Accepted Manuscript

Variance-corrected tests for covariance structures with high-dimensional data

Guangyu Mao

PII: S0047-259X(17)30129-X
DOI: http://dx.doi.org/10.1016/j.jmva.2017.08.003
Reference: YJMVA 4280

To appear in: *Journal of Multivariate Analysis*

Received date : 4 March 2017

Variance-corrected tests for covariance structures with high-dimensional data

Guangyu Mao

School of Economics and Management, Beijing Jiaotong University, Beijing, 100044, P.R. China

Abstract

It has been reported in the literature that the identity and sphericity tests of Chen et al. [4] suffer from severe size distortion when they are applied to heavy-tailed data. This paper provides a theoretical explanation for this observation. New, variance-corrected identity and sphericity tests are constructed. The proposed tests are simple extensions of the tests due to Chen et al. [4] but simulation results show that they have much better statistical performance than the latter, and two other existing tests.

Keywords: Covariance structures, heavy tails, high dimension, identity test, sphericity test.

1. Introduction

Testing for the covariance structure of a random vector is an important topic in multivariate analysis. For a variety of reasons, when analyzing a data set, we may be interested in whether or not the covariance matrix of the underlying population is equal to a specified positive-definite matrix or that it is proportional to an identity matrix, or even that it has a diagonal structure, etc. When the dimension of the random vector p is far less than the sample size n, tests for different covariance structures under the assumption that p is fixed and $n \to \infty$ have been studied for a very long time, and they are well documented; see, e.g., Muirhead [11].

Over the past decade, due to the growing availability of high-dimensional data sets, which are typically such that p is comparable to, or even far larger than n, more and more attention has been paid to testing covariance structures for these kinds of data. It is now well understood that to construct effective tests under high-dimensional settings, it is more appropriate to postulate that $p \to \infty$ and $n \to \infty$ simultaneously, denoted by $(p, n) \to \infty$ below. Under this assumption, a considerable body of literature for testing high-dimensional covariance structures has been developed. Related contributions include, but are not limited to, Ledoit and Wolf [9], Srivastava [13], Bai et al. [2], Chen et al. [4], Fisher et al. [6], Cai and Jiang [3], Srivastava et al. [14], Fisher [5], Qiu and Chen [12], Wang and Yao [17], Srivastava et al. [16], Zou et al. [20], Zheng et al. [19], He and Chen [7], and Jiang and Wang [8].

In this paper, we are interested in testing two hypotheses about a high-dimensional random vector: (i) the covariance matrix is equal to an identity matrix; (ii) the covariance matrix is proportional to an identity matrix. Concretely, suppose $x_1, \ldots, x_n$ is a random sample generated from the population of a p-dimensional random vector with unknown positive-definite covariance matrix Σ_p. We aim to test

$$\mathcal{H}_0 : \Sigma_p = I_p \text{ vs. } \mathcal{H}_1 : \Sigma_p \neq I_p,$$

and

$$\tilde{\mathcal{H}}_0 : \Sigma_p = \sigma^2 I_p \text{ vs. } \tilde{\mathcal{H}}_1 : \Sigma_p \neq \sigma^2 I_p,$$

respectively, where I_p is an identity matrix of dimension p, and σ^2 is an unknown but finite positive constant. For simplicity, we refer to the tests for $\mathcal{H}_0$ and $\tilde{\mathcal{H}}_0$ as an identity test and a sphericity test, respectively.

In the literature, under the high-dimensional setting $p > n$, various tests for $\mathcal{H}_0$ and $\tilde{\mathcal{H}}_0$ have been developed by Ledoit and Wolf [9], Srivastava [13], Chen et al. [4], Fisher et al. [6], Srivastava et al. [14], Fisher [5], Wang and Yao

Email address: gymao@bjtu.edu.cn (Guangyu Mao)

[17], Srivastava et al. [16] and Zou et al. [20], to name a few. Of these papers, Chen et al. [4] is the earliest one that investigates the two tests for non-normal populations; it has attracted much attention. However, as partly noted by Zou et al. [20], the tests proposed by Chen, Zhang and Zhong (henceforth CZZ) may suffer from severe size distortion when they are applied to heavy-tailed data, even in large samples.

Motivated by the findings in Zou et al. [20], in the present paper we try to explore theoretically why CZZ's tests perform poorly when the data come from heavy-tailed distributions. Roughly speaking, we find that when approximating the distributions of their test statistics, CZZ neglected $O(p^{-1})$ bias terms in the asymptotic variances of the test statistics. Even though the bias terms are asymptotically negligible when $p \to \infty$, as will be explained below, they generally deviate much from zero even for large p in the presence of heavy tails. Therefore, omitting the $O(p^{-1})$ bias terms is the source of severe size distortion. Based on our theoretical analysis, we propose a simple method to correct the bias in the asymptotic variance of CZZ's test statistics. This leads to two *variance-corrected (VC) tests*: a VC identity test and a VC sphericity test. Simulation studies show that the VC tests can bring remarkable improvement when the data are heavy-tailed.

The rest of this paper is organized as follows. In the next section, we explore theoretically why CZZ's tests have severe size distortion when they are applied to heavy-tailed data, and we propose the VC tests. Section 3 is devoted to simulation studies about the VC tests. A short conclusion is provided in Section 4. All proofs are postponed to Appendix A.

2. Variance-corrected tests

Before proceeding to the formal theoretical analysis, we first introduce the following assumption.

Assumption 1. *Let $x_1 = (x_{11}, \ldots, x_{p1})^\top, \ldots, x_n = (x_{1n}, \ldots, x_{pn})^\top$ be p-dimensional random vectors such that, for each $j \in \{1, \ldots, n\}$,*

$$x_j = \mu + \Sigma_p^{1/2} z_j,$$

where μ is a p-dimensional constant vector, and $z_1 = (z_{11}, \ldots, z_{p1})^\top, \ldots, z_n = (z_{1n}, \ldots, z_{pn})^\top$ are iid p-dimensional random vectors satisfying: (i) $\mathrm{E}(z_1) = 0$; (ii) $\mathrm{var}(z_1) = I_p$; (iii) $\mathrm{E}(z_{i1}^3) = \xi$ for all $i \in \{1, \ldots, p\}$; (iv) $\mathrm{E}(z_{i1}^4) = \kappa$ for all $i \in \{1, \ldots, p\}$; (v) moments of z_{i1} are finite up to the eighth order for all $i \in \{1, \ldots, p\}$, and

$$\mathrm{E}(z_{i_1 1}^{\ell_1} \cdots z_{i_q 1}^{\ell_q}) = \mathrm{E}(z_{i_1 1}^{\ell_1}) \cdots \mathrm{E}(z_{i_q 1}^{\ell_q})$$

for all mutually distinct integers $i_1, \ldots, i_q$ and $\ell_1, \ldots, \ell_q \geq 0$ with $\ell_1 + \cdots + \ell_q = 8$.

The assumption is similar to that adopted by CZZ. When $z_{11}, \ldots, z_{p1}$ are mutually independent, Assumption (v) automatically holds, but the latter is sufficient to ensure the validity of our theoretical results. In the literature, besides CZZ, (v) was used, e.g., by Bai and Saranadasa [1], Srivastava and Kubokawa [15], Zhang et al. [18], Srivastava et al. [16], and He and Chen [7].

CZZ observed that $\mathrm{tr}(\Sigma_p^2)/p - 2\mathrm{tr}(\Sigma_p)/p \geq -1$ with equality if and only if $\mathcal{H}_0$ is true, and $p\mathrm{tr}(\Sigma_p^2)/\{\mathrm{tr}(\Sigma_p)\}^2 \geq 1$ with equality if and only if $\tilde{\mathcal{H}}_0$ is true. This led them to construct, under Assumption 1, unbiased estimators $T_{1,n}$ and $T_{2,n}$ for $\mathrm{tr}(\Sigma_p)$ and $\mathrm{tr}(\Sigma_p^2)$, respectively. CZZ then proposed statistics for testing $\mathcal{H}_0$ and $\tilde{\mathcal{H}}_0$. The two estimators are given by

$$T_{1,n} = Y_{1,n} - Y_{3,n}, \tag{1}$$
$$T_{2,n} = Y_{2,n} - 2Y_{4,n} + Y_{5,n}, \tag{2}$$

where

$$Y_{1,n} = \frac{1}{n}\sum_{i=1}^{n} x_i^\top x_i,\; Y_{3,n} = \frac{1}{n_1}\sum_{i,j}^{*} x_i^\top x_j,$$
$$Y_{2,n} = \frac{1}{n_1}\sum_{i,j}^{*}(x_i^\top x_j)^2,\; Y_{4,n} = \frac{1}{n_2}\sum_{i,j,k}^{*} x_i^\top x_j x_j^\top x_k,\; Y_{5,n} = \frac{1}{n_3}\sum_{i,j,k,\ell}^{*} x_i^\top x_j x_k^\top x_\ell,$$

in which $n_i = n\prod_{j=1}^{i}(n-j)$ for each $i \in \{1,2,3\}$, and $\sum^*$ denotes summation over mutually distinct indices. For example, $\sum^*_{i,j,k}$ means summation over $\{(i,j,k) : i \neq j, j \neq k, k \neq i\}$.

CZZ proved that under Assumption 1, one has, as $(p,n) \to \infty$,

$$nV_n \rightsquigarrow \mathcal{N}(0,4) \text{ under } \mathcal{H}_0 \quad \text{and} \quad nU_n \rightsquigarrow \mathcal{N}(0,4) \text{ under } \tilde{\mathcal{H}}_0, \tag{3}$$

where $\rightsquigarrow$ denotes convergence in distribution, and

$$V_n = T_{2,n}/p - 2T_{1,n} + 1/p, \quad U_n = p(T_{2,n}/T_{1,n}^2) - 1.$$

It can thus be expected that under the null hypotheses, the distributions of $nV_n/\sqrt{\mathrm{var}(nV_n)}$ and $nU_n/\sqrt{\mathrm{var}(nU_n)}$ can be well approximated by the standard Normal distribution, $\mathcal{N}(0,1)$.

As can be surmised from Eq. (3), CZZ employed the asymptotic variances of nV_n and nU_n when $(p,n) \to \infty$, namely 4, to estimate $\mathrm{var}(nV_n)$ and $\mathrm{var}(nU_n)$, respectively. If $\mathrm{var}(nV_n)$ and $\mathrm{var}(nU_n)$ deviate much from 4 for finite p and n, however, their tests may then suffer from severe size distortion. To see the difference between $\mathrm{var}(nV_n)$ and $\mathrm{var}(nU_n)$, and their limits, it is helpful to first compute the exact variances of $T_{1,n}$ and $T_{2,n}$, as well as their covariance.

Proposition 1. *Under Assumption 1 and* $\tilde{\mathcal{H}}_0$,

(i) $\mathrm{var}(T_{1,n}) = \sigma^4 p(\kappa-1)/n + 2\sigma^4 p/n_1$;

(ii) $\mathrm{var}(T_{2,n}) = 2\sigma^8 p(b_1 n^3 + b_2 n^2 + b_3 n + b_4)/n_3$, *where*

$$b_1 = 2\kappa - 2, \quad b_2 = 2p + \kappa^2 - 14\kappa - 4\xi^2 + 15,$$

$$b_3 = -6p - 5\kappa^2 + 36\kappa + 28\xi^2 - 45, \quad b_4 = 4p + 6\kappa^2 - 36\kappa - 48\xi^2 + 62;$$

(iii) $\mathrm{cov}(T_{1,n}, T_{2,n}) = 2\sigma^6 p\{(\kappa-1)n - \kappa - \xi^2 + 3\}/n_1$.

Calling on this proposition, we can show the following results.

Theorem 1. (*i*) *Under Assumption 1 and* $\mathcal{H}_0$,

$$\mathrm{var}(nV_n) = \frac{n^2}{pn_3}(c_1 n^2 + c_2 n + c_3), \tag{4}$$

where $c_1 = 4p + 2\kappa^2 - 4\kappa - 2$, $c_2 = -12p - 10\kappa^2 + 28\kappa + 16\xi^2 - 6$, *and* $c_3 = 8p + 12\kappa^2 - 48\kappa - 48\xi^2 + 52$.

(*ii*) *Under Assumption 1 and* $\tilde{\mathcal{H}}_0$, $\mathrm{var}(nU_n)$ *is approximately equal to the right-hand side of Eq. (4) via the delta method.*

As we can see from (i), the leading term of $\mathrm{var}(nV_n)$ is $4 + 2(\kappa^2 - 2\kappa - 1)/p$. As a result, $\lim_{(p,n)\to\infty} \mathrm{var}(nV_n) = 4$, which confirms CZZ's results. However, the approximation $\mathrm{var}(nV_n) \approx 4$ may be poor if $z_1, \ldots, z_n$ and hence $x_1, \ldots, x_n$ are heavy-tailed, because κ may be much large in this case.

To see this, suppose that for each $j \in \{1, \ldots, n\}$, $z_{1j}, \ldots, z_{pj}$ are iid random variables generated from the standardized $\mathcal{LN}(0,1)$, where $\mathcal{LN}(0,1)$ refers to the standard log-normal distribution. It is easy to verify that $\kappa = 114$ by the probabilistic properties of the log-normal distribution. Thus, $\mathrm{var}(nV_n) = 4 + 25534/p + o(1)$. Suppose we are now testing $\mathcal{H}_0$ at the 5% level using CZZ's identity test and the dimension of the data set is 25,534. Since

$$\Pr(nV_n/2 \leq 1.645) = \Pr(nV_n/\sqrt{5} \leq 3.29/\sqrt{5}) \approx \Phi(3.29/\sqrt{5}) \approx 0.93,$$

where Φ is the cumulative distribution function of the standard Normal distribution, the actually controlled size is about 7%, not 5%. In other words, the approximation $\mathrm{var}(nV_n) \approx 4$ is not sufficiently accurate. Therefore, when facing heavy-tailed data, CZZ's identity test may be oversized due to large deviation of the $O(p^{-1})$ bias from zero. The same analysis goes for CZZ's sphericity test. Theorem 1 also shows that increasing p is instrumental in alleviating the size distortion for CZZ's tests, but this is not true for n.

Since we have derived the exact forms of the $O(p^{-1})$ bias terms in $\mathrm{var}(nV_n)$ and $\mathrm{var}(nU_n)$, and given that they are the main sources of size distortion, it is natural to correct these biases. Because in the $O(p^{-1})$ bias terms the fourth moment of z_i is the unique unknown quantity, we can reach our aim if κ can be consistently estimated. Now, suppose $\hat{\kappa}_V$ and $\hat{\kappa}_U$ are two consistent estimators of κ under $\mathcal{H}_0$ and $\tilde{\mathcal{H}}_0$ respectively. We have the following results.

Theorem 2. (*i*) *Under Assumption 1 and* $\mathcal{H}_0$*,*

$$\frac{nV_n}{\sqrt{4 + 2(\hat{\kappa}_V^2 - 2\hat{\kappa}_V - 1)/p}} \rightsquigarrow \mathcal{N}(0,1) \text{ as } (p,n) \to \infty;$$

(*ii*) *Under Assumption 1 and* $\tilde{\mathcal{H}}_0$*,*

$$\frac{nU_n}{\sqrt{4 + 2(\hat{\kappa}_U^2 - 2\hat{\kappa}_U - 1)/p}} \rightsquigarrow \mathcal{N}(0,1) \text{ as } (p,n) \to \infty.$$

We refer to the identity test and sphericity test based on the theorem as *variance-corrected (VC) tests.* The theorem holds since the test statistics are asymptotically equivalent to CZZ's when

$$\hat{\kappa}_V \xrightarrow{p} \kappa \quad \text{and} \quad \hat{\kappa}_U \xrightarrow{p} \kappa,$$

where $\xrightarrow{p}$ means convergence in probability. Note that when $\hat{\kappa}_V$ and $\hat{\kappa}_U$ are inconsistent, Theorem 2 may still hold provided that $\hat{\kappa}_V^2/p = o_p(1)$ and $\hat{\kappa}_U^2/p = o_p(1)$. Here we require the consistency of the two estimators because our aim is to reduce the influence of the $O(p^{-1})$ biases.

Let $\tilde{X}$ be the $p \times n$ matrix whose (i,j) element is $x_{ij} - \sum_{j=1}^n x_{ij}/n$. Accordingly, the sample covariance matrix can be defined by $\hat{\Sigma}_p = \tilde{X}\tilde{X}^\top/(n-1)$. To estimate κ under $\mathcal{H}_0$, it is reasonable to consider the estimator

$$\hat{\kappa}_V = \frac{1}{np}\sum_{i=1}^p\sum_{j=1}^n \tilde{X}_{ij}^4, \tag{5}$$

where $\tilde{X}_{ij}$ is the (i,j) element of $\tilde{X}$. The estimator $\hat{\kappa}_V$ was proposed by Wang and Yao [17] and is generally consistent as discussed in their Remark 2.1. To estimate κ under $\tilde{\mathcal{H}}_0$, we can use

$$\hat{\kappa}_U = \frac{\sum_{i=1}^p\sum_{j=1}^n \tilde{X}_{ij}^4/(np)}{\{\mathrm{tr}(\hat{\Sigma}_p)/p\}^2}, \tag{6}$$

due to the fact that $\mathrm{tr}(\hat{\Sigma}_p)/p$ is a consistent estimator of σ^2 according to Chen et al. [4].

For the estimators in Eqs. (5) and (6), we may expect that $\hat{\kappa}_V - \kappa = O_p(1/\sqrt{np})$ under $\mathcal{H}_0$ and $\hat{\kappa}_U - \kappa = O_p(1/\sqrt{np})$ under $\tilde{\mathcal{H}}_0$. If this is true, both the leading terms in the differences between $\mathrm{var}(nV_n)$ and $\mathrm{var}(nU_n)$, and their corresponding corrected variances in Theorem 2 are $-12n^3/n_3$, which is of order $O(n^{-1})$ and does not depend on κ. As a consequence, the corrected variances can largely remove the influence of κ. Besides, if $\hat{\kappa}_V^2/p = o_p(1)$ and $\hat{\kappa}_U^2/p = o_p(1)$ hold under the alternatives, consistency of the VC tests can be justified by CZZ's Theorem 3 and 4 since they are asymptotically equivalent to CZZ's tests.

Remark 1. *By virtue of Theorem 1, we may do a complete correction for CZZ's tests, i.e., we may estimate* $\mathrm{var}(nV_n)$ *or* $\mathrm{var}(nU_n)$ *by* $n^2(\hat{c}_1 n^2 + \hat{c}_2 n + \hat{c}_3)/(pn_3)$*, where* $\hat{c}_1$*,* $\hat{c}_2$ *and* $\hat{c}_3$ *are obtained by replacing* κ *and* ξ *in* c_1*,* c_2 *and* c_3 *by consistent estimators respectively. In this case, we can estimate* ξ *by* $\hat{\xi}_V = \sum_{i=1}^p\sum_{j=1}^n \tilde{X}_{ij}^3/(np)$ *under* $\mathcal{H}_0$*, or* $\hat{\xi}_U = (np)^{-1}\sum_{i=1}^p\sum_{j=1}^n \tilde{X}_{ij}^3/\{\mathrm{tr}(\hat{\Sigma}_p)/p\}^{3/2}$ *under* $\tilde{\mathcal{H}}_0$*. However, our extensive simulation studies (not reported in this paper) show that further improvement is generally limited if the complete correction is carried out. Therefore, we recommend the correction introduced above for practical use.*

3. Simulation studies

To evaluate the proposed VC tests, we compared several existing tests in the literature with ours by simulation. When testing $\mathcal{H}_0$, we made a comparison between IT_{CZZ} (CZZ's identity test) with IT_{VC} (the VC identity test). When testing $\tilde{\mathcal{H}}_0$, we compared ST_{CZZ} (CZZ's sphericity test), ST_{VC} (the VC sphericity test), ST_{CJ} (the corrected John's test of Wang and Yao [17]), and ST_{BCS} (the bias-corrected sign test of Zou et al. [20]).

As we can see from (1) and (2), the forms of CZZ's statistics are complicated. However, as noted recently by Mao [10], $T_{1,n}$ and $T_{2,n}$ are identical to two other unbiased estimators for $\mathrm{tr}(\Sigma_p)$ and $\mathrm{tr}(\Sigma_p^2)$ proposed by Srivastava et al. [16], viz.

$$T_{1,n} = \mathrm{tr}(\hat{\Sigma}_p), \quad T_{2,n} = \frac{1}{n_3}\Big\{(n-2)(n-1)^3\mathrm{tr}(\hat{\Sigma}_p^2) + (n-1)^2\mathrm{tr}(\hat{\Sigma}_p)^2 - n_1\sum_{j=1}^{n}(\tilde{x}_j^\top\tilde{x}_j)^2\Big\},$$

where $\tilde{x}_j = x_j - \sum_{\ell=1}^{n} x_\ell/n$. Thus, CZZ's test statistics and ours can be quickly computed by using the traces of $\hat{\Sigma}_p$ and $\hat{\Sigma}_p^2$.

The corrected John's test of Wang and Yao [17] is based on the statistic $\{(n-1)\tilde{U}_n - p - \hat{\kappa}_U + 2\}/2$, where $\tilde{U}_n = p\,\mathrm{tr}(\hat{\Sigma}_p^2)\mathrm{tr}(\hat{\Sigma}_p)^{-2} - 1$. Under $\tilde{\mathcal{H}}_0$, Wang and Yao [17] proved that the statistic converges in distribution to $\mathcal{N}(0,1)$ as $(p,n) \to \infty$ and $p/n \to c \in (0,\infty)$. The bias-corrected sign test of Zou et al. [20] are designed to test $\tilde{\mathcal{H}}_0$ for p-dimensional random vectors from an elliptical distribution. The test statistic is of the form $\tilde{\sigma}_0(\tilde{Q} - p\delta_{n,p})$, where $\tilde{\sigma}_0 = 2\sqrt{(p-1)/\{n_1(p+2)\}}$ and

$$\tilde{Q} = \frac{p}{n_1}\sum_{i,j}^{*}\frac{(x_i - \hat{\theta}_{n,p})^\top(x_j - \hat{\theta}_{n,p})}{\| x_i - \hat{\theta}_{n,p} \| \times \| x_j - \hat{\theta}_{n,p} \|} - 1,$$

where $\hat{\theta}_{n,p} = \arg\min_\theta \sum_{j=1}^{n} \| x_j - \theta \|$, $\delta_{n,p}$ is an unknown term defined by Eq. (5) in Zou et al. [20]. When $\delta_{n,p}$ is appropriately estimated (see discussion therein), the authors showed that the resulting statistic converges in distribution to $\mathcal{N}(0,1)$ as $(p,n) \to \infty$ and $p = O(n^2)$.

Let (e_{ij}) be a double array of iid random variables. In our simulation, we considered three scenarios for generating the e_{ij}'s, viz.

(i) $e_{ij} \sim \mathcal{G}(4, \sqrt{2}/2)$, Gamma distribution with shape parameter 4 and scale parameter $\sqrt{2}/2$;

(ii) $e_{ij} \sim \mathcal{P}(9)$, basic Pareto distribution with shape parameter 9;

(iii) $e_{ij} \sim \mathcal{LN}(0,1)$.

The kurtosis of these three distributions is 4.5, 19.8, and 114, respectively. Therefore, only the first scenario corresponds to the case of light tails. All e_{ij} were standardized to be of zero mean and unit variance, and then used to generate $z_1,\ldots,z_n$ or $x_1,\ldots,x_n$ directly. The performance of the tests was compared in terms of their empirical size and power. Our simulation considers different combinations of $p \in \{60, 120, 240, 480, 600\}$ and $n \in \{50, 80, 120, 180, 240\}$. For each combination, we generated 2500 independent replicates. All the tests were performed at the 5% level.

First, we investigate the empirical size of IT_{CZZ} and IT_{VC}. Under $\mathcal{H}_0$, $z_1,\ldots,z_n$ were produced by setting, for each $j \in \{1,\ldots,n\}$, $(z_{1j},\ldots,z_{pj})^\top = (e_{1j},\ldots,e_{pj})^\top$. Simulation results are summarized in Table 1. As we can see, when the data are Gamma, both IT_{CZZ} and IT_{VC} perform well. However, as the kurtosis grows, IT_{CZZ} becomes noticeably oversized. In contrast, IT_{VC} still can maintain acceptable size. Besides, we can find that increasing n is unhelpful to relieve the size distortion of IT_{CZZ}, but increasing p does help. These phenomena confirm our theoretical analysis above.

To study the empirical power of IT_{CZZ} and IT_{VC}, we consider three alternatives: (i) x_j has neighbor correlation; (ii) x_j has non-unit variance; (iii) x_j has both neighbor correlation and non-unit variance. In case (i), we set

$$x_j = (x_{1j},\ldots,x_{pj})^\top = \omega_1 \times (e_{1j},\ldots,e_{pj})^\top + \omega_2 \times (e_{2j},\ldots,e_{p+1,j})^\top, \tag{7}$$

where $\omega_1 = 84/85$ and $\omega_2 = 13/85$ for scenario (i) and (ii); $\omega_1 = 15/17$ and $\omega_2 = 8/17$ for scenario (iii). Thus, x_{ij} is only correlated with its neighbors: $x_{i-1,j}$ and $x_{i+1,j}$.

As Table 2 shows, the empirical power of IT_{VC} tends to 1 as $(p,n) \to \infty$ for all scenarios. Therefore, IT_{VC} is a consistent test. In case (ii), we let $(z_{1j},\ldots,z_{pj})^\top = (e_{1j},\ldots,e_{pj})^\top$ and $\sigma^2 = 1.5$. The results in Table 3 also confirm the consistency of IT_{VC}. In case (iii), to generate x_j, we continued to employ the sampling scheme of case (i) but scale x_{ij} such that $\mathrm{var}(x_{ij}) = 1.5$. Compared to Tables 2–3, the results in Table 4 suggest that IT_{VC} will possess higher power when both the neighbor correlation and the non-unit variance are present.

Table 1: Empirical size (%) of identity tests

Scenario	$p \mid n$	IT_{CZZ} 50	80	120	180	240	IT_{VC} 50	80	120	180	240
$\mathcal{GA}(4, \sqrt{2}/2)$	60	7.16	5.40	6.96	6.32	6.12	6.48	4.96	6.04	5.84	5.36
	120	6.60	6.52	6.08	5.88	5.56	6.32	6.12	5.76	5.56	5.20
	240	5.04	4.88	5.24	5.68	5.76	4.88	4.76	5.12	5.60	5.68
	480	5.64	5.08	4.84	5.44	5.32	5.52	4.96	4.84	5.44	5.28
	600	6.20	5.56	5.48	5.48	5.68	6.12	5.44	5.48	5.40	5.64
$\mathcal{PA}(9)$	60	15.64	16.64	18.00	19.12	18.68	6.40	6.52	6.36	6.20	5.88
	120	11.92	12.24	13.48	14.80	14.40	5.56	5.40	5.64	6.28	5.76
	240	9.48	9.80	9.92	10.40	10.36	5.88	5.44	5.32	5.72	5.48
	480	8.08	7.52	7.96	7.64	7.68	5.36	5.24	5.24	5.12	5.36
	600	7.28	7.64	7.36	7.20	7.96	4.92	5.60	5.24	5.36	6.08
$\mathcal{LN}(0, 1)$	60	25.24	27.16	28.48	31.00	32.04	6.16	6.60	5.92	5.80	5.04
	120	22.72	24.72	25.84	28.68	29.28	5.36	5.72	5.52	5.36	5.16
	240	20.56	23.12	25.16	26.36	25.76	5.04	5.48	5.16	5.36	5.24
	480	18.12	20.12	21.84	23.60	22.64	5.16	5.40	5.76	5.64	5.04
	600	16.88	19.56	20.76	22.52	22.36	5.28	5.72	5.92	5.20	4.92

[*] IT_{CZZ}, the identity test of Chen et al. [4]; IT_{VC}, the variance-corrected identity test proposed in this paper.

Next, we shift our attention to the four tests for $\tilde{\mathcal{H}}_0$. To generate data for studying the size of the tests, the sampling scheme for Table 1 was employed, except that we let $\sigma^2 = 2$. Table 5 collects the simulation results. It shows that for the scenario of light tails, the four tests perform similarly except that ST_{BCS} tends to be oversized when p is small. Since ST_{BCS} is justified under the condition $p = O(n^2)$, this finding may be caused by the violation of the condition. For the other two scenarios, all the tests except ST_{VC} suffer from obvious size distortion. Therefore, ST_{VC} outperforms the other rivals. Besides, the simulation results confirm our theoretical analysis again that increasing p will alleviate the size distortion of ST_{CZZ}.

To investigate the power performance, we continue to sample $x_1, \ldots, x_n$ by (7) but scale x_{ij} such that $\mathrm{var}(x_{ij}) = 2$.

Table 2: Empirical power (%) of identity tests against neighbor correlation

Scenario	$p \mid n$	IT_{CZZ} 50	80	120	180	240	IT_{VC} 50	80	120	180	240
$\mathcal{GA}(4, \sqrt{2}/2)$	60	32.96	54.12	81.40	97.76	100.00	31.52	52.20	79.84	97.32	100.00
	120	30.88	56.00	81.60	98.44	99.92	30.20	54.56	80.92	98.24	99.92
	240	31.28	54.24	83.56	99.16	99.96	30.84	53.96	83.28	99.12	99.96
	480	30.40	56.52	84.28	98.88	100.00	30.16	56.28	84.04	98.80	100.00
	600	32.00	56.04	84.64	99.24	100.00	31.52	55.88	84.60	99.24	100.00
$\mathcal{PA}(9)$	60	34.76	49.96	69.60	88.64	97.24	17.24	27.08	42.68	68.64	86.88
	120	34.56	52.40	74.80	93.92	98.96	20.48	33.64	55.88	82.92	96.40
	240	33.28	53.52	79.44	96.20	99.68	23.52	40.52	67.64	92.52	99.24
	480	34.20	55.60	80.44	97.48	99.88	27.88	47.80	74.28	96.04	99.80
	600	34.12	55.80	82.16	98.12	99.96	28.96	49.08	77.24	97.16	99.80
$\mathcal{LN}(0, 1)$	60	96.16	99.16	99.84	99.92	99.96	72.80	87.16	93.44	97.40	98.16
	120	98.32	99.80	99.92	100.00	100.00	81.68	93.04	96.52	98.20	99.24
	240	99.36	100.00	99.96	100.00	100.00	89.36	96.28	98.24	99.12	99.44
	480	99.72	100.00	100.00	100.00	100.00	93.72	97.52	98.80	99.36	99.64
	600	99.80	100.00	100.00	100.00	100.00	95.00	98.00	99.04	99.52	99.68

[*] IT_{CZZ}, the identity test of Chen et al. [4]; IT_{VC}, the variance-corrected identity test proposed in this paper.

Table 3: Empirical power (%) of identity tests against non-unit variance

Scenario	$p \mid n$	IT_{CZZ} 50	80	120	180	240	IT_{VC} 50	80	120	180	240
	60	95.80	99.88	100.00	100.00	100.00	94.44	99.76	100.00	100.00	100.00
	120	96.92	100.00	100.00	100.00	100.00	96.24	100.00	100.00	100.00	100.00
$\mathcal{GA}(4, \sqrt{2}/2)$	240	97.52	99.96	100.00	100.00	100.00	97.12	99.96	100.00	100.00	100.00
	480	96.80	100.00	100.00	100.00	100.00	96.68	100.00	100.00	100.00	100.00
	600	97.68	100.00	100.00	100.00	100.00	97.52	100.00	100.00	100.00	100.00
	60	84.92	97.20	99.92	100.00	100.00	50.60	79.72	96.28	99.64	100.00
	120	90.08	99.00	99.96	100.00	100.00	66.08	92.56	99.52	99.96	99.96
$\mathcal{PA}(9)$	240	94.32	99.60	100.00	100.00	100.00	82.76	97.76	99.88	99.92	99.96
	480	96.48	99.88	100.00	100.00	100.00	90.36	99.60	99.96	99.96	100.00
	600	96.28	100.00	100.00	100.00	100.00	91.96	99.64	99.96	100.00	100.00
	60	56.00	72.20	85.64	95.40	98.96	12.08	16.24	25.12	42.64	60.16
	120	63.44	79.28	92.40	98.96	99.84	14.96	22.76	38.40	62.52	79.44
$\mathcal{LN}(0, 1)$	240	70.96	88.60	97.48	99.92	100.00	20.28	33.76	56.56	80.92	91.28
	480	78.80	94.96	99.60	100.00	100.00	28.48	50.16	75.96	92.28	96.60
	600	81.04	95.72	99.80	100.00	100.00	33.44	57.04	81.76	94.52	97.00

* IT_{CZZ}, the identity test of Chen et al. [4]; IT_{VC}, the variance-corrected identity test proposed in this paper.

Table 4: Empirical power (%) of identity tests against neighbor correlation and non-unit variance simultaneously

Scenario	$p \mid n$	IT_{CZZ} 50	80	120	180	240	IT_{VC} 50	80	120	180	240
	60	99.80	100.00	100.00	100.00	100.00	99.52	100.00	100.00	100.00	100.00
	120	99.72	100.00	100.00	100.00	100.00	99.68	100.00	100.00	100.00	100.00
$\mathcal{GA}(4, \sqrt{2}/2)$	240	99.88	100.00	100.00	100.00	100.00	99.88	100.00	100.00	100.00	100.00
	480	99.88	100.00	100.00	100.00	100.00	99.88	100.00	100.00	100.00	100.00
	600	99.92	100.00	100.00	100.00	100.00	99.92	100.00	100.00	100.00	100.00
	60	94.84	99.80	100.00	100.00	100.00	73.16	95.04	99.28	99.92	99.96
	120	98.04	99.96	100.00	100.00	100.00	88.08	98.84	99.96	99.96	99.96
$\mathcal{PA}(9)$	240	98.88	100.00	100.00	100.00	100.00	95.36	99.84	99.96	100.00	100.00
	480	99.44	100.00	100.00	100.00	100.00	98.72	99.96	100.00	100.00	100.00
	600	99.56	100.00	100.00	100.00	100.00	98.88	99.96	100.00	100.00	100.00
	60	99.64	100.00	100.00	100.00	100.00	85.32	94.24	96.80	98.72	99.12
	120	100.00	100.00	100.00	100.00	100.00	91.68	97.20	98.60	99.32	99.52
$\mathcal{LN}(0, 1)$	240	100.00	100.00	100.00	100.00	100.00	95.52	98.28	99.24	99.56	99.72
	480	100.00	100.00	100.00	100.00	100.00	96.96	98.68	99.20	99.64	99.84
	600	100.00	100.00	100.00	100.00	100.00	97.64	98.88	99.48	99.64	99.88

* IT_{CZZ}, the identity test of Chen et al. [4]; IT_{VC}, the variance-corrected identity test proposed in this paper.

Table 5: Empirical size (%) of sphericity tests

Scenario	$p \mid n$	ST_{CZZ} 50	80	120	180	240	ST_{VC} 50	80	120	180	240
$\mathcal{GA}(4, \sqrt{2}/2)$	60	6.96	5.24	6.52	5.72	5.88	6.48	4.52	5.72	5.20	5.12
	120	6.56	6.44	6.12	5.88	5.48	6.36	6.08	5.68	5.20	5.04
	240	5.12	5.04	5.24	5.64	5.68	5.04	4.76	5.12	5.52	5.40
	480	5.80	5.00	4.80	5.36	5.36	5.68	5.00	4.64	5.36	5.20
	600	6.28	5.60	5.40	5.40	5.64	6.12	5.52	5.36	5.32	5.56
$\mathcal{PA}(9)$	60	14.44	15.72	16.52	17.60	16.60	5.96	5.92	5.40	5.52	5.12
	120	11.60	11.16	12.76	14.24	13.40	5.40	4.84	5.36	5.92	5.56
	240	9.52	9.56	9.76	10.16	10.08	5.80	5.32	5.08	5.44	5.00
	480	7.80	7.40	7.92	7.28	7.48	5.28	5.20	5.16	4.88	5.40
	600	6.88	7.52	7.16	7.08	7.96	4.84	5.52	5.16	5.12	5.76
$\mathcal{LN}(0, 1)$	60	22.48	23.52	24.08	26.96	28.24	4.68	4.68	4.28	4.36	3.76
	120	20.96	22.96	23.52	26.16	26.64	4.52	5.04	4.88	4.56	4.48
	240	19.92	21.72	23.88	24.96	24.28	4.88	5.36	4.68	4.80	4.84
	480	17.48	19.80	21.28	22.56	21.76	4.92	5.36	5.68	5.36	4.92
	600	16.28	18.60	20.64	21.56	21.88	5.32	5.68	5.72	5.12	4.72

Scenario	$p \mid n$	ST_{CJ} 50	80	120	180	240	ST_{BCS} 50	80	120	180	240
$\mathcal{GA}(4, \sqrt{2}/2)$	60	8.16	6.36	7.00	6.16	6.08	9.92	9.00	9.80	10.04	9.68
	120	7.36	6.80	6.40	5.88	5.92	5.44	6.04	6.08	5.28	5.00
	240	6.84	5.20	5.60	5.76	6.00	3.92	4.60	5.04	4.96	5.24
	480	6.52	5.80	5.48	5.64	5.36	5.56	4.72	4.60	5.32	4.88
	600	7.08	6.52	6.00	5.44	6.12	5.84	4.96	5.00	5.28	5.44
$\mathcal{PA}(9)$	60	23.08	22.32	21.76	20.48	19.48	10.64	11.12	12.56	12.84	12.76
	120	22.44	20.00	18.16	17.24	15.64	8.56	8.28	9.20	9.48	9.20
	240	23.48	19.00	17.40	15.16	14.12	8.28	7.40	7.96	7.72	7.76
	480	24.48	19.32	17.12	13.80	13.04	7.00	7.04	7.40	6.56	6.76
	600	24.04	20.04	16.76	14.00	12.60	6.20	6.64	7.08	6.08	6.92
$\mathcal{LN}(0, 1)$	60	33.64	32.56	30.24	31.56	30.72	10.48	10.08	11.64	11.44	11.04
	120	38.28	34.76	33.44	32.92	31.24	10.56	10.44	11.32	11.04	10.72
	240	40.36	36.68	35.52	34.36	32.24	9.68	9.20	9.60	8.96	8.36
	480	42.12	39.52	37.68	35.28	33.48	10.80	12.88	12.96	13.28	12.40
	600	41.96	39.72	37.40	34.60	34.00	10.56	13.80	14.72	14.16	13.48

* ST_{CZZ}, the sphericity test of Chen et al. [4]; ST_{VC}, the variance-corrected sphericity test proposed in this paper; ST_{CJ}, the corrected John's test of Wang and Yao [17]; ST_{BCS}, the bias-corrected sign test of Zou et al. [20].

Table 6: Empirical power (%) of sphericity tests against neighbor correlation

Scenario	$p \mid n$	ST_{CZZ} 50	80	120	180	240	ST_{VC} 50	80	120	180	240
	60	31.84	53.60	80.56	97.52	99.96	30.68	51.40	79.40	97.24	99.96
	120	30.92	54.84	81.40	98.36	99.92	29.68	53.84	80.48	98.20	99.92
$\mathcal{GA}(4, \sqrt{2}/2)$	240	31.28	54.24	83.52	99.12	99.96	31.08	53.72	83.28	99.08	99.96
	480	30.44	56.84	84.36	98.92	100.00	30.24	56.48	84.08	98.84	100.00
	600	31.68	56.04	84.64	99.20	100.00	31.44	55.68	84.56	99.20	100.00
	60	32.52	47.48	66.72	87.56	96.84	16.24	25.00	40.44	66.32	85.44
	120	33.20	50.40	73.24	93.16	98.84	20.28	32.44	54.08	82.08	96.20
$\mathcal{PA}(9)$	240	32.16	52.40	78.60	95.96	99.72	23.24	39.76	66.64	92.28	99.04
	480	33.36	55.24	79.88	97.40	99.88	27.52	47.48	73.72	95.92	99.80
	600	33.48	55.40	81.92	98.08	99.92	28.92	48.52	77.04	97.08	99.80
	60	93.68	98.80	99.52	99.88	99.88	71.00	86.28	93.16	96.92	97.96
	120	97.44	99.44	99.84	99.96	99.96	80.80	92.32	96.24	98.16	99.08
$\mathcal{LN}(0,1)$	240	98.96	99.96	99.96	99.96	100.00	88.68	96.24	98.16	99.08	99.36
	480	99.56	100.00	100.00	100.00	100.00	93.28	97.44	98.76	99.36	99.64
	600	99.64	99.96	100.00	100.00	100.00	94.64	97.96	99.00	99.52	99.68

Scenario	$p \mid n$	ST_{CJ} 50	80	120	180	240	ST_{BCS} 50	80	120	180	240
	60	35.44	56.40	82.12	97.92	100.00	20.32	32.72	48.28	61.28	64.20
	120	33.92	57.16	83.16	98.56	99.88	27.80	51.60	77.72	95.32	97.32
$\mathcal{GA}(4, \sqrt{2}/2)$	240	34.64	57.20	84.60	99.24	99.96	29.84	53.16	83.44	99.00	99.92
	480	35.04	58.56	85.88	99.08	100.00	29.76	55.92	83.60	98.88	99.96
	600	35.64	58.92	85.64	99.20	100.00	30.68	55.12	83.88	99.12	100
	60	44.76	56.68	71.64	89.76	97.08	35.88	59.12	82.88	98.12	99.96
	120	47.80	60.68	77.88	94.48	99.04	35.36	55.92	81.40	97.52	99.84
$\mathcal{PA}(9)$	240	49.64	61.44	80.52	96.04	99.64	32.76	55.36	83.20	98.32	99.92
	480	49.56	63.00	80.32	97.20	99.92	32.64	56.52	83.48	98.56	100.00
	600	50.92	64.80	83.16	97.52	99.92	33.24	56.04	83.68	98.44	99.96
	60	100.00	100.00	100.00	100.00	100.00	6.92	7.00	6.60	6.68	6.56
	120	100.00	100.00	100.00	100.00	100.00	9.08	8.44	7.40	6.96	7.36
$\mathcal{LN}(0,1)$	240	100.00	100.00	100.00	100.00	100.00	10.24	10.76	10.04	9.24	9.96
	480	100.00	100.00	100.00	100.00	100.00	12.16	13.20	12.32	12.36	11.80
	600	100.00	100.00	100.00	100.00	100.00	13.08	13.72	14.88	15.36	14.08

* ST_{CZZ}, the sphericity test of Chen et al. [4]; ST_{VC}, the variance-corrected sphericity test proposed in this paper; ST_{CJ}, the corrected John's test of Wang and Yao [17]; ST_{BCS}, the bias-corrected sign test of Zou et al. [20].

When the neighbor correlation is present, as we can see from Table 6, ST_{VC} is generally consistent. It is worth noting that in the log-normal case, ST_{BCS} suffers not only from an incorrect empirical size under $\tilde{\mathcal{H}}_0$ but also from a low empirical power under the alternative considered.

4. Conclusion

In this paper, motivated by the finding that CZZ's identity test and sphericity test suffer from severe size distortion when the data are heavy-tailed, we offered a theoretical analysis for this phenomenon. We showed that inappropriately omitting the $O(p^{-1})$ bias terms in the variances of CZZ's test statistics is the main source of size distortion. Based on the analysis, we constructed two new tests called variance-corrected tests in this paper. They are simple extensions to CZZ's tests, but our simulation results suggest that they remarkably outperform the tests of Chen et al. [4], the corrected John's test of Wang and Yao [17], and the bias-corrected sign test of Zou et al. [20].

Acknowledgments. The author would like to thank two anonymous referees, an Associate Editor, and the Editor-in-Chief Professor Christian Genest for providing valuable comments. The author also would like to express his deep gratitude to Professor Christian Genest for his help in polishing the paper. The present research was supported by the National Natural Science Foundation of China (Grant No. 71601016) and the Humanities and Social Sciences Youth Foundation of Ministry of Education of China (Grant No. 16YJC790074).

Appendix A. Proofs

Before proving our main results, we first introduce a useful lemma.

Lemma 1. *Under Assumption 1, we have (i)* $\mathrm{E}(z_1^\top z_2)^4 = p\kappa^2 + 3p(p-1)$*; (ii)* $\mathrm{E}\{(z_1^\top z_2)^2(z_2^\top z_3)^2\} = p\kappa + p(p-1)$*; (iii)* $\mathrm{E}\{(z_1^\top z_2)^2(z_3^\top z_4)^2\} = p^2$*; (iv)* $\mathrm{E}(z_1^\top z_2)^3 = p\xi^2$*; (v)* $\mathrm{E}(z_1^\top z_2 z_3^\top z_4 z_2^\top z_3 z_4^\top z_1) = p$*; (vi)* $\mathrm{E}\{z_1^\top z_1(z_2^\top z_3)^2\} = p^2$*; (vii)* $\mathrm{E}\{z_1^\top z_1(z_1^\top z_2)^2\} = p\kappa + p(p-1)$.

Proof. (i) Since $\mathrm{E}(z_1^\top z_2)^4 = \sum_{u=1}^p \sum_{v=1}^p \sum_{s=1}^p \sum_{t=1}^p \mathrm{E}(z_{u1}z_{v1}z_{s1}z_{t1})\mathrm{E}(z_{u2}z_{v2}z_{s2}z_{t2})$, we find

$$\mathrm{E}(z_1^\top z_2)^4 = \sum_{u=1}^p \mathrm{E}(z_{u1}^4)\mathrm{E}(z_{u2}^4) + 3\sum_{u,v}^{*} \mathrm{E}(z_{u1}^2 z_{v1}^2)\mathrm{E}(z_{u2}^2 z_{v2}^2) = p\kappa^2 + 3p(p-1).$$

(ii) Because

$$\mathrm{E}(z_1^\top z_1)^2 = \sum_{u=1}^p \sum_{v=1}^p \mathrm{E}(z_{u1}^2 z_{v1}^2) = \sum_{u=1}^p \mathrm{E}(z_{u1}^4) + \sum_{u,v}^{*} \mathrm{E}(z_{u1}^2 z_{v1}^2) = p\kappa + p(p-1), \tag{A.1}$$

we have

$$\begin{aligned}\mathrm{E}\{(z_1^\top z_2)^2(z_2^\top z_3)^2\} &= \mathrm{E}(z_2^\top z_1 z_1^\top z_2 z_2^\top z_3 z_3^\top z_2) = \mathrm{E}\{\mathrm{tr}(z_2 z_2^\top z_3 z_3^\top z_2 z_2^\top)\}\\ &= \mathrm{E}\{\mathrm{tr}(z_2 z_2^\top z_2 z_2^\top)\} = \mathrm{E}(z_2^\top z_2)^2 = p\kappa + p(p-1).\end{aligned}$$

(iii) Since

$$\mathrm{E}(z_1^\top z_2)^2 = \mathrm{E}\{\mathrm{tr}(z_2 z_2^\top z_1 z_1^\top)\} = p, \tag{A.2}$$

we have

$$\mathrm{E}\{(z_1^\top z_2)^2(z_3^\top z_4)^2\} = \mathrm{E}(z_1^\top z_2)^2\mathrm{E}(z_3^\top z_4)^2 = p^2.$$

(iv) Since $\mathrm{E}(z_1^\top z_2)^3 = \sum_{u=1}^p \sum_{v=1}^p \sum_{s=1}^p \mathrm{E}(z_{u1}z_{v1}z_{s1})\mathrm{E}(z_{u2}z_{v2}z_{s2})$, we have

$$\mathrm{E}(z_1^\top z_2)^3 = \sum_{u=1}^p \mathrm{E}(z_{u1}^3)\mathrm{E}(z_{u2}^3) = p\xi^2. \tag{A.3}$$

(v) Using (A.2),

$$\mathrm{E}(z_1^\top z_2 z_3^\top z_4 z_2^\top z_3 z_4^\top z_1) = \mathrm{E}\{\mathrm{tr}(z_2 z_3^\top z_4 z_2^\top z_3 z_4^\top)\} = \mathrm{E}(z_2^\top z_3 z_3^\top z_4 z_4^\top z_2) = \mathrm{E}(z_3^\top z_4)^2 = p.$$

(vi) Using (A.2) again, we find that $\mathrm{E}\{z_1^\top z_1(z_2^\top z_3)^2\} = \mathrm{E}(z_1^\top z_1)\mathrm{E}(z_2^\top z_3)^2 = p^2$.

(vii) It follows from (A.1) that $\mathrm{E}\{z_1^\top z_1(z_1^\top z_2)^2\} = \mathrm{E}(z_1^\top z_1 z_1^\top z_2 z_2^\top z_1) = \mathrm{E}(z_1^\top z_1)^2 = p\kappa + p(p-1)$.

This concludes the proof of Lemma 1. □

The following lemma is instrumental in computing the variance of $T_{2,n}$.

Lemma 2. *Under Assumption 1 and $\tilde{\mathcal{H}}_0$, (i) $\mathrm{var}(Y_{2,n}) = 2\sigma^8 n_1^{-1} p\{2(\kappa-1)n + 2p + \kappa^2 - 4\kappa + 1\}$; (ii) $\mathrm{var}(Y_{4,n}) = 2\sigma^8 n_2^{-1} p(n+p+\kappa+2\xi^2-4)$; (iii) $\mathrm{var}(Y_{5,n}) = 8\sigma^8 n_3^{-1} p(p+2)$; (iv) $\mathrm{cov}(Y_{2,n}, Y_{4,n}) = 2\sigma^8 n_1^{-1} p\xi^2$; (v) $\mathrm{cov}(Y_{2,n}, Y_{5,n}) = 0$; (vi) $\mathrm{cov}(Y_{4,n}, Y_{5,n}) = 0$;*

Proof. (i) Since

$$\begin{aligned}
\mathrm{var}\Big\{\sum_{i,j}^{*}(z_i^\top z_j)^2\Big\} &= \mathrm{E}\Big\{\sum_{i,j}^{*}(z_i^\top z_j)^2\Big\}^2 - \Big\{\sum_{i,j}^{*}\mathrm{E}(z_i^\top z_j)^2\Big\}^2 \\
&= \mathrm{E}\Big\{\sum_{i,j}^{*}\sum_{k,\ell}^{*}(z_i^\top z_j)^2(z_k^\top z_\ell)^2\Big\} - \{n_1\mathrm{E}(z_1^\top z_2)^2\}^2 \\
&= 2\sum_{i,j}^{*}\mathrm{E}(z_i^\top z_j)^4 + 4\sum_{i,j,k}^{*}\mathrm{E}\{(z_i^\top z_j)^2(z_j^\top z_k)^2\} + \sum_{i,j,k,\ell}^{*}\mathrm{E}\{(z_i^\top z_j)^2(z_k^\top z_\ell)^2\} - \{n_1\mathrm{E}(z_1^\top z_2)^2\}^2 \\
&= 2n_1\mathrm{E}(z_1^\top z_2)^4 + 4n_2\mathrm{E}\{(z_1^\top z_2)^2(z_2^\top z_3)^2\} + n_3\mathrm{E}\{(z_1^\top z_2)^2(z_3^\top z_4)^2\} - \{n_1\mathrm{E}(z_1^\top z_2)^2\}^2,
\end{aligned}$$

we can use (i)–(iii) in Lemma 1 and (A.2) to conclude that

$$\begin{aligned}
\mathrm{var}(Y_{2,n}) &= \mathrm{var}\Big\{\frac{1}{n_1}\sum_{i,j}^{*}(x_i^\top x_j)^2\Big\} = \frac{\sigma^8}{n_1^2}\mathrm{var}\Big\{\sum_{i,j}^{*}(z_i^\top z_j)^2\Big\} \\
&= \frac{\sigma^8 p}{n_1}\{2\kappa^2 + 6p - 6 + 4(n-2)\kappa + 4(n-2)(p-1) + (n-2)(n-3)p - n_1 p\} \\
&= \frac{2\sigma^8 p}{n_1}\{2(\kappa-1)n + 2p + \kappa^2 - 4\kappa + 1\}.
\end{aligned}$$

(ii) Since $\mathrm{E}(\sum_{i,j,k}^{*} z_i^\top z_j z_j^\top z_k) = 0$, one has $\mathrm{var}\{\sum_{i,j,k}^{*} z_i^\top z_j z_j^\top z_k\} = \mathrm{E}\{\sum_{i,j,k}^{*} z_i^\top z_j z_j^\top z_k\}^2$. Note that when i, j and k are mutually different, $\mathrm{E}(z_1^\top z_2 z_2^\top z_3 z_i^\top z_j z_j^\top z_k) \neq 0$ only if (i,j,k) falls into the following eight cases:

$$\mathrm{E}(z_1^\top z_2 z_2^\top z_3 z_i^\top z_j z_j^\top z_k) = \begin{cases} \mathrm{E}(z_1^\top z_2 z_2^\top z_3)^2 & \text{if } (i,j,k) = (1,2,3) \text{ or } (3,2,1); \\ \mathrm{E}(z_1^\top z_2)^3 & \text{if } (i,j,k) = (1,3,2),\ (2,3,1),\ (2,1,3) \text{ or } (3,1,2); \\ p & \text{if } (i,k) = (1,3) \text{ or } (3,1), \text{ and } j \neq 2. \end{cases}$$

Therefore,

$$\begin{aligned}
\mathrm{var}\Big(\sum_{i,j,k}^{*} z_i^\top z_j z_j^\top z_k\Big) &= 2\sum_{i,j,k}^{*}\mathrm{E}(z_i^\top z_j z_j^\top z_k)^2 + 4\sum_{i,j,k}^{*}\mathrm{E}(z_i^\top z_j z_j^\top z_k z_j^\top z_k z_k^\top z_i) + 2\sum_{i,j,k,\ell}^{*}\mathrm{E}(z_i^\top z_j z_j^\top z_k z_i^\top z_\ell z_\ell^\top z_k) \\
&= 2n_2\mathrm{E}\{(z_1^\top z_2)^2(z_2^\top z_3)^2\} + 4n_2\mathrm{E}(z_1^\top z_2)^3 + 2n_3 p.
\end{aligned}$$

In view of (ii) and (iv) in Lemma 1,

$$\mathrm{var}(Y_{4,n}) = \mathrm{var}\Big(\frac{1}{n_2}\sum_{i,j,k}^{*} x_i^\top x_j x_j^\top x_k\Big) = \frac{\sigma^8}{n_2^2}\mathrm{var}\Big(\sum_{i,j,k}^{*} z_i^\top z_j z_j^\top z_k\Big) = \frac{2\sigma^8 p}{n_2}(n + p + \kappa + 2\xi^2 - 4).$$

(iii) In a similar way, we can verify that

$$\begin{aligned}
\operatorname{var}\Big(\sum_{i,j,k,\ell}^{*} z_i^\top z_j z_k^\top z_\ell\Big) &= \mathrm{E}\Big(\sum_{i,j,k,\ell}^{*} z_i^\top z_j z_k^\top z_\ell\Big)^2 \\
&= 8\sum_{i,j,k,\ell}^{*} \mathrm{E}(z_i^\top z_j z_k^\top z_\ell)^2 + 16\sum_{i,j,k,\ell}^{*} \mathrm{E}(z_i^\top z_j z_k^\top z_\ell z_j^\top z_k z_\ell^\top z_i) \\
&= 8n_3 \mathrm{E}\{(z_1^\top z_2)^2 (z_3^\top z_4)^2\} + 16 n_3 \mathrm{E}(z_1^\top z_2 z_3^\top z_4 z_2^\top z_3 z_4^\top z_1).
\end{aligned}$$

By (ii) and (v) in Lemma 1, we can then conclude that

$$\operatorname{var}(Y_{5,n}) = \operatorname{var}\Big(\frac{1}{n_3}\sum_{i,j,k,\ell}^{*} x_i^\top x_j x_k^\top x_\ell\Big) = \frac{\sigma^8}{n_3^2}\operatorname{var}\Big(\sum_{i,j,k,\ell}^{*} z_i^\top z_j z_k^\top z_\ell\Big) = \frac{8\sigma^8 p(p+2)}{n_3}.$$

(iv) Note that $\operatorname{cov}(Y_{2,n}, Y_{4,n}) = \mathrm{E}(Y_{2,n} Y_{4,n})$ due to $\mathrm{E}(Y_{4,n}) = 0$. Since

$$\begin{aligned}
\sum_{i_0,j_0}^{*}\sum_{i_1,j_1,k_1}^{*} \mathrm{E}\{(z_{i_0}^\top z_{j_0})^2 z_{i_1}^\top z_{j_1} z_{j_1}^\top z_{k_1}\} &= 2\sum_{i_0,j_0,j_1}^{*} \mathrm{E}\{(z_{i_0}^\top z_{j_0})^2 z_{i_0}^\top z_{j_1} z_{j_1}^\top z_{j_0}\} \\
&= 2n_2 \mathrm{E}\{(z_1^\top z_2)^2 z_1^\top z_3 z_3^\top z_2\} = 2n_2 \mathrm{E}(z_1^\top z_2)^3,
\end{aligned}$$

using (A.3), we find

$$\operatorname{cov}(Y_{2,n}, Y_{4,n}) = \frac{\sigma^8}{n_1 n_2}\sum_{i_0,j_0}^{*}\sum_{i_1,j_1,k_1}^{*} \mathrm{E}\{(z_{i_0}^\top z_{j_0})^2 z_{i_1}^\top z_{j_1} z_{j_1}^\top z_{k_1}\} = \frac{2\sigma^8 p\xi^2}{n_1}.$$

(v) Because $\mathrm{E}(Y_{5,n}) = 0$, we have

$$\operatorname{cov}(Y_{2,n}, Y_{5,n}) = \frac{\sigma^8}{n_1 n_3}\sum_{i_0,j_0}^{*}\sum_{i_1,j_1,k_1,\ell_1}^{*} \mathrm{E}\{(z_{i_0}^\top z_{j_0})^2 z_{i_1}^\top z_{j_1} z_{k_1}^\top z_{\ell_1}\} = 0.$$

(vi) It is easy to see that

$$\operatorname{cov}(Y_{4,n}, Y_{5,n}) = \frac{\sigma^8}{n_2 n_3}\sum_{i_0,j_0,k_0}^{*}\sum_{i_1,j_1,k_1,\ell_1}^{*} \mathrm{E}(z_{i_0}^\top z_{j_0} z_{j_0}^\top z_{k_0} z_{i_1}^\top z_{j_1} z_{k_1}^\top z_{\ell_1}) = 0.$$

This concludes the proof of Lemma 2. □

Lemma 3. *(i)* $\operatorname{cov}(Y_{1,n}, Y_{2,n}) = 2\sigma^6 n^{-1} p(\kappa - 1)$*; (ii)* $\operatorname{cov}(Y_{1,n}, Y_{4,n}) = 0$*; (iii)* $\operatorname{cov}(Y_{1,n}, Y_{5,n}) = 0$*; (iv)* $\operatorname{cov}(Y_{3,n}, Y_{2,n}) = 2\sigma^6 n_1^{-1} p\xi^2$*; (v)* $\operatorname{cov}(Y_{3,n}, Y_{4,n}) = 2\sigma^6 n_1^{-1} p$*; (vi)* $\operatorname{cov}(Y_{3,n}, Y_{5,n}) = 0$*;*

Proof. (i) Since $\mathrm{E}(Y_{1,n}) = p\sigma^2$, $\mathrm{E}(Y_{2,n}) = p\sigma^4$, and according to (vi) and (vii) in Lemma 1,

$$\begin{aligned}
\mathrm{E}\Big\{\sum_{i=1}^{n}\sum_{j,k}^{*} z_i^\top z_i (z_j^\top z_k)^2\Big\} &= \sum_{i,j,k}^{*} \mathrm{E}\{z_i^\top z_i (z_j^\top z_k)^2\} + 2\sum_{i,j}^{*} \mathrm{E}\{z_i^\top z_i (z_i^\top z_j)^2\} \\
&= n_2 \mathrm{E}\{z_1^\top z_1 (z_2^\top z_3)^2\} + 2n_1 \mathrm{E}\{z_1^\top z_1 (z_1^\top z_2)^2\} \\
&= n_1 p(np + 2\kappa - 2).
\end{aligned}$$

Thus we have

$$\operatorname{cov}(Y_{1,n}, Y_{2,n}) = \frac{\sigma^6}{n n_1}\mathrm{E}\Big\{\sum_{i=1}^{n}\sum_{j,k}^{*} z_i^\top z_i (z_j^\top z_k)^2\Big\} - p^2\sigma^6 = \frac{2\sigma^6 p(\kappa - 1)}{n}.$$

(ii) Since $\mathrm{E}(Y_{4,n}) = 0$ and $\mathrm{E}(x_i^\top x_i x_j^\top x_k x_k^\top x_\ell) = 0$ for mutually different j, k and ℓ,

$$\mathrm{cov}(Y_{1,n}, Y_{4,n}) = \frac{1}{nn_2} \sum_{i=1}^{n} \sum_{j,k,k}^{*} \mathrm{E}(x_i^\top x_i x_j^\top x_k x_k^\top x_\ell) - \mathrm{E}(Y_{1,n})\mathrm{E}(Y_{4,n}) = 0.$$

(iii) Proceeding as in (ii), we can easily verify that $\mathrm{cov}(Y_{1,n}, Y_{5,n}) = 0$.

(iv) Since $\mathrm{E}(Y_{3,n}) = 0$ and by (A.3),

$$\mathrm{E}\Big\{\sum_{i,j}^{*} \sum_{k,\ell}^{*} z_i^\top z_j (z_k^\top z_\ell)^2\Big\} = 2 \sum_{i,j}^{*} \mathrm{E}(z_i^\top z_j)^3 = 2n_1 p\xi^2.$$

Thus we have

$$\mathrm{cov}(Y_{3,n}, Y_{2,n}) = \frac{\sigma^6}{n_1^2} \mathrm{E}\Big\{\sum_{i,j}^{*} \sum_{k,\ell}^{*} z_i^\top z_j (z_k^\top z_\ell)^2\Big\} = \frac{2\sigma^6 p\xi^2}{n_1}.$$

(v) Using (A.2),

$$\begin{aligned}\mathrm{cov}(Y_{3,n}, Y_{4,n}) &= \frac{\sigma^6}{n_1 n_2} \sum_{i_0,j_0}^{*} \sum_{i_1,j_1,k_1}^{*} \mathrm{E}(z_{i_0}^\top z_{j_0} z_{i_1}^\top z_{j_1} z_{j_1}^\top z_{k_1}) = \frac{2\sigma^6}{n_1 n_2} \sum_{i_1,j_1,k_1}^{*} \mathrm{E}(z_{i_1}^\top z_{k_1} z_{i_1}^\top z_{j_1} z_{j_1}^\top z_{k_1}) \\ &= \frac{2\sigma^6}{n_1} \mathrm{E}(z_1^\top z_2 z_1^\top z_3 z_3^\top z_2) = \frac{2\sigma^6}{n_1} \mathrm{E}(z_1^\top z_2)^2 = \frac{2\sigma^6 p}{n_1}.\end{aligned}$$

(vi) It is easy to see $\mathrm{cov}(Y_{3,n}, Y_{5,n}) = 0$.

This concludes the proof of Lemma 3. □

Now, we are in a position to show Proposition 1.

Proof of Proposition 1. (i) It is straightforward to verify $\mathrm{var}(Y_{1,n}) = \sigma^4 p(\kappa - 1)/n$ and $\mathrm{var}(Y_{3,n}) = 2\sigma^4 p/n_1$ by (i) and (iii) in Proposition A.2 in Chen et al. [4]. The result follows immediately from (v) in that proposition, which shows $\mathrm{cov}(Y_{1,n}, Y_{3,n}) = 0$.

(ii) In view of (v) and (vi) in Lemma 2,

$$\mathrm{var}(T_{2,n}) = \mathrm{var}(Y_{2,n}) + 4\mathrm{var}(Y_{4,n}) + \mathrm{var}(Y_{5,n}) - 4\mathrm{cov}(Y_{2,n}, Y_{4,n}).$$

Thus, the result can be directly verified by (i)–(iv) of Lemma 2.

(iii) According to (ii), (iii) and (vi) in Lemma 3,

$$\mathrm{cov}(T_{1,n}, T_{2,n}) = \mathrm{cov}(Y_{1,n}, Y_{2,n}) - \mathrm{cov}(Y_{3,n}, Y_{2,n}) + 2\mathrm{cov}(Y_{3,n}, Y_{4,n}).$$

The result follows from (i), (iv), and (v) of Lemma 3. □

Finally, we prove Theorem 1.

Proof of Theorem 1. (i) The result can be directly shown by letting $\sigma^2 = 1$ and using Proposition 1.

(ii) Let $a_{i0} = \lim_{p\to\infty} \mathrm{tr}(\Sigma_p^i)/p$ for $i \in \{1, 2\}$. We know that $T_{1,n}/p \xrightarrow{p} a_{10}$ and $T_{2,n}/p \xrightarrow{p} a_{20}$ in terms of Chen et al. [4]. Then, by the delta method,

$$\begin{aligned} nU_n &= n\left\{\frac{T_{2,n}/p}{(T_{1,n})^2/p} - 1\right\} \approx n\left(\frac{a_{20}}{a_{10}^2} - 1\right) + \frac{1}{a_{10}^2} n\left(\frac{1}{p}T_{2,n} - a_{20}\right) - 2\left(\frac{a_{20}}{a_{10}^3}\right) n\left(\frac{1}{p}T_{1,n} - a_{10}\right) \\ &\approx n\left\{\frac{1}{\sigma^4}\left(\frac{1}{p}T_{2,n} - \sigma^4\right) - 2\left(\frac{1}{\sigma^2}\right)\left(\frac{1}{p}T_{1,n} - \sigma^2\right)\right\}.\end{aligned}$$

As a consequence,

$$\operatorname{var}(nU_n) \approx n^2 \left\{ \frac{1}{\sigma^8 p^2} \operatorname{var}(T_{2,n}) + \frac{4}{\sigma^4 p^2} \operatorname{var}(T_{1,n}) - \frac{4}{\sigma^6 p^2} \operatorname{cov}(T_{2,n}, T_{1,n}) \right\},$$

which is identical $\operatorname{var}(nV_n)$ under $\mathcal{H}_0$. $\square$

References

[1] Z. Bai, H. Saranadasa, Effects of high dimension: By an example of a two sample problem, Statist. Sinica 6 (1996) 311–329.
[2] Z. Bai, D. Jiang, J. Yao, S. Zheng, Corrections to LRT on large-dimensional covariance matrix by RMT, Ann. Statist. 37 (2009) 3822–3840.
[3] T.T. Cai, T. Jiang, Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices, Ann. Statist. 39 (2011) 1496–1525.
[4] S.X. Chen, L. Zhang, P. Zhong, Tests for high-dimensional covariance matrices, J. Amer. Statist. Assoc. 105 (2010) 810–819.
[5] T.J. Fisher, On testing for an identity covariance matrix when the dimensionality equals or exceeds the sample size, J. Statist. Plann. Inference 142 (2012) 312–326.
[6] T.J. Fisher, X. Sun, C.M. Gallagher, A new test for sphericity of the covariance matrix for high dimensional data, J. Multivariate Anal. 101 (2012) 2554–2570.
[7] J. He, S.X. Chen, Testing super-diagonal structure in high dimensional covariance matrices, J. Econometrics 194 (2016) 283–297.
[8] S. Jiang, S. Wang, Moderate deviation principles for classical likelihood ratio tests of high-dimensional normal distributions, J. Multivariate Anal. 156 (2017) 57–69.
[9] O. Ledoit, M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compare to the sample size, Ann. Statist. 30 (2002) 1081–1102.
[10] G. Mao, A note on tests for high-dimensional covariance matrices, Statist. Probab. Lett. 117 (2016) 89–92.
[11] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[12] Y. Qiu, S.X. Chen, Test for bandedness of high-dimensional covariance matrices and bandwidth estimation, Ann. Statist. 40 (2012) 1285–1314.
[13] M.S. Srivastava, Some tests concerning the covariance matrix in high dimensional data, J. Japan Statist. Soc. 35 (2005) 251–272.
[14] M.S. Srivastava, T. Kollo, D. von Rosen, Some tests for the covariance matrix with fewer observations than the dimension under non-normality, J. Multivariate Anal. 102 (2011) 1090–1103.
[15] M.S. Srivastava, T. Kubokawa, Tests for multivariate analysis of variance in high dimension under non-normality, J. Multivariate Anal. 115 (2013) 204–216.
[16] M.S. Srivastava, H. Yanagihara, T. Kubokawa, Tests for covariance matrices in high dimension with less sample size, J. Multivariate Anal. 130 (2014) 289–309.
[17] Q. Wang, J. Yao, On the sphericity test with large-dimensional observations, Electron. J. Statist. 7 (2013) 2164–2192.
[18] R. Zhang, L. Peng, R. Wang, Tests for covariance matrix with fixed or divergent dimension, Ann. Statist. 41 (2013) 2075–2096.
[19] S. Zheng, Z. Bai, J. Yao, Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing, Ann. Statist. 43 (2015) 546–591.
[20] C. Zou, L. Peng, L. Feng, Z. Wang, Multivariate sign-based high-dimensional tests for sphericity, Biometrika 101 (2014) 229–236.