the ability to integrate information across three difference measures, the combination tests generally yield strong power and are more powerful than the worst-performing individual test in all scenarios. Between T_{tn} and T_{cn}, we also observe that the empirical power of the latter test is slightly higher in most cases.

In theory, the testing size of T_{tn} as well as T_{cn} can be well controlled at or below the nominal level of α. A more sensible approach would be to construct a combination test using the limiting joint distribution of the three individual test statistics. However, as shown in Zou et al. (2021), which studies the two-sample tests for high-dimensional covariance matrices, the derivation of the correlation between T_{1n} and T_{2n} is non-trivial, and the finite-sample performance based on such a joint test is particularly similar to that of the Tippett's minimum p-value test. Moreover, the proposed test based on T_{1n} is applicable only when the inverse of $\widehat{\mathbf{R}}_0$ exists. In cases where $\widehat{\mathbf{R}}_0$ is non-invertible, the development of a ratio-based test statistic and its associated testing procedure is left as future research.

Supplementary Materials

In the supplement, we give the detailed proofs of Lemma 1, Theorems 1, 2, 3, 5, and Corollary 3. We also present the simulation results when the observations are generated from the Gamma population.

Acknowledgements

The authors thank the Co-Editor, Dr. Huixia Judy Wang, the Associate Editor, and three reviewers for their constructive and insightful comments that greatly improved this paper.

Table 3: Empirical size and power for the Tippett's minimum p-value test T_{tn} and the Cauchy combination test T_{cn} under scenarios 1–8, where n observations with dimension p are generated from the Gaussian population.

Scenario		1		2		3		4		5		6		7		8	
n	p	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}
		Empirical size (%)															
100	50	4.2	5.0	4.5	4.8	3.9	4.7	3.5	4.7	4.1	4.6	3.4	4.7	4.3	4.8	4.7	4.8
	100	4.8	5.2	4.6	4.9	4.2	4.9	3.6	4.6	4.4	4.9	3.2	4.6	4.3	4.6	4.4	4.5
	300	5.0	5.2	4.8	4.9	4.5	4.9	3.3	4.3	4.4	4.6	3.0	4.3	4.2	4.6	4.5	4.5
	500	4.9	5.3	4.8	5.0	4.7	4.8	3.6	4.7	4.4	4.7	3.1	4.7	4.8	5.2	4.2	4.2
	1000	5.2	5.3	5.3	5.3	4.9	5.1	3.7	4.7	4.6	4.7	3.4	4.8	4.7	4.8	4.3	4.1
300	50	4.4	5.0	4.4	4.6	3.9	4.6	3.2	4.4	4.0	4.5	3.0	4.5	4.1	4.5	3.8	4.0
	100	4.5	5.0	4.6	5.0	4.2	5.0	3.4	4.7	4.4	5.0	3.3	4.8	4.3	4.9	4.0	4.3
	300	4.6	5.1	4.8	5.0	4.4	5.0	3.5	4.7	4.3	4.8	3.3	4.8	4.3	4.8	4.1	4.0
	500	4.3	4.6	4.6	4.6	4.2	4.5	3.7	5.0	3.9	4.3	3.2	4.8	4.2	4.6	3.9	3.8
	1000	4.7	4.8	5.0	5.1	4.7	5.1	4.0	5.2	4.6	4.7	3.5	5.1	4.9	5.3	3.5	3.5
		Empirical power (%)															
100	50	100.0	100.0	52.7	54.9	68.0	67.8	14.7	17.9	89.1	90.9	61.8	67.4	18.1	18.4	44.8	46.0
	100	100.0	100.0	93.8	95.0	52.7	52.8	19.0	22.0	90.0	91.7	62.9	68.8	33.1	33.2	36.7	37.9
	300	100.0	100.0	100.0	100.0	30.8	31.3	41.3	42.4	99.8	99.9	65.0	69.9	91.0	90.8	43.8	45.1
	500	100.0	100.0	100.0	100.0	23.2	23.4	66.5	66.1	100.0	100.0	66.7	71.2	99.8	99.7	57.4	59.1
	1000	100.0	100.0	100.0	100.0	15.7	16.2	98.5	98.5	100.0	100.0	71.4	74.7	100.0	100.0	87.5	88.2
300	50	100.0	100.0	88.2	89.1	99.9	99.9	44.0	49.8	100.0	100.0	99.8	99.9	34.3	34.4	97.1	97.4
	100	100.0	100.0	100.0	100.0	99.6	99.5	53.9	58.2	100.0	100.0	99.9	99.9	56.3	56.0	95.7	96.0
	300	100.0	100.0	100.0	100.0	97.7	97.5	79.6	80.7	100.0	100.0	100.0	100.0	98.3	98.3	97.0	97.4
	500	100.0	100.0	100.0	100.0	96.1	95.8	93.4	93.2	100.0	100.0	100.0	100.0	100.0	100.0	98.7	98.8
	1000	100.0	100.0	100.0	100.0	91.6	91.4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.9	99.9

Shurong Zheng's research was supported by the National Key R&D Program of China (No. 2024YFA1012200), the National Natural Science Foundation of China (Nos. 12326606 and 12231011). Tingting Zou's research was supported by the National Natural Science Foundation of China (No. 12301339). Guangren Yang's research was supported by the National Social Science Fund of China grant (24BTJ070). Guo-Liang Tian's research was partially supported by the National Natural Science Foundation of China (No. 12171225).

References

Aitkin, M. A. (1969). Some tests for correlation matrices. *Biometrika 56*(2), 443–446.

Bartlett, M. S. and D. V. Rajalakshman (1953). Goodness of fit tests for simultaneous autoregressive series. *Journal of the Royal Statistical Society (Series B) 15*, 107–124.

Cai, T. and T. T. Jiang (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *The Annals of Statistics 39*, 1496–1525.

Cai, T. and T. T. Jiang (2012). Phase transition in limiting distributions of coherence of high-dimensional random matrices. *Journal of Multivariate Analysis 107*, 24–39.

Cai, T., W. Liu, and Y. Xia (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association 108*(501), 265–277.

Cai, T. T. and A. Zhang (2016). Inference for high-dimensional differential correlation matrices. *Journal of Multivariate Analysis 143*, 107–126.

Chen, J., X. Y. Wang, S. R. Zheng, B. S. Liu, and N.-Z. Shi (2020). Tests for high dimensional covariance matrices. *Random Matrices: Theory and Applications 9*(3), 1–25.

Gao, J. T., X. Han, G. M. Pan, and Y. R. Yang (2017). High-dimensional correlation matrices: the central limit theorem and its application. *Journal of the Royal Statistical Society (Series B) 79*(3), 677–693.

Jiang, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *The Annals of Applied Probability 14*(2), 865–880.

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika 43*(4), 443–477.

Kullback, S. (1967). On testing correlation matrices. *Journal of the Royal Statistical Society (Series C) 16*(1), 80–85.

Larzelere, R. E. and S. A. Mulaik (1977). Single-sample tests for many correlations. *Psychological Bulletin 84*(3), 557–569.

Leung, D. and M. Drton (2018). Testing independence in high dimensions with sums of rank correlations. *The Annals of Statistics 46*(1), 280–307.

Li, D., W. D. Liu, and A. Rosalsky (2010). Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probability Theory and Related Fields 148*, 5–35.

Li, D., Y. Qi, and A. Rosalsky (2012). On jiang's asymptotic distribution of the largest entry of a sample correlation matrix. *Journal of Multivariate Analysis 111*, 256–270.

Li, D. and A. Rosalsky (2006). Some strong limit theorems for the largest entries of sample correlation matrices. *The Annals of Applied Probability 16*, 423–447.

Liu, W. D., Z. Lin, and Q. M. Shao (2008). The asymptotic distribution and berry-esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *The Annals of Applied Probability 18*, 2337–2366.

Liu, Y. and J. Xie (2020). Cauchy combination test: a powerful test with analytic p-value calculation under arbitrary dependency structures. *Journal of the American Statistical Association 115*(529), 393–402.

McDonald, R. P. (1975). Testing pattern hypotheses for correlation matrices. *Psychometrika 40*(40), 253–255.

Mestre, X. and P. Vallet (2017). Correlation tests and linear spectral statistics of the sample correlation matrix. *IEEE Transactions on Information Theory 63*(7), 4585–4618.

Noureddine, E. K. (2009). Concentration of measure and spectra of random matrices: with applications to correlation matrices, elliptical distributions and beyond. *The Annals of Applied Probability 19*(6), 2362–2405.

Oldham, M. C., S. Horvath, and D. H. Geschwind (2006). Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences 103*(47), 17973–17978.

Opgen-Rhein, R. and K. Strimmer (2007). From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC systems biology 1*(1), 37.

Ramsay, J. O. and B. W. Silverman (2002). *Applied functional data analysis: methods and case studies*. Springer.

Schott, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika 92*(4), 951–956.

Steiger, J. H. (1980). Testing pattern hypotheses on correlation matrices: Alternative statistics and some empirical results. *Multivariate Behavioral Research 15*, 335–352.

Xiao, H. and W. Wu (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Processes and their Applications 123*, 2899–2920.

Yin, Y., C. Li, G.-L. Tian, and S. Zheng (2022). Spectral properties of rescaled sample correlation matrix. *Statistica Sinica 32*(4), 2007–2022.

Yin, Y. and Y. Ma (2022). Properties of eigenvalues and eigenvectors of large-dimensional sample correlation matrices. *The Annals of Applied Probability 32*(6), 4763–4802.

Yong, A. G. and S. Pearce (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology 9*(2), 79–94.

Zheng, S. R., Z. Chen, H. J. Cui, and R. Z. Li (2019). Hypothesis testing on linear structures of high-dimensional covariance matrix. *The Annals of Statistics 47*(6), 3300–3334.

Zheng, S. R., G. H. Cheng, J. H. Guo, and H. T. Zhu (2019). Test for high-dimensional correlation matrices. *The Annals of Statistics 47*(5), 2887–2921.

Zhong, P.-S., W. Lan, P. X. K. Song, and C.-L. Tsai (2017). Tests for covariance structures with high-dimensional repeated measurements. *The Annals of Statistics 45*(3), 1185–1213.

Zhou, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Transactions of the American Mathematical Society 359*(11), 5345–5363.

Zou, T., R. Lin, S. Zheng, and G.-L. Tian (2021). Two-sample tests for high-dimensional covariance matrices using both difference and ratio. *Electronic Journal of Statistics 15*(1), 135–210.

Tingting Zou

School of Mathematics, Jilin University, Changchun, China

E-mail: zoutt260@jlu.edu.cn

Guangren Yang

Department of Statistics, School of Economics, Jinan University, Guangzhou, China

E-mail: tygr@jnu.edu.cn

Ruitao Lin

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, USA

E-mail: rlin@mdanderson.org

Guo-Liang Tian

Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, China

E-mail: tiangl@sustech.edu.cn

Shurong Zheng

School of Mathematics and Statistics, Northeast Normal University, Changchun, China

E-mail: zhengsr@nenu.edu.cn