Statistica Sinica **32** (2022), 719-735
doi:https://doi.org/10.5705/ss.202019.0319

HYPOTHESIS TESTING FOR BLOCK-STRUCTURED CORRELATION FOR HIGH-DIMENSIONAL VARIABLES

Shurong Zheng[1], Xuming He[2] and Jianhua Guo[1]

[1]*Northeast Normal University and* [2]*University of Michigan*

Abstract: Testing the independence or block independence of high-dimensional random vectors is important in multivariate statistical analysis. Recent works on high-dimensional block-independence tests aim to extend their validity beyond specific distributions (e.g., Gaussian) or restrictive block sizes. In this paper, we propose a new and powerful test for the block-structured correlation of high-dimensional random vectors, for sparse or nonsparse alternatives, without strict distributional assumptions. The statistical properties of the proposed test are developed under the asymptotic regime that the dimension grows proportionally with the sample size. Empirically, we find that the proposed test outperforms existing tests for a variety of alternatives, and works quite well when there are few existing tests at our disposal.

Key words and phrases: High-dimension, multivariate statistical analysis, nonsparse alternatives, sparse alternatives, testing block-independence.

1. Introduction

Driven by a wide range of scientific applications, testing the independence of random vectors is of great importance in multivariate statistical analysis. In the conventional low-dimensional setting with $p/n \to 0$, where p is the dimension of the random vector and n is the sample size, complete and block independence tests are well established. For complete independence, Anderson (2003) proposed a *likelihood ratio test* (LRT) for Gaussian populations. For block independence, Wilks (1935) and Sugiura and Fujikoshi (1969) developed effective LRTs for Gaussian populations and derived their asymptotic distributions under regularity conditions.

In the high-dimensional setting, the classical LRT is invalid or cannot be defined as the dimension p becomes greater than the sample size n. In recent years, researchers have made great advances related to high-dimensional independence tests. For complete independence, Bai et al. (2009) proposed a corrected LRT when $p/n \to y \in (0, 1)$. Jiang and Yang (2013) studied the LRT when

Corresponding author: Shurong Zheng, Northeast Normal University and KLAS, Changchun, Jilin, China. E-mail: zhengsr@nunu.edu.cn.

$p/n \to y \in (0,1]$. Schott (2005) developed a test based on the Frobenius norm of the sample correlation matrix for $p > n$. Zhou (2007) and Cai and Jiang (2011) extended the results of Jiang (2004) to obtain the extreme distribution of coherence of the sample correlation matrices. Li and Xue (2015) proposed a quadratic-type statistic and an extreme-value-type statistic. For high-dimensional block independence, Jiang, Bai and Zheng (2013) developed a corrected LRT and trace test when $p/n \to y \in (0,1)$. Jiang and Yang (2013) studied the LRT for Gaussian populations when $p/n \to (0,1]$. Bao et al. (2017) proposed a Schott-type statistic when the dimension of every block of random variables is less than the sample size. Yamada, Hyodo and Nishiyama (2017) allowed a more general setting by using the Frobenius norm of the sample covariance matrix. Paindaveine and Verdebout (2016) proposed a high-dimensional sign test for the block-structured correlation between the random variables of two blocks under appropriate symmetry assumptions.

This study develops a new and powerful test for the block-structured correlation of a high-dimensional random vector, for sparse or nonsparse alternatives and with no strict distributional assumptions, under the asymptotic regime of $p/n \to y \in (0,\infty)$. To this end, we propose a two-term test statistic. The first term is $T_{n1} = \mathrm{tr}[\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11},\ldots,\mathbf{S}_{KK})]^2$, where the sample covariance matrix $\mathbf{S}_n$ is a natural estimator of the population covariance matrix, and the block-diagonal matrix $\mathrm{diag}(\mathbf{S}_{11},\ldots,\mathbf{S}_{KK})$ is a population covariance matrix estimator under a block-structured correlation. The statistic T_{n1} does not impose any conditions on the dimension because it does not involve a matrix inversion. The statistic T_{n1} is the sum of the squared entries of $\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11},\ldots,\mathbf{S}_{KK})$, and captures the overall difference between $\mathbf{S}_n$ and $\mathrm{diag}(\mathbf{S}_{11},\ldots,\mathbf{S}_{KK})$, even if the individual entries of $\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11},\ldots,\mathbf{S}_{KK})$ are small. That is, T_{n1}, similarly to the test of Yamada, Hyodo and Nishiyama (2017), has good power for nonsparse alternatives. The second term is a screening term, T_{n0}, which is added to T_{n1} to enhance the power under sparse alternatives. Thus, the proposed test statistic $T_{n1} + T_{n0}$ is effective for both nonsparse and sparse alternatives. To examine the performance of the proposed test statistic, the limiting null distribution is derived as $p/n \to y \in (0,\infty)$, allowing y to be greater than one. Simulation studies show that the type-I errors of the proposed test can be well maintained. Moreover, under the alternative hypothesis, the limiting distribution of the proposed test is discussed, and the asymptotic unbiasedness of the proposed test is proved. When the dimension is smaller than the sample size, simulation studies are conducted to compare our proposed test with existing tests for Gaussian populations. In the empirical power comparison, our proposed test outperforms other tests designed for high dimensions. Even when the population is nonGaussian and the dimension is greater than the sample size, our proposed test performs well.

The remainder of the paper is organized as follows. In Section 2, we propose the test statistic, derive its limiting distribution under the null and alternative hypotheses, and present the asymptotic power function to show that the proposed test is asymptotically unbiased. In Section 3, we conduct simulation studies to compare the proposed test with several existing tests. A real data set is analyzed in Section 4 for illustration. Section 5 concludes the paper.

2. Test on Block-Structured Correlation

Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a random sample from the p-dimensional population random vector $\mathbf{x} = (x_1, \ldots, x_p)^\top$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Let $\bar{\mathbf{x}} = n^{-1} \sum_{i=1}^n \mathbf{x}_i$ and $\mathbf{S}_n = (n-1)^{-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top$ be the sample mean and sample covariance matrix, respectively. Without loss of generality, the random vector $\mathbf{x} = (x_1, \ldots, x_p)^T$ can be formulated using K random variable blocks: $\{x_1, \ldots, x_{p_1}\}$, $\{x_{p_1+1}, \ldots, x_{p_1+p_2}\}, \ldots, \{x_{p_1+p_2+\cdots+p_{K-1}+1}, \ldots, x_p\}$, where $p = p_1 + \cdots + p_K$, and K is permitted to increase with n at some rate. Let $\boldsymbol{\Sigma}_{ij}$ be the covariance matrix of the ith and jth random variable blocks. The population and sample covariance matrices can be partitioned into $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{i,j=1}^K$ and $\mathbf{S}_n = (\mathbf{S}_{ij})_{i,j=1}^K$, respectively. Testing the block-structured correlation of $\mathbf{x}$ can be formulated as testing

$$H_0 : \boldsymbol{\Sigma} = \operatorname{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{KK}), \tag{2.1}$$

where $\operatorname{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{KK})$ is the block-diagonal matrix from K blocks $\{\boldsymbol{\Sigma}_{kk}, k = 1, \ldots, K\}$. A natural estimator of $\boldsymbol{\Sigma}$ is $\mathbf{S}_n$. Under the null hypothesis, a natural estimator of $\boldsymbol{\Sigma}$ is $\operatorname{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$. For a Gaussian population, the LRT statistic is Wilks (1935)

$$\log |\mathbf{S}_n| - \log |\operatorname{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})|,$$

which is the entropy loss of $\mathbf{S}_n$ and $\operatorname{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$. The entropy loss for the covariance matrix estimation can be found in James and Stein (1961) and Muirhead (1982). Jiang, Bai and Zheng (2013) proposed the following trace test statistic for the case of $K = 2$:

$$\operatorname{tr}\left[\left(\mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2}\right)\left(\mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2}\right)^\top\right],$$

which is the quadratic loss of $\mathbf{S}_n$ and $\mathrm{diag}(\mathbf{S}_{11}, \mathbf{S}_{22})$. The quadratic loss for the covariance matrix estimation can be found in Olkin and Selliah (1977), Haff (1980), and Muirhead (1982). For the block-structured correlation, regardless of the entropy loss or quadratic loss for the covariance matrix estimation, the inversion of a sample covariance matrix or log-determinant of $\mathbf{S}_{kk}$ is involved; as a result, the block dimension cannot be larger than the sample size.

We propose a test statistic with two terms, where one term is the distance between $\mathbf{S}_n$ and $\mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$, and the other term is a screening term. Motivated by the Frobenius distance between matrices, we propose the following statistic:

$$T_{n1} = \mathrm{tr}[\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})]^2.$$

Note that the statistic T_{n1} as used in Yamada, Hyodo and Nishiyama (2017) is the sum of the squared entries of $\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$, which captures the overall difference even when the individual entries of $\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$ are small nonzero numbers. Therefore, the statistic T_{n1} is not only suitable for low and high dimensions, but is also expected to perform well for nonsparse alternatives. Furthermore, to enhance the power of T_{n1} when $\mathbf{\Sigma} - \mathrm{diag}(\mathbf{\Sigma}_{11}, \ldots, \mathbf{\Sigma}_{KK})$ is very sparse, a screening term T_{n0} is added to T_{n1}. A similar idea is used in Fan, Liao and Yao (2015). Let the screening term be

$$T_{n0} = p^2 \delta_{\{\max_{(\ell_1,\ell_2)\in A_0} n (s_{\ell_1\ell_2})^2 (\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p)\}},$$

where $\delta_{\{\cdot\}}$ is an indicator function, $s^*(n,p)$ is a threshold depending on (n,p), $\mathbf{S}_n = (s_{\ell_1\ell_2})_{\ell_1,\ell_2=1}^p$, $\hat{\theta}_{\ell_1\ell_2} = n^{-1}\sum_{i=1}^n [(x_{\ell_1 i} - \bar{x}_{\ell_1})(x_{\ell_2 i} - \bar{x}_{\ell_2}) - s_{\ell_1\ell_2}]^2$, and the set

$$A_0 = \{(\ell_1, \ell_2) : \ell_1 \in \{\tilde{p}_{i-1}+1, \ldots, \tilde{p}_i\},\ \ell_2 \in \{\tilde{p}_{j-1}+1, \ldots, \tilde{p}_j\},\ 1 \le i < j \le K\}, \tag{2.2}$$

with $\tilde{p}_i = p_1 + \cdots + p_i$, $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})^\top$, $\bar{x}_{\ell_1} = n^{-1}\sum_{i=1}^n x_{\ell_1 i}$, and $\bar{x}_{\ell_2} = n^{-1}\sum_{i=1}^n x_{\ell_2 i}$. The screening term T_{n0} shows that if some s_{ℓ_1,ℓ_2} is sufficiently large, then T_{n0} is at least of order p^2. Thus, the screening term T_{n0} captures the difference between $\mathbf{S}_n$ and $\mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})$, even when $\mathbf{\Sigma} - \mathrm{diag}(\mathbf{\Sigma}_{11}, \ldots, \mathbf{\Sigma}_{KK})$ is very sparse. Our proposed test statistic is the sum of the two terms; that is,

$$\begin{aligned} T_n &= T_{n1} + T_{n0} \qquad (2.3)\\ &= \mathrm{tr}[\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})]^2 + p^2 \delta_{\{\max_{(\ell_1,\ell_2)\in A_0} n (s_{\ell_1\ell_2})^2 (\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p)\}}. \end{aligned}$$

This is expected to perform well for both nonsparse and sparse alternatives. The conditions needed on the threshold s^* are given later.

2.1. Limiting null distribution of T_n

To facilitate the formulation, we use the following independent component structure model for the data.

Assumption 1. *Let* $\{\mathbf{x}_i\}_{i=1}^n$ *satisfy the independent component structure* $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})^T = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{w}_i$*, where* $\mathbf{w}_i = (w_{1i}, \ldots, w_{pi})^\top$*, and all elements* $\{w_{ji} : j = 1, \ldots, p,\ i = 1, \ldots, n\}$ *are independent and identically distributed (i.i.d.) with* $\mathrm{E}(w_{ji}) = 0$*,* $\mathrm{E}(w_{ji}^2) = 1$*, and finite fourth moments.*

Remark 1. In fact, by (1.8) of Bai and Silverstein (2004), the existence of the finite fourth moment of w_{ji} implies that there exists a sequence $\{\eta_n\}$ satisfying $\eta_n \to 0$, $\eta_n n^{1/4} \to +\infty$, and $\eta_n^{-4}\mathrm{E}w_{ji}^4\delta_{(|w_{ji}|>\eta_n\sqrt{n})} \to 0$.

Assumption 2. *Assume that the number of blocks satisfies* $K\eta_n^2 = o(1)$*. Moreover, the spectral norm of* $\boldsymbol{\Sigma}$ *is bounded uniformly in* p*. The convergence regime* $p/n \to y \in (0, \infty)$*, for some constant* y*, is satisfied.*

In Assumption 1, moment conditions are imposed that are distribution free. For example, the Gaussian distribution and many other distributions readily satisfy the independent component structure. In Assumption 2, $K\eta_n^2 = o(1)$ allows K to increase with n at some rate. In particular, for the Gaussian distribution, we have

$$\begin{aligned}\eta_n^{-4}\mathrm{E}w_{ji}^4\delta_{(|w_{ji}|>\eta_n\sqrt{n})} &\le \eta_n^{-(4+m)}n^{-m/2}\mathrm{E}w_{ji}^{4+m}\delta_{(|w_{ji}|>\eta_n\sqrt{n})}\\ &= o(\eta_n^{-(4+m)}n^{-m/2}) = o(1),\end{aligned}$$

for any even m, if $\eta_n^{-2} = O(n^{m/(m+4)})$. Then, K can be of order $o(n^{1-\epsilon})$, for any $\epsilon > 0$.

Lemma 1. *Under Assumptions* 1 *and* 2*, and under* H_0 *specified by* (2.1)*, we have*

$$\frac{T_{n1} - \mu}{\sigma} \to N(0,1) \quad \textit{and} \quad \frac{T_{n1} - \hat{\mu}}{\sigma_0} \to N(0,1),$$

where

$$\mu = \frac{(n^2 - n - 1)[(\mathrm{tr}\boldsymbol{\Sigma})^2 - \sum_{k=1}^K(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2]}{n(n-1)^2},$$

$$\hat{\mu} = \frac{(n^2 - n - 1)[(\mathrm{tr}\mathbf{S}_n)^2 - \sum_{k=1}^K(\mathrm{tr}\mathbf{S}_{kk})^2]}{n(n-1)^2},$$

$$\sigma_0^2 = 4(n^{-1}\mathrm{tr}\mathbf{\Sigma}^2)^2 - 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}^2)^2,$$

$$\sigma^2 = \sigma_0^2 + 4n^{-3}\sum_{k=1}^{K}(\mathrm{tr}\mathbf{\Sigma}_{kk} - \mathrm{tr}\mathbf{\Sigma})^2\left[2\,\mathrm{tr}\mathbf{\Sigma}_{kk}^2 + \beta_w\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top\mathbf{\Sigma}_{kk}\mathbf{e}_{\ell k})^2\right],$$

$$\beta_w = \mathrm{E}(w_{ji}^4) - 3.$$

Here, $\mathbf{e}_\ell$ *is a* p*-dimensional vector with the* ℓ*th element equal to one and all other elements equal to zero, and* $\mathbf{e}_{\ell k}$ *is a* p_k*-dimensional vector with the* ℓ*th element equal to one and all other elements equal to zero.*

Note that we have suppressed the subscript n in many of the quantities we use, such as μ and σ^2. The proof of Lemma 1 is provided in supplementary file 1. The asymptotic variance σ_0^2 depends on the unknown parameters $\mathrm{tr}(\mathbf{\Sigma}^2)$ and $\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)$, for $k = 1, \ldots, K$. However,

$$(n-2)^{-1}[\mathrm{tr}(\mathbf{S}_{kk}^2) - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2] - n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2) = o_p(1), \quad k = 1, \ldots, K,$$

which can be used to estimate σ_0^2; see the proof in supplementary file 1. Moreover, under H_0, we have $\mathrm{tr}(\mathbf{\Sigma}^2) = \sum_{k=1}^{K}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)$; thus,

$$(n-2)^{-1}\sum_{k=1}^{K}[\mathrm{tr}(\mathbf{S}_{kk}^2) - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2] - n^{-1}\mathrm{tr}(\mathbf{\Sigma}^2) = o_p(1).$$

Therefore, σ_0^2 can be consistently estimated by

$$\hat{\sigma}_0^2 = 4(n-2)^{-2}\left\{\sum_{k=1}^{K}[\mathrm{tr}(\mathbf{S}_{kk}^2) - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2]\right\}^2 \\ -4(n-2)^{-2}\sum_{k=1}^{K}[\mathrm{tr}(\mathbf{S}_{kk}^2) - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2]^2.$$

Bai and Saranadasa (1996) suggested a uniformly minimum variance unbiased estimator of $\mathrm{tr}(\mathbf{\Sigma}^2)$ under the normality assumption, but we have used an asymptotic approximation with a finite-sample correction factor to better control type-I errors. Let

$$p_0^2 = p^2 - p_1^2 - \cdots - p_K^2. \tag{2.4}$$

The following result provides the asymptotic justification for the proposed test.

Theorem 1. *Under Assumptions* 1 *and* 2, *and under* H_0 *specified by* (2.1), *if* $\liminf_{n\to\infty} \inf_{(i,j)\in A_0} \mathrm{var}[(x_{1i} - \mathrm{E}x_{1i})(x_{1j} - \mathrm{E}x_{1j})][\mathrm{var}(x_{1i})\mathrm{var}(x_{1j})]^{-1/2} > 0$, $s^*(n,p) - 4\log p_0 \to +\infty$, *and* $\sup_{1\le\ell\le p} \mathrm{E}\exp(t_0|x_{\ell 1}|^{m_0}) < \infty$, *for some constants* $t_0 > 0$ *and* $0 < m_0 \le 2$, *then we have*

$$\hat{\sigma}_0^{-1}(T_n - \hat{\mu}) \to N(0,1).$$

Note that T_n has the same null distribution as T_{n1} in the asymptotic sense, and the second term T_{n0} plays a role mainly when the alternative hypothesis is true. The one-sided rejection region for H_0 at the nominal level α is

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n : T_n - \hat{\mu} > \hat{\sigma}_0 q_{1-\alpha}\}, \tag{2.5}$$

where q_α is the αth quantile of the standard normal distribution.

Remark 2. To apply the proposed test in practice, we need to choose the threshold $s^*(n,p)$. There are many choices for the threshold, as long as it satisfies $s^*(n,p) - 4\log p_0 \to +\infty$. For simplicity, in this paper, the threshold is taken to be

$$s^*(n,p) = [4 + (\log\log n - 1)^2](\log p_0 - 0.25\log\log p_0) + q, \tag{2.6}$$

where q satisfies $\exp[-(8\pi)^{-1/2}\exp(-q/2)] = 0.99$. The threshold ensures that even if n and p_0 are small, the probability of the event $T_{n0} = 0$ is bounded by 0.01 under H_0, because $\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2\, \hat{\theta}_{\ell_1\ell_2}^{-1} - 4\log p_0 + \log\log p_0$ converges to a type-I extreme value distribution, $\exp[-(8\pi)^{-1/2}\exp(-t/2)]$, under the null hypothesis (see Xiao and Wu (2013)). The probability of the event $T_{n0} = 0$ becomes negligible under H_0 when either n or p_0 is moderately large. For example, if $n = 200$ and $p_0 = 250$, the relevant probability is only 0.002.

Remark 3. Our proposed hypothesis test (2.5) is a global test on correlations between different blocks. If the null hypothesis is rejected, under the sparsity assumption, we may use the multiple testing method of Cai and Liu (2016) to identify individual nonzero correlations in two steps. Let

$$T_{ij} = \frac{\sum_{\ell=1}^n (x_{i\ell} - \bar{x}_i)(x_{j\ell} - \bar{x}_j)}{\sqrt{n\hat{\theta}_{ij}}}, \tag{2.7}$$

where $\hat{\theta}_{ij} = n^{-1}\sum_{\ell=1}^n [(x_{i\ell} - \bar{x}_i)(x_{j\ell} - \bar{x}_j) - s_{ij}]^2$.

Step 1: bootstrap procedure. Let $\{x_{j1}^*, \ldots, x_{jn}^*\}$ be a sample drawn randomly with replacement from $\{x_{j1}, \ldots, x_{jn}\}$, for every $j \in \{1, \ldots, p\}$. Let $\mathbf{x}_\ell^* = (x_{1\ell}^*, \ldots, x_{p\ell}^*)^T$, for $\ell = 1, \ldots, n$, and compute the bootstrap test statistic T_{ij}^* from $\mathbf{x}_1^*, \ldots, \mathbf{x}_n^*$, as in (2.7). When the above bootstrap procedure is repeated N times, we have N bootstrap test statistics $T^*_{ij1},\ldots,T^*_{ijN}$. Let

$$G^*_{n,N}(t)=\frac{2}{Np_0^2}\sum_{\ell=1}^{N}\sum_{(i,j)\in A_0} I\{|T^*_{ij\ell}|\ge t\},$$

where A_0 is given in (2.2).

Step 2: Large-scale correlation tests with bootstrap given in Cai and Liu (2016). Let

$$\hat{t}=\inf\left\{0\le t\le\sqrt{4\log p_0-2\log(\log p_0)}:\right. \\ \left.\frac{G^*_{n,N}(t)(p_0^2)/2}{\max\{\sum_{(i,j)\in A_0} I\{|T_{ij}|\ge t\},1\}}\le\alpha\right\}.$$

If $\hat{t}$ does not exist, then let $\hat{t}=\sqrt{4\log p_0}$. We reject $H_{0ij}:\sigma_{ij}=0$ whenever $|T_{ij}|\ge\hat{t}$, for $(i,j)\in A_0$.

Remark 4. On the surface, it seems that we need the eighth moment of $\mathbf{x}_i$ to calculate the variance of T_{n1}. In fact, Yamada, Hyodo and Nishiyama (2017) require a finite eighth moment condition. However, our Lemma 1 and Theorem 2 require only the fourth moment of $\mathbf{x}_i$.

2.2. Limiting distribution of T_n under the alternative hypothesis

Next, we study the theoretical property of the proposed statistic T_n under the alternative hypothesis. Let the difference between the null hypothesis and the alternative hypothesis be $\mathbf{A}=\mathbf{\Sigma}^2-\mathrm{diag}(\mathbf{\Sigma}^2_{11},\ldots,\mathbf{\Sigma}^2_{KK})$.

Theorem 2. *Under Assumptions* 1 *and* 2*, we have*

$$\sigma_1^{-1}(T_{n1}-\hat{\mu}-\mu_1)\to N(0,1),$$

where $\mu_1=(n^2-n+2)\mathrm{tr}\mathbf{A}/(n-1)^2$ *and*

$$\sigma_1^2=\sigma_0^2+4\left[2n^{-1}\mathrm{tr}\mathbf{A}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}.(\mathbf{e}_\ell^\top\mathbf{A}\mathbf{e}_\ell)^2\right].$$

Here, $\mathbf{e}_\ell$ *is a* p*-dimensional vector with the* ℓ*th element equal to one and all other elements equal to zero, and* $\beta_w=\mathrm{E}w_{ij}^4-3$.

The asymptotic power function of T_n is $\beta_{T_n}(\mathbf{A})=P(T_n-\hat{\mu}>\hat{\sigma}_0q_{1-\alpha})$. We have $P(T_n-\hat{\mu}>\hat{\sigma}_0q_{1-\alpha})-[1-\Phi(\sigma_1^{-1}(\sigma_0q_{1-\alpha}-\mu_1))]=o(1)$. Because

$\mathrm{tr}\mathbf{A} = \mathrm{tr}\mathbf{\Sigma}^2 - \sum_{k=1}^{K} \mathrm{tr}\mathbf{\Sigma}_{kk}^2 = \sum_{1\le k_1\ne k_2\le K} \mathrm{tr}\mathbf{\Sigma}_{k_1k_2}\mathbf{\Sigma}_{k_2k_1} \ge 0$, it is easy to see that $\sigma_1^2 \ge \sigma_0^2$ and $\mu_1 \ge 0$. If the population covariance matrix departs from the null hypothesis (in the sense that $\mathrm{tr}\mathbf{A} > \epsilon_0 > 0$, for any positive constant ϵ_0), then $\sigma_1^2 > \sigma_0^2$ and $\mu_1 > 0$. Under such an alternative hypothesis, we have $(\sigma_0 q_{1-\alpha} - \mu_1)/\sigma_1 < q_{1-\alpha}$; that is,

$$\beta_{T_n}(\mathbf{A}) > \alpha.$$

Thus, the proposed test T_n is asymptotically unbiased. In fact, when n is sufficiently large, $\beta_{T_n}(\mathbf{A})$ is an increasing function of $\mathrm{tr}\mathbf{A}$, where $\mathrm{tr}\mathbf{A}$ measures the departure from the null hypothesis.

Theorem 3. *Under Assumptions* 1 *and* 2 *and* $\mathbf{\Sigma}^2 = \mathrm{diag}(\mathbf{\Sigma}_{11}^2, \ldots, \mathbf{\Sigma}_{KK}^2) + \mathbf{A}$,

(1) we have $\beta_{T_n}(\mathbf{A}) \ge \alpha$ *when n is sufficiently large; in particular, when* $\mathrm{tr}\mathbf{A} > \epsilon_0 > 0$, *for any positive constant* ϵ_0, *we have* $\beta_{T_n}(\mathbf{A}) > \alpha$ *for sufficiently large n; and*

(2) if $\mathrm{tr}\mathbf{A}$ *tends to infinity or* $P(\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2(\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p))$ *converges to one, then we have* $\beta_{T_n}(\mathbf{A}) \to 1$ *as* $n \to \infty$.

Theorem 3 shows that the proposed test T_n is asymptotically unbiased. If the absolute value of at least one entry of $\mathbf{A}$ is greater than $\sqrt{(\log p_0 \log n)/n}$, then there exists $(\ell_1, \ell_2) \in A_0$ such that $n(s_{\ell_1\ell_2})^2(\hat{\theta}_{\ell_1\ell_2})^{-1}\ (s^*(n,p))^{-1} \approx c\log n/\log \log n$ converges to infinity in probability under the conditions of Theorem 1. Thus, $P(\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2(\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p)) \to 1$ holds by Remark 2, and the power converges to one.

Remark 5. Support recovery of $\mathbf{\Sigma}$: Following the proof of Theorem 5 in Cai, Liu and Xia (2013), under the conditions

$$\frac{p}{n} \to y \in (0, +\infty), \quad \min_{(i,j)\in A_0} \theta_{ij}(\sigma_{ii}\sigma_{jj})^{-1/2} > \tau,$$
$$\mathrm{E}|(x_{j1} - \mathrm{E}x_{j1})(\sigma_{jj})^{-1/2}|^{8+\epsilon} \le c_0, \ \forall\ 1 \le j \le p,$$

for some $c_0 > 0$, $\epsilon > 0$, $\tau > 0$, with the set A_0 defined in (2.2), we have

$$\liminf_{\mathbf{\Sigma}\in W_0} P(\hat{\Psi} = \Psi) \to 1,$$

where

$$\Psi = \{(i,j) : \sigma_{ij} \ne 0, (i,j) \in A_0\},$$
$$\hat{\Psi} = \{(i,j) : n(s_{ij} - \sigma_{ij})^2(\hat{\theta}_{ij})^{-1} \ge 4\log p_0, (i,j) \in A_0\},$$

$$W_0 = \left\{ \mathbf{\Sigma} : \min_{(i,j)\in\Psi} n^{1/2}|\sigma_{ij}|(\theta_{ij})^{-1/2} \geq 4\sqrt{\log p_0}, (i,j) \in A_0 \right\},$$

with $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^{p}$ and $p_0^2 = p^2 - p_1^2 - \cdots - p_K^2$ given in (2.4).

3. Simulation Studies

In this section, we evaluate the finite-sample performance of the proposed test in terms of its type-I error rates and power. Because the proposed test uses the Frobenius distance between the covariance matrices, we denote it as FDS. The test proposed by Paindaveine and Verdebout (2016) was developed for variables with mean zero. When applied to the centered variables (by removing the sample mean) in high dimensions, the test has seriously inflated type-I errors; therefore, we exclude it from the comparisons. The test used by Jiang, Bai and Zheng (2013) is the same as the test of Bao et al. (2017) when $K = 2$, but has slightly poorer performance when $K = 3$; thus, we include the latter test only. The following three competing tests are used in our comparisons:

- "CLRT": the test of Jiang and Yang (2013);
- "BHPZ": the test of Bao et al. (2017);
- "YHN": the test of Yamada, Hyodo and Nishiyama (2017);

We generate samples of size n from $\mathbf{x}_i = \mathbf{1}_p + \mathbf{\Sigma}^{1/2}\mathbf{w}_i$, for $i = 1, \ldots, n$, where $\mathbf{1}_p$ is a p-dimensional vector with all elements equal to one, $\mathbf{w}_i = (w_{1i}, \ldots, w_{pi})^{\top}$, and $\{w_{ji}, i = 1, \ldots, n, j = 1, \ldots, p\}$ are i.i.d. as $N(0, 1)$. To consider different structures of $\mathbf{\Sigma}$, we use $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \sum_{i=1}^{3} \theta_i \mathbf{\Sigma}_i$ for some values $(\theta_1, \theta_2, \theta_3)$, where $\mathbf{\Sigma}_1 = (0.5^{|i-j|})_{i,j=1}^{p}$ is approximately sparse in structure, $\mathbf{\Sigma}_2 = \mathbf{I}_p + 0.5(\delta_{\{|i-j|=1\}})_{i,j=1}^{p}$ is sparse, and $\mathbf{\Sigma}_3 = 0.98\mathbf{I}_p + 0.02\mathbf{1}_p\mathbf{1}_p^T$ is a dense structure. For each setting, we conduct 5,000 Monte Carlo simulations. For the type-I error estimates, the standard errors are approximately 0.006.

At the sample size $n = 200$, we consider the dimension $p = 60, 120, 180$, and the number of blocks $K = 2, 3$, with block sizes $p_1 = \cdots = p_K = p/K$. The ROC curves for the competing tests are plotted in Figure 1 under the null hypothesis $\mathbf{\Sigma} = 0.2\mathbf{I}_p$ and the alternative hypotheses $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \mathbf{\Sigma}_i$, for $i = 1, 2, 3$, at $n = 200$ and $p_1 = p_2 = p_3 = 20$. Clearly, the FDS test performs best for the non-dense $\mathbf{\Sigma}$. When $\mathbf{\Sigma}$ is dense, FDS and YHN are similar, but YHN is the worst performer for the sparse alterative. Moreover, the empirical size and power of each test are listed in Table 1 for a variety of settings. All methods maintain type-I errors well. The proposed FDS test outperforms the other tests in terms of power. In particular, when $(p_1, p_2, p_3) = (20, 20, 20)$ and $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \mathbf{\Sigma}_1$, the empirical power of the FDS test is about 98%, and that of the other tests is between 36% and 53%. For $(p_1, p_2, p_3) = (60, 60, 60)$ and $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \mathbf{\Sigma}_2$, the empirical power of the FDS test is about 88%, whereas that of the other tests ranges between at most 10% and 14%. Overall, the proposed FDS test is more powerful than its competitors. When $\mathbf{\Sigma}$ is dense, FDS and YHN are similar, and both lead the comparison.

When the dimension is much greater than the sample size, we examine the performance of FDS, BHPZ, and YHN only, because CLRT cannot handle such cases. In the simulation, the null hypothesis is $\mathbf{\Sigma} = 0.2\mathbf{I}_p$ and the alternative hypothesis is $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \theta_1\mathbf{\Sigma}_1 + \theta_2\mathbf{\Sigma}_2^* + \theta_3\mathbf{\Sigma}_3$, where $\mathbf{\Sigma}_2^* = \mathbf{I}_p + \rho_0(\delta_{\{|i-j|=1\}})_{i,j=1}^p$, with $\rho_0 = 0.3 + 0.3\exp(0.009p)/(0.15 + \exp(0.009p))$ and $\theta_i = 0$ or 1, for $i = 1, 2, 3$. The distribution of w_{ji} is taken to be $N(0, 1)$ or Gamma(4, 2)-2. In this study, we consider the sample sizes $n = 150, 300$, dimensions $p = 180, 360, 900$, and number of blocks $K = 2, 3$, with block sizes $p_1 = \cdots = p_K = p/K$. The empirical size and power of each test are listed in Tables 2 and 3. The type-I errors are all close to the nominal level of 0.05. Moreover, as the dimension increases, the empirical power of the tests increases with n. For example, when $\mathbf{\Sigma} = 0.2\mathbf{I}_p + \mathbf{\Sigma}_2^*$, $p = 180$ and $K = 2$, the power of FDS increases from 71.24% to 99.96% quickly as the sample size increases from $n = 150$ to 300, whereas that of other tests rises much less. To save space, Table 3 is given in supplementary file 1.

Note that the proposed FDS test does not always dominate the others when p is small. We refer to the ROC curve in Figure 1 under the null hypothesis $\mathbf{\Sigma} = 0.2\mathbf{I}_p$ and the alternative hypotheses $\mathbf{\Sigma} = \mathbf{\Sigma}_4 = 1.2\mathbf{I}_p + 0.18(\delta_{\{|i-j|=1\}})_{i,j=1}^p + 0.1(\delta_{\{|i-j|=3\}})_{i,j=1}^p$, with a sample size $n = 200$, dimension $p = 6$, and $K = 3$ blocks of equal sizes, $p_1 = p_2 = p_3 = 2$. In this case, the population is Gaussian and the likelihood is correctly specified, so it is not surprising that CLRT shows slightly better performance than FDS.

To check the sensitivity of the threshold $s^*(n, p)$ and any scaled version of T_{n0}, we consider the rejection region

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n : T_n(c_1, c_2) - \hat{\mu} > \hat{\sigma}_0 q_{1-\alpha}\}, \tag{3.1}$$

which is similar to (2.5), where $\hat{\mu}$ and $\hat{\sigma}_0$ are in (2.4), and

$$T_n(c_1, c_2) = T_{n1} + c_1 \cdot T_{n0}(c_2),$$

with $T_{n1} = \mathrm{tr}[\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})]^2$ and

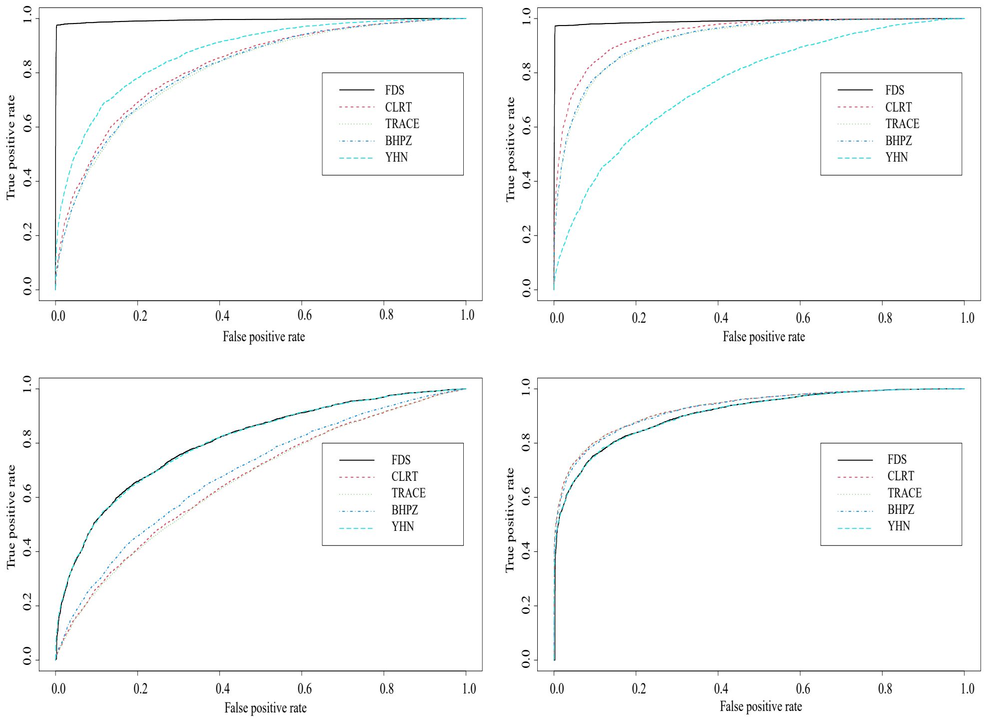

Figure 1. The first three ROC curves are the results from three simulation settings given in Section 3 with different specifications $\boldsymbol{\Sigma}_1$ (upper left panel), $\boldsymbol{\Sigma}_2$ (upper right), $\boldsymbol{\Sigma}_3$ (lower left), with w_{ij} being i.i.d from $N(0,1)$, $(n,p)=(200,60)$, and $p_1=p_2=p_3=20$. The ROC curve in the lower-right panel refers to the case of $(n,p)=(200,6)$ with $K=3$ equal block sizes and Σ_4. The curves for FDS and YHN are nearly identical in the lower-left panel and lower-right panel.

$$T_{n0}(c_2) = p^2\delta_{\{\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2(\hat{\theta}_{\ell_1\ell_2})^{-1}>s^*(n,p,c_2)\}},$$
$$s^*(n,p,c_2) = c_2\cdot[4+(\log\log n-1)^2](\log p_0 - 0.25\log\log p_0)+q.$$

We have $s^*(n,p)=s^*(n,p,1)$, $T_{n0}=T_{n0}(1)$, and $T_n=T_n(1,1)$. We consider the sample size n=200, dimension $p=60,120,180$, and number of blocks $K=2,3$, with block sizes $p_1=\cdots=p_K=p/K$. The parameters c_1 and c_2 are taken as $c_1=0.001,0.5,2$ and $c_2=0.5,1,2$. The empirical test sizes and power for different values of c_1 and c_2 are listed in Tables 4 and 5. The simulation results in Table 4 show that when c_1 is small or large, the empirical test sizes and empirical power values are similar for the different values of c_1. The simulation results in Table 5 show that when c_2 is small, the empirical test size cannot be controlled. Furthermore, when c_2 is large, although the empirical test size can be controlled, the empirical power decreases. Thus, the penalty T_{n0} is somewhat

Table 1. Empirical test sizes and power (in percentage) for comparison of four methods with $n = 200$, $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$, and $K = 2, 3$ for Gaussian variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (2.5).

$(\theta_1, \theta_2, \theta_3)$	Methods	$p = 60$	120	180	60	120	180
		$K = 2$			$K = 3$		
		Empirical test sizes					
(0, 0, 0)	FDS	4.50	4.95	4.94	5.10	4.85	4.88
	CLRT	4.74	5.52	4.86	5.02	5.30	5.12
	BHPZ	4.58	5.12	4.52	4.88	5.09	4.68
	YHN	4.64	5.07	5.07	5.18	4.94	4.88
		Empirical powers					
(1, 0, 0)	FDS	87.86	76.52	69.28	98.06	93.20	88.42
	CLRT	19.52	9.40	6.98	38.74	14.28	8.38
	BHPZ	17.46	8.80	6.64	36.08	14.72	9.55
	YHN	27.28	13.22	9.72	52.48	22.78	14.83
(0, 1, 0)	FDS	86.70	75.52	68.62	97.50	92.68	88.02
	CLRT	38.28	13.26	7.86	75.42	24.86	10.92
	BHPZ	30.86	11.82	7.82	66.78	23.62	13.26
	YHN	15.68	92.50	7.60	26.12	14.18	10.02
(0, 0, 1)	FDS	32.46	69.86	90.90	38.48	78.90	95.32
	CLRT	12.82	12.38	8.78	15.62	15.90	11.70
	BHPZ	11.92	11.32	9.00	18.10	20.20	17.62
	YHN	32.62	70.20	91.02	38.96	79.16	95.42

sensitive for the threshold $s^*(n, p)$, but is not sensitive for the scaled version of T_{n0}. Moreover, to show that our test is valid for $p/n \to y = 0$, Table 6 presents simulation results with $n = 500, 750, 1{,}000$ and $p = 6, 12, 18$. To save the space, Tables 4–6 are given in supplementary file 1.

4. Demonstration with a Real-Data Example

To further demonstrate the power of the proposed test, we use data from a major supermarket in northern China (see Wang (2009)). In the data set, each record contains the daily sales volume of individual products over a 463-day period. We are interested in understanding the correlation between vegetable sale volumes and dairy sale volumes. We have 26 major vegetables and 58 dairy products in the study; that is, $(p_1, p_2) = (26, 58)$.

To evaluate the power of various tests at small sample sizes, we randomly draw the sale volumes of vegetables and dairy products using $p_1 + p_2 + 2$ days;

Table 2. Empirical test sizes and power (in percentage) for comparison of three methods with $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$ and $K = 2, 3$ for Gaussian variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (2.5). When a test is not applicable, the corresponding entries are marked $-$.

$(\theta_1,\theta_2,\theta_3)$	n	Methods	p=180	360	900	180	360	900
			$K=2$			$K=3$		
			Empirical test sizes					
$(0,0,0)$	150	FDS	5.11	4.72	4.22	4.86	4.78	4.48
		BHPZ	4.62	—	—	5.08	4.76	—
		YHN	5.50	4.94	5.06	5.26	4.86	5.24
	300	FDS	5.08	4.92	4.93	5.08	5.08	5.02
		BHPZ	5.08	4.70	—	5.26	5.30	—
		YHN	5.04	5.08	5.33	5.42	5.32	5.12
			Empirical powers					
$(1,0,0)$	150	FDS	38.22	25.78	14.06	57.02	38.85	21.80
		BHPZ	6.14	—	—	7.84	5.26	—
		YHN	8.74	6.22	5.44	12.41	7.66	5.66
	300	FDS	97.74	94.16	87.52	99.95	99.51	97.74
		BHPZ	8.74	5.92	—	13.76	7.48	—
		YHN	12.42	8.14	6.60	22.86	11.36	7.72
$(0,1,0)$	150	FDS	71.24	59.54	41.78	89.52	80.20	61.92
		BHPZ	9.32	—	—	20.72	7.10	—
		YHN	7.68	5.86	5.32	10.22	7.18	5.24
	300	FDS	99.96	99.88	99.74	100	100	100
		BHPZ	32.22	10.50	—	74.24	27.82	—
		YHN	10.42	7.2	6.70	16.02	9.85	7.00
$(0,0,1)$	150	FDS	76.18	98.48	100	84.28	99.38	100
		BHPZ	7.24	—	—	11.20	6.48	—
		YHN	76.87	98.52	100	84.56	99.46	100
	300	FDS	99.36	100	100	99.82	100	100
		BHPZ	14.84	9.16	—	34.16	21.02	—
		YHN	99.34	100	100	99.82	100	100

that is, the sample size is $n = p_1 + p_2 + 2$. Based on 10,000 random draws at this sample size, FDS and YHN reject the null hypothesis that the sale volumes of vegetables and dairy products are uncorrelated 100% of the time. The tests CLRT and BHPZ reject the null hypothesis 58.71% and 84.22% of the time, respectively. For the sensitivity analysis with $(c_1, c_2) = (0.001, 1), (5, 1), (1, 0.5), (1, 2)$, the proposed FDS test still rejects the null hypothesis 100% of the time.

When we take a small number of days randomly from the data set, autocorrelation is negligible. To use the whole sample to understand or confirm the correlation between the prices of these two products, we use an autoregressive AR(1) model to fit the data, and then examine the residuals. In this case, all the tests we considered reject the null hypothesis of no correlation at the level 0.001. The fact that the proposed test is able to detect the correlation with high power, even when the sample size is slightly above the total dimension, indicates that the test is valuable in the analysis of moderately high-dimensional problems.

5. Discussion

We have proposed a test for detecting block-structured correlation in high-dimensional variables. The validity of the test is established under a framework where the dimension of the variables grows linearly with the sample size. For an explanation of why the framework of p/n tending to a constant is useful for high-dimensional data analysis, refer to Marcenko and Pastur (1967) and Bai and Silverstein (2010). The test can be used in a wide range of problems for Gaussian or nonGaussian variables, and attains good power for sparse or nonsparse alternatives. Our simulations show that the proposed test performs very well in terms of both the type-I error rate and power relative to existing tests, when the latter are applicable. Unlike the other tests, the proposed method does not invert any covariance matrices and requires only finite fourth moments of the random variables. More importantly, the proposed test performs quite well, even when the dimension exceeds the sample size. When p is small and n is large, and the data are Gaussian, the proposed test loses some power against the LRT, but the loss of power is limited even in these situations in our empirical studies.

Supplementary Material

The first online Supplementary Material file contains proofs of Lemma 1 and Theorems 1–3. The second file contains three lemmas and detailed proofs of (S2.6)–(S2.8) in the first file. These proofs are conducted under Assumptions 1–2. The sample covariance matrix $\mathbf{S}_n$ of 84 major vegetables and 58 dairy products in Section 4 is available at `https://math127.nenu.edu.cn/shuxue/HData/webpage/covariancematrix.zip`.

Acknowledgments

We thank the associate editor, referees, and editor for their constructive and insightful comments. Shurong Zheng's research was supported by NSFC grant 12071066. Jianhua Guo's research was supported by NSFC grants 11690012 and National Key Research and Development Program of China (No. 2020YFA07141 02). Xuming He's research was supported by National Science of Foundation (USA) Awards DMS-1607840 and DMS-1914496.

References

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* 3rd Edition. John Wiley & Sons, New York.

Bai, Z. D. and Saranadasa, H. (1996). Effect of high dimension: By an example of a two sample problem. *Statistica Sinica* **6**, 311–329.

Bai, Z. D., Jiang, D. D., Yao, J. F. and Zheng, S. R. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics* **37**, 3822–3840.

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability* **32**, 553–605.

Bai, Z. D. and Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices.* 2nd Edition. Springer, New York.

Bao, Z. G., Hu, J., Pan, G. M. and Zhou, W. (2017). Test of independence for high-dimensional random vectors based on freeness in block correlation matrices. *Electronic Journal of Statistics* **11**, 1527–1548

Cai, T. T. and Jiang, T. F. (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. T. and Liu, W. D. (2016). Large-scale multiple testing of correlations. *Journal of the American Statistical Association* **111**, 229–240.

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

Fan, J. Q., Liao, Y. and Yao, J. W. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica* **83**, 1497–1541.

Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *The Annals of Statistics* **8**, 586–597.

James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability* **1**, 361–379. University of California Press, Berkeley.

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

Jiang, D. D., Bai, Z. D. and Zheng, S. R. (2013). Testing the independence of sets of large-dimensional variables. *Science China: Mathematics* **56**, 135–147.

Jiang, T. F. and Yang, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *The Annals of Statistics* **41**, 2029–2074.

Li, D. N. and Xue, L. Z. (2015). Joint limiting laws for high-dimensional independence tests. *ArXiv: 1512.08819v1.*

Marčenko, V. A. and Pastur, L. A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sb.* **1**, 457-483.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley & Sons, New York.

Olkin, I. and Selliah, J. B. (1977). Estimating covariances in a multivariate normal distribution. In *Statistical Decision Theory and Related Topics* (Edited by S. S. Gupta and D. S. Moore) **II**, 313–326. Academic Press, New York.

Paindaveine, D. and Verdebout, T. (2016). On high-dimensional sign tests. *Bernoulli* **22**, 1745–1769.

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

Sugiura, N. and Fujikoshi, Y. (1969). Asymptotic expansions of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence. *The Annals of Mathematical Statistics* **40**, 942–952.

Wang, H. S. (2009). Forward regression for ultra-high dimensional variable screening. *Journal of the American Statistical Association* **104**, 1512–1524.

Wilks, S. S. (1935). On the indepdendence of k sets of normally distributed statistical variables. *Econometrica* **3**, 309–326.

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

Yamada, Y., Hyodo, M. and Nishiyama, T. (2017). Testing block-diagonal covariance structure for hig-dimensional data under non-normality. *Journal of Multivariate Analysis* **155**, 305–316.

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

Shurong Zheng

Northeast Normal University and KLAS, Changchun, Jilin, China.

E-mail: zhengsr@nenu.edu.cn

Xuming He

University of Michigan, Ann Arbor, MI 48109, USA.

E-mail: xmhe@umich.edu

Jianhua Guo

Northeast Normal University and KLAS, Changchun, Jilin, China.

E-mail: jhguo@nenu.edu.cn

(Received September 2019; accepted August 2020)