ORIGINAL ARTICLE

Journal Section

Testing Equality of Two High-dimensional Spatial Sign Covariance Matrices

Guanghui Cheng[1] | **Baisen Liu**[2] | **Liuhua Peng**[3] | **Baoxue Zhang**[4*] | **Shurong Zheng**[5*]

[1]School of Economics and Statistics, Guangzhou University, Guangdong Engineering and Technology Research Center of Intelligent Finance, Accounting & Taxation, China

[2]School of Statistics, Dongbei University of Finance and Economics, China

[3]School of Mathematics and Statistics, University of Melbourne, Australia

[4]School of Statistics, Capital University of Economics and Business, China

[5]KLAS and School of Mathematics and Statistics, Northeast Normal University, China

Correspondence
KLAS and School of Mathematics and Statistics, Northeast Normal University, China
Email: zhengsr@nenu.edu.cn

Funding information
NSFC 11671268, 11522105 and 11690012, Department of Education of Liaoning Province (NO. LN2017ZD001).

This paper is concerned with testing the equality of two high-dimensional spatial sign covariance matrices with applications to testing the proportionality of two high-dimensional covariance matrices. It is interesting that these two testing problems are completely equivalent for the class of elliptically symmetric distributions. This paper develops a new test for testing the equality of two high-dimensional spatial sign covariance matrices based on the Frobenius norm of the difference between two spatial sign covariance matrices. Asymptotic normality of the proposed testing statistic is derived under the null and alternative hypotheses when the dimension and sample sizes both tend to infinity. Moreover, the asymptotic power function is also presented. Simulation studies show that the proposed test performs very well in a wide range of settings and can be allowed to the case of large dimension and small sample sizes.

KEYWORDS
Elliptically symmetric distribution; high dimension; spatial sign covariance matrix; U-statistic.

1 | INTRODUCTION

Consider two populations $\mathbf{X}$ and $\mathbf{Y}$ with p-dimensional mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and $p \times p$-dimensional covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, respectively. Many interests have been focused on learning the relationship between $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. One of them is to test

TABLE 1 Reviews for developed tests for proportionality of covariance matrices on different population distributions.

Method	p	distributions
Flury (1986)	fixed	two multivariate normal populations
Eriksen (1987)	fixed	$k \geq 2$ multivariate normal populations
Schott (1999)	fixed	$k \geq 2$ populations with finite fourth moments
Xu et al. (2014)	$p/n_l \in (0,1), l = 1,2$	two populations with finite fourth moments
Liu et al. (2014)	$p/n_1 \in (0,\infty)$, $p/n_2 \in (0,1)$	two populations with finite fourth moments

where n_1 and n_2 are sample sizes of two populations.

the following hypothesis

$$H_0 : \boldsymbol{\Sigma}_1 = c\boldsymbol{\Sigma}_2, \qquad \text{versus} \qquad H_1 : \boldsymbol{\Sigma}_1 \neq c\boldsymbol{\Sigma}_2, \tag{1.1}$$

for some unknown scalar $c > 0$. It is interesting that for elliptically symmetric distributions, testing the proportionality $\boldsymbol{\Sigma}_1 = c\boldsymbol{\Sigma}_2$ is completely equivalent to testing $\mathbf{S}_1 = \mathbf{S}_2$ where $\mathbf{S}_1 = \mathrm{E}[\|\mathbf{X} - \boldsymbol{\mu}_1\|^{-2}(\mathbf{X} - \boldsymbol{\mu}_1)(\mathbf{X} - \boldsymbol{\mu}_1)^T]$ and $\mathbf{S}_2 = \mathrm{E}[\|\mathbf{Y} - \boldsymbol{\mu}_2\|^{-2}(\mathbf{Y} - \boldsymbol{\mu}_2)(\mathbf{Y} - \boldsymbol{\mu}_2)^T]$ are two spatial sign covariance matrices with $\|\cdot\|$ denoting the L_2-norm. The equivalence will be proved in the Section 2.2. The target of this paper is to develop a novel nonparametric procedure to test

$$H_0^{'} : \mathbf{S}_1 = \mathbf{S}_2, \qquad \text{versus} \qquad H_1^{'} : \mathbf{S}_1 \neq \mathbf{S}_2, \tag{1.2}$$

with applications to test the proportionality $H_0 : \boldsymbol{\Sigma}_1 = c\boldsymbol{\Sigma}_2$ for elliptically symmetric distributions.

There are rare literatures about testing the equality of high-dimensional spatial sign covariance matrices but the literatures for testing the proportionality of two high-dimensional covariance matrices are rich. The proportionality test of (1.1) is widely used in various areas, for example, discriminant analysis, principal components analysis (Flury and Riedwyl, 1988; Schott, 1991), and it is also very attractive in economics and genetics of these scientific fields (Flury, 1986; Jensen and Madsen, 2004). Since Federer (1951) first studied the proportional relationship of two covariance matrices with dimension $p \leq 3$, a number of methods have been developed on the proportionality test of covariance matrices (for example, Kim, 1971; Rao, 1983; Flury, 1986; Eriksen, 1987; Schott, 1999). However, all of these tests are based on the classical limit theorem which assumes that the sample size tends to infinity but the dimension p is fixed. With rapid development and wide applications of computer techniques, the dimension p is no longer very small with respect to the sample size and even diverges to infinity as the sample size goes to infinity. This often leads to the classical tests performing poorly, being unstable for large p and even being unavailable when p is larger than $n_1 + n_2$ (Bai and Saranadasa, 1996). Thus, it is very urgent to develop some new methods to deal with the proportionality test of (1.1) for high-dimensional data. Assuming that the dimension p increases proportionally with the sample size n_l, i.e., $p/n_l \to q_l \in (0,1), l = 1,2$, Xu et al. (2014) developed a pseudo-likelihood ratio test (PLRT) which extended the traditional likelihood ratio test, and established its asymptotic normality under the existence of finite fourth population moments. Moreover, Liu et al. (2014) proposed a method which allowed $p/n_1 \to q_1 \in (0,\infty)$ and showed that the test statistic satisfied asymptotic normality. These methods are summarized in Table 1.

A shared drawback of Xu et al. (2014) and Liu et al. (2014) is that both of them have unsatisfactory size and power performance when the population distribution is heavy-tailed (see Table 3 in Section 4). Moreover, when $p > \max(n_1, n_2)$, they cannot be defined due to the non-existence of the inverse of sample covariance matrix.

In multivariate data analysis, it is often assumed the observations are from a Gaussian population. However, it is well-known that, in real data analysis, the observed data seldom follows a Gaussian distribution but often displays some non-normal features, like heavy-tails and skewness etc. For example, it is observed that the marginal distributions of the micro-array expressions are non-normal even after log transformation and have heavy tails based on values of their marginal kurtosises in the gene expression data (Purdom and Holmes, 2005; Wang, Peng and Li, 2015). Furthermore, "outliers in micro-array expression data frequently arise due to the array chip artifacts such as uneven spray of reagents within arrays" (Wang, Peng and Li, 2015). A good alternative approach is to consider a more flexible class of distributions, like multivariate t-distribution which has been shown being more robust than multivariate normal distribution. This motivates us to consider a class of elliptically symmetric distributions which includes multivariate Gaussian distribution, multivariate t-distribution, multivariate logistic distribution and Pearson II type multivariate distribution as special ones.

In this paper, we propose a new high-dimensional test for the hypothesis (1.2) based on sample spatial sign covariance matrices. On the one hand, the proposed test can also be used to test the hypothesis (1.1) for elliptically symmetric distributions. On the other hand, the proposed test is robust against the heavy-tailed data and can be allowed to the case of $p > \max(n_1, n_2)$. The proposed testing statistic is asymptotically normal under the null and alternative hypotheses. Simulation studies show that our procedure can keep the empirical sizes well at the nominal size and achieve great empirical powers for a wide range of dimensions, sample sizes and population distributions.

The rest of this paper is organized as follows. We first prove the equivalence between testing the equality of two spatial sign covariance matrices and testing the proportionality of two covariance matrices for elliptically symmetric distributions in Section 2. Section 3 introduces our high-dimensional nonparametric test statistic and its limiting distributions under the null and alternative hypotheses. Monte Carlo simulation studies are conducted in Section 4. Section 5 includes conclusions and discussions. Technical proofs are given in the Appendix.

2 | EQUIVALENCE OF EQUALITY OF TWO SPATIAL SIGN COVARIANCE MATRICES AND PROPORTIONALITY OF TWO COVARIANCE MATRICES FOR ELLIPTICALLY SYMMETRIC DISTRIBUTIONS

2.1 | Review of elliptically symmetric distributions

In this section, we briefly review some theories of elliptically symmetric distributions. A p-dimensional random vector $\mathbf{X}$ is said to follow an elliptically symmetric distribution if it has the following stochastic representation (Fang, Kotz and Ng, 1990)

$$\mathbf{X} = \boldsymbol{\mu} + \xi \boldsymbol{\Gamma} \mathbf{W},$$

where $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{\Gamma}$ is a non-random and invertible $p \times p$-dimensional matrix, ξ is a univariate nonnegative random variable and $\mathbf{W}$ is a p-dimensional random vector from a uniform distribution on the unit sphere in R^p with ξ and $\mathbf{W}$ being independent, $\mathrm{E}(\mathbf{W}\mathbf{W}^T) = p^{-1}\mathbf{I}_p$ and $\mathbf{I}_p$ is the $p \times p$ dimensional identity matrix. The covariance matrix $\boldsymbol{\Sigma}$ and shape matrix $\boldsymbol{\Lambda}$ of the elliptical symmetric population $\mathbf{X}$ satisfies

$$\boldsymbol{\Sigma} = p^{-1}\mathrm{E}(\xi^2) \cdot \boldsymbol{\Lambda}, \ \ \boldsymbol{\Lambda} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T, \ \ \boldsymbol{\Sigma} \propto \boldsymbol{\Lambda},$$

where A^T denotes the transpose of a vector or a matrix A. The class of elliptical distributions includes many useful multivariate distributions, such as multivariate Gaussian distribution, multivariate t-distribution, multivariate logistic distribution, Pearson II type multivariate distribution and so on.

2.2 | Equivalence of two testing hypotheses (1.1) and (1.2) for elliptically symmetric distributions

Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are from two p-dimensional elliptically symmetric populations with mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, respectively. The samples $\mathbf{x}_1, ..., \mathbf{x}_{n_1}$ are independent and identically distributed (i.i.d) from $\mathbf{X}$ and samples $\mathbf{y}_1, ..., \mathbf{y}_{n_2}$ are i.i.d from $\mathbf{Y}$. The shape matrices are $\boldsymbol{\Lambda}_l = c_{lp}^{-1}\boldsymbol{\Sigma}_l$ for some scalar $c_{lp} > 0$, $l = 1, 2$. Without loss of generality, assume that $\mathrm{tr}(\boldsymbol{\Lambda}_1) = \mathrm{tr}(\boldsymbol{\Lambda}_2) = p$ where $\mathrm{tr}(\cdot)$ denotes the trace of a square matrix. For a given vector $\mathbf{x}$, the corresponding multivariate spatial sign function is defined by $U(\mathbf{x}) = \|\mathbf{x}\|^{-1}\mathbf{x}$ where $\|\mathbf{x}\|$ is the L_2-norm of $\mathbf{x}$. Let

$$\mathbf{u}_i = U(\mathbf{x}_i - \boldsymbol{\mu}_1) \ \text{ and } \ \mathbf{v}_j = U(\mathbf{y}_j - \boldsymbol{\mu}_2), i = 1, \dots n_1, j = 1, \dots n_2,$$

be the spatial sign functions of the centralized $\mathbf{x}_i$ and $\mathbf{y}_j$, respectively. When the center parameters $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are known, the corresponding sample spatial sign covariance matrices are defined as

$$\mathbf{S}_{1n_1} = n_1^{-1}\sum_{i=1}^{n_1}\mathbf{u}_i\mathbf{u}_i^T \ \text{ and } \ \mathbf{S}_{2n_2} = n_2^{-1}\sum_{j=1}^{n_2}\mathbf{v}_j\mathbf{v}_j^T,$$

respectively with $\mathbf{S}_l = \mathrm{E}(\mathbf{S}_{ln_l}), l = 1, 2$. About the relationship of $\boldsymbol{\Sigma}_l$ and $\mathbf{S}_l$, there is a very useful lemma which is from Magyar and Tyler (2014).

Lemma 2.1 *(see Eq. 3.9 of Magyar and Tyler, 2014) For the p-dimensional elliptically symmetric population* $\mathbf{X}$*, let* $\lambda_1 \geq \cdots \geq \lambda_p$ *be the eigenvalues of* $\boldsymbol{\Sigma}_1$ *and* $\psi_1 \geq \cdots \geq \psi_p$ *be the eigenvalues of* $\mathbf{S}_1$*. Then we have*

$$\psi_h = \mathrm{E}\left(\lambda_h\chi^2_{1,h} / \sum_{r=1}^{p}\lambda_r\chi^2_{1,r}\right), \quad h = 1, \dots, p,$$

where $\chi^2_{1,1}, \dots, \chi^2_{1,p}$ *are mutually independent* χ^2*-variates with degrees of freedom 1.*

Lemma 2.1 shows that, when the eigenvalues of the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are proportional, the spatial sign covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ have the same eigenvalues. Similarly, we are fortunate to obtain that, when the spatial sign covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ have the same eigenvalues, the eigenvalues of the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are proportional.

Theorem 2.1 *For two p-dimensional elliptically symmetric populations* $\mathbf{X}$ *and* $\mathbf{Y}$*, if* $\psi_h = \psi_h'$*,* $h = 1, \dots, p$*, then we have*

$$\lambda_1/\lambda_1' = \ldots = \lambda_p/\lambda_p',$$

where $\psi_1 \geq \cdots \geq \psi_p$ *and* $\lambda_1 \geq \cdots \geq \lambda_p$ *are the eigenvalues of* $\mathbf{S}_1$ *and* $\boldsymbol{\Sigma}_1$*, and* $\psi_1' \geq \cdots \geq \psi_p'$ *and* $\lambda_1' \geq \cdots \geq \lambda_p'$ *are the eigenvalues of* $\mathbf{S}_2$ *and* $\boldsymbol{\Sigma}_2$*, respectively.*

The proof of Theorem 2.1 is placed in the Appendix.

From Section 3.1 of Magyar and Tyler (2014), we know that $\boldsymbol{\Sigma}_l$ and $\mathbf{S}_l$ have the same eigenvectors, $l = 1, 2$ for elliptically symmetric populations. Then by Lemma 2.1 and Theorem 2.1, we know that the null hypothesis (1.1) is completely equivalent to the null hypothesis $H_0' : \mathbf{S}_1 = \mathbf{S}_2$, that is,

$$H_0 : \boldsymbol{\Sigma}_1 = c\boldsymbol{\Sigma}_2 \Longleftrightarrow H_0' : \mathbf{S}_1 = \mathbf{S}_2.$$

Thus, testing the proportionality of two covariance matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ is equivalent to testing the equality of two spatial sign covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ for elliptically symmetric distributions.

3 | TESTING EQUALITY OF TWO HIGH-DIMENSIONAL SPATIAL SIGN COVARIANCE MATRICES FOR ELLIPTICAL SYMMETRIC DISTRIBUTIONS

To test the equality of two high-dimensional spatial sign covariance matrices, we require two assumptions.

Assumption 3.1 *Assume that the moments* $\mathrm{E}(R_{1i}^{-4})$ *and* $\mathrm{E}(R_{2j}^{-4})$ *exist for* $1 \le i \le n_1, 1 \le j \le n_2$ *with* $R_{1i} = \|\mathbf{x}_i - \boldsymbol{\mu}_1\|$ *and* $R_{2j} = \|\mathbf{y}_j - \boldsymbol{\mu}_2\|$. *Moreover, for* $2 \le k \le 4$, *we have* $\mathrm{E}(R_{1i}^{-k})/\mathrm{E}^k(R_{1i}^{-1}) \to d_{1k} \in [1, +\infty)$ *and* $\mathrm{E}(R_{2j}^{-k})/\mathrm{E}^k(R_{2j}^{-1}) \to d_{2k} \in [1, +\infty)$ *as* $p \to \infty$ *where* d_{1k} *and* d_{2k} *are constants.*

Assumption 3.2 *Assume* $n_l/(n_1+n_2) \to \eta_l \in (0,1)$ *as* $\min\{n_1, n_2\} \to \infty$, *and* $p = O(n_l^2)$, *for* $l = 1, 2$. *Moreover, as* $p \to \infty$, *assume* $\mathrm{tr}(\mathbf{\Sigma}_{l_1}\mathbf{\Sigma}_{l_2}\mathbf{\Sigma}_{l_3}\mathbf{\Sigma}_{l_4}) = o(\mathrm{tr}(\mathbf{\Sigma}_{l_1}\mathbf{\Sigma}_{l_2})\mathrm{tr}(\mathbf{\Sigma}_{l_3}\mathbf{\Sigma}_{l_4}))$ *for* $l_1, l_2, l_3, l_4 \in \{1, 2\}$.

For simplicity, we let $n = n_1 + n_2$. Similar to Zou et al. (2014), Assumption 3.1 is a necessary condition to ensure the validity of the second-order expansions. Assumption 3.2 is similar to Li and Chen (2012). When all the eigenvalues of both $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are bounded, Assumption 3.2 holds. But for the matrix having the maximum eigenvalue being $O(p)$, e.g., the compound symmetric matrix, Assumption 3.2 does not hold. Moreover, Assumption 3.2 requires that the convergence order between dimension and sample sizes, that is, $p = O(n_l^2)$ for $l = 1, 2$.

3.1 | Testing statistic

The Frobenius norm for any $m \times n$ matrix D is defined as $||D||_F^2 = \mathrm{tr}(D^T D)$. It is known that $\mathbf{S}_{ln_l}$ is a sample version of $\mathbf{S}_l$ for $l = 1, 2$. This motivates us to construct a random variable to measure the difference between $\mathbf{S}_{1n_1}$ and $\mathbf{S}_{2n_2}$ under Frobenius norm as follows

$$T_{n_1,n_2}^* = p\mathrm{tr}(\mathbf{S}_{1n_1} - \mathbf{S}_{2n_2})^2 = p\left\{ n_1^{-1} + n_1^{-2}\sum_{i\neq i'}^{n_1}(\mathbf{u}_i^T\mathbf{u}_{i'})^2 + n_2^{-1} + n_2^{-2}\sum_{j\neq j'}^{n_2}(\mathbf{v}_j^T\mathbf{v}_{j'})^2 - 2n_1^{-1}n_2^{-1}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(\mathbf{u}_i^T\mathbf{v}_j)^2 \right\}.$$

Ignoring the non-random terms n_1^{-1}, n_2^{-1}, using $\{n_1(n_1-1)\}^{-1}$, $\{n_2(n_2-1)\}^{-1}$ as a multiplier instead of n_1^{-2} and n_2^{-2} in the first two random terms, we arrive at a modified version of T_{n_1,n_2}^* as $T_{n_1,n_2} = p(A_{n_1} + B_{n_2} - 2C_{n_1,n_2})$ where

$$A_{n_1} = \frac{1}{n_1(n_1-1)}\sum_{i\neq i'}^{n_1}(\mathbf{u}_i^T\mathbf{u}_{i'})^2,\ B_{n_2} = \frac{1}{n_2(n_2-1)}\sum_{j\neq j'}^{n_2}(\mathbf{v}_j^T\mathbf{v}_{j'})^2,\ C_{n_1,n_2} = \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(\mathbf{u}_i^T\mathbf{v}_j)^2. \tag{3.1}$$

It is clear that $\mathrm{E}(T_{n_1,n_2}) = p\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$. In practice, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are unknown. By replacing $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ with the spatial median estimator $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ in T_{n_1,n_2} (Mottonen and Oja, 1995), we propose the statistic for testing the equality of two spatial sign covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ as

$$T_{n_1,n_2}' = \frac{p}{n_1(n_1-1)}\sum_{i\neq i'}^{n_1}(\hat{\mathbf{u}}_i^T\hat{\mathbf{u}}_{i'})^2 + \frac{p}{n_2(n_2-1)}\sum_{j\neq j'}^{n_2}(\hat{\mathbf{v}}_j^T\hat{\mathbf{v}}_{j'})^2 - \frac{2p}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(\hat{\mathbf{u}}_i^T\hat{\mathbf{v}}_j)^2,$$

where $\hat{\mathbf{u}}_i = U(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1), i = 1, ..., n_1$ and $\hat{\mathbf{v}}_j = U(\mathbf{y}_j - \hat{\boldsymbol{\mu}}_2), j = 1, ..., n_2$. Consider $\mathbf{r}_i = \mathcal{B}\mathbf{x}_i$, $\mathbf{w}_j = \mathcal{B}\mathbf{y}_j$, where $\mathcal{B}$ is an orthogonal matrix. Note that the spatial median $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ are orthogonal invariant, $\hat{\boldsymbol{\mu}}_r = \mathcal{B}\hat{\boldsymbol{\mu}}_1$, $\hat{\boldsymbol{\mu}}_w = \mathcal{B}\hat{\boldsymbol{\mu}}_2$. Then we have

$$\hat{\mathbf{u}}_{\mathbf{r}_i} = U(\mathbf{r}_i - \hat{\boldsymbol{\mu}}_r) = \mathcal{B}U(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1) = \mathcal{B}\hat{\mathbf{u}}_i,$$

and $(\hat{\mathbf{u}}_{\mathbf{r}_i}^T \hat{\mathbf{u}}_{\mathbf{r}_j})^2 = (\hat{\mathbf{u}}_i^T \mathcal{B}^T \mathcal{B}\hat{\mathbf{u}}_j)^2 = (\hat{\mathbf{u}}_i^T \hat{\mathbf{u}}_j)^2$, $(\hat{\mathbf{v}}_{\mathbf{w}_i}^T \hat{\mathbf{v}}_{\mathbf{w}_j})^2 = (\hat{\mathbf{v}}_i^T \hat{\mathbf{v}}_j)^2$, $(\hat{\mathbf{u}}_{\mathbf{r}_i}^T \hat{\mathbf{v}}_{\mathbf{w}_j})^2 = (\hat{\mathbf{u}}_i^T \hat{\mathbf{v}}_j)^2$. Hence, T'_{n_1,n_2} is also orthogonal invariant when $\mathbf{x}_i$ and $\mathbf{y}_j$ take the same orthogonal transformations.

3.2 | Limiting distributions of the testing statistic

To consider the bias of the testing statistic T'_{n_1,n_2}, the following lemma will first give the expectation of T'_{n_1,n_2}.

Lemma 3.1 *Under Assumptions 3.1–3.2 and elliptically symmetric distributions, as* $\min\{n_1, n_2\} \to \infty$ *and* $p \to \infty$ *we have*

$$\mathrm{E}(T'_{n_1,n_2}) = p\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2 + p\delta_{n_1,n_2}, \tag{3.2}$$

where

$$\begin{aligned}\delta_{n_1,n_2} \quad = \quad & n_1^{-2}\left\{2 - \frac{2\mathrm{E}(R_{11}^{-2})}{\mathrm{E}^2(R_{11}^{-1})} + \frac{\mathrm{E}^2(R_{11}^{-2})}{\mathrm{E}^4(R_{11}^{-1})}\right\} + n_2^{-2}\left\{2 - \frac{2\mathrm{E}(R_{2j}^{-2})}{\mathrm{E}^2(R_{21}^{-1})} + \frac{\mathrm{E}^2(R_{21}^{-2})}{\mathrm{E}^4(R_{21}^{-1})}\right\} \\ & + n_1^{-3}\left\{-\frac{6\mathrm{E}^2(R_{11}^{-2})}{\mathrm{E}^4(R_{11}^{-1})} + \frac{2\mathrm{E}(R_{11}^{-2})\mathrm{E}(R_{11}^{-3})}{\mathrm{E}^5(R_{11}^{-1})} + \frac{8\mathrm{E}(R_{11}^{-2})}{\mathrm{E}^2(R_{11}^{-1})} - \frac{2\mathrm{E}(R_{11}^{-3})}{\mathrm{E}^3(R_{11}^{-1})}\right\} \\ & + n_2^{-3}\left\{-\frac{6\mathrm{E}^2(R_{21}^{-2})}{\mathrm{E}^4(R_{21}^{-1})} + \frac{2\mathrm{E}(R_{21}^{-2})\mathrm{E}(R_{21}^{-3})}{\mathrm{E}^5(R_{21}^{-1})} + \frac{8\mathrm{E}(R_{21}^{-2})}{\mathrm{E}^2(R_{21}^{-1})} - \frac{2\mathrm{E}(R_{21}^{-3})}{\mathrm{E}^3(R_{21}^{-1})}\right\}.\end{aligned} \tag{3.3}$$

The proof of Lemma 3.1 is provided in the supplementary material.

Then we obtain the central limit theorem of the proposed testing statistic under both null and alternative hypotheses in the following theorem.

Theorem 3.1 *Under Assumptions 3.1–3.2 and elliptically symmetric distributions, as* $\min\{n_1, n_2\} \to \infty$ *and* $p \to \infty$*, we have*
(1). $\upsilon_{n_1,n_2}^{-1}\{T'_{n_1,n_2} - p\delta_{n_1,n_2} - p\mathrm{tr}[(\mathbf{S}_1 - \mathbf{S}_2)^2]\} \xrightarrow{D} N(0,1)$ *where* $\xrightarrow{D}$ *denotes convergence in distribution and*

$$\begin{aligned}\upsilon_{n_1,n_2}^2 = & \frac{4}{n_1(n_1-1)}\frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_1^2)}{(p+2)^2} + \frac{8}{n_1}\frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1^4) - \mathrm{tr}^2(\boldsymbol{\Lambda}_1^2)}{p^2(p+2)} \\ & + \frac{4}{n_2(n_2-1)}\frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_2^2)}{(p+2)^2} + \frac{8}{n_2}\frac{p\mathrm{tr}(\boldsymbol{\Lambda}_2^4) - \mathrm{tr}^2(\boldsymbol{\Lambda}_2^2)}{p^2(p+2)} \\ & + \frac{8}{n_1 n_2}\frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)}{(p+2)^2} + \left(\frac{8}{n_1} + \frac{8}{n_2}\right)\frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)^2 - \mathrm{tr}^2(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)}{p^2(p+2)} \\ & - \frac{16}{n_1}\frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1^3\boldsymbol{\Lambda}_2) - \mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)\mathrm{tr}(\boldsymbol{\Lambda}_1^2)}{p^2(p+2)} - \frac{16}{n_2}\frac{p\mathrm{tr}(\boldsymbol{\Lambda}_2^3\boldsymbol{\Lambda}_1) - \mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)\mathrm{tr}(\boldsymbol{\Lambda}_2^2)}{p^2(p+2)};\end{aligned}$$

(2). Under $H_0 : \mathbf{S}_1 = \mathbf{S}_2$*, we have* $\upsilon_{0,n_1,n_2}^{-1}\{T'_{n_1,n_2} - p\delta_{n_1,n_2}\} \xrightarrow{D} N(0,1)$ *where* δ_{n_1,n_2} *is in (3.3), and* $\upsilon_{0,n_1,n_2}^2 = 4(n_1^{-1} + n_2^{-1})^2(p+2)^{-2}\mathrm{tr}^2(\boldsymbol{\Lambda}^2)$ *with* $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2$ *under* H_0.

Remark 3.1 *To formulate our testing procedure, we need to estimate the unknown quantities in δ_{n_1,n_2} and υ_{0,n_1,n_2} under $H_0 : \mathbf{S}_1 = \mathbf{S}_2$. The moment estimate of $\mathrm{E}(R_{l1}^{-k})$ is $\hat{\mathrm{E}}(R_{l1}^{-k}) = n_l^{-1}\sum_{i=1}^{n_l}\hat{R}_{li}^{-k}$ for $l = 1, 2$ with $\hat{R}_{1i} = \|\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1\|$ and $\hat{R}_{2i} = \|\mathbf{y}_i - \hat{\boldsymbol{\mu}}_2\|$. In the Supplementary Material (page 6), we proved that*

$$p\mathrm{tr}(\mathbf{S}_l^2) = p^{-1}\mathrm{tr}(\boldsymbol{\Lambda}_l^2)\{1 + o(1)\}, l = 1, 2.$$

Under H_0, we have $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2$. A straightforward approach to estimate $p^{-1}\mathrm{tr}(\boldsymbol{\Lambda}^2)$ is

$$p^{-1}\widehat{\mathrm{tr}(\boldsymbol{\Lambda}^2)} = p(n_1 + n_2)^{-1}\{n_1(A_{n_1}^{'} - \hat{\delta}_{n_1}) + n_2(B_{n_1}^{'} - \hat{\delta}_{n_2})\},$$

where $A_{n_1}^{'} = n_1^{-1}(n_1 - 1)^{-1}\sum_{i\neq i'}(\hat{\mathbf{u}}_i^T\hat{\mathbf{u}}_{i'})^2$, $B_{n_2}^{'} = n_2(n_2 - 1)^{-1}\sum_{j\neq j'}(\hat{\mathbf{v}}_j^T\hat{\mathbf{v}}_{j'})^2$ and

$$\hat{\delta}_{n_l} = \frac{1}{n_l^2}\left\{2 - \frac{2\hat{\mathrm{E}}(R_{l1}^{-2})}{\hat{\mathrm{E}}^2(R_{l1}^{-1})} + \frac{\hat{\mathrm{E}}^2(R_{l1}^{-2})}{\hat{\mathrm{E}}^4(R_{l1}^{-1})}\right\} + \frac{1}{n_l^3}\left\{-\frac{6\hat{\mathrm{E}}^2(R_{l1}^{-2})}{\hat{\mathrm{E}}^4(R_{l1}^{-1})} + \frac{2\hat{\mathrm{E}}(R_{l1}^{-2})\hat{\mathrm{E}}(R_{l1}^{-3})}{\hat{\mathrm{E}}^5(R_{l1}^{-1})} + \frac{8\hat{\mathrm{E}}(R_{l1}^{-2})}{\hat{\mathrm{E}}^2(R_{l1}^{-1})} - \frac{2\hat{\mathrm{E}}(R_{l1}^{-3})}{\hat{\mathrm{E}}^3(R_{l1}^{-1})}\right\},$$

for $l = 1, 2$. By replacing $p^{-1}\mathrm{tr}(\boldsymbol{\Lambda}^2)$ and $\mathrm{E}(R_{l1}^{-k})$ by $p^{-1}\widehat{\mathrm{tr}(\boldsymbol{\Lambda}^2)}$ and $\hat{\mathrm{E}}(R_{l1}^{-k})$ in υ_{0,n_1,n_2}^2 and δ_{n_1,n_2}, we have $\hat{\upsilon}_{0,n_1,n_2}^2$ and $\hat{\delta}_{n_1,n_2}$.

The next theorem shows that the estimator $\hat{\upsilon}_{0,n_1,n_2}^2$ is ratio consistent. And the proof is given in the Appendix.

Theorem 3.2 *Under Assumptions 3.1–3.2, the null hypothesis (1.2) and elliptically symmetric distributions, as $\min\{n_1, n_2\} \to \infty$ and $p \to \infty$, we have*

$$\hat{\upsilon}_{0,n_1,n_2}/\upsilon_{0,n_1,n_2} \xrightarrow{P} 1,$$

where $\xrightarrow{P}$ denotes convergence in probability.

Remark 3.2 *Then by the Slutsky's Theorem, $\hat{\upsilon}_{0,n_1,n_2}^{-1}(T_{n_1,n_2}^{'} - p\hat{\delta}_{n_1,n_2}) \xrightarrow{D} N(0, 1)$, as long as $\hat{\upsilon}_{0,n_1,n_2}^{-1}p(\hat{\delta}_{n_1,n_2} - \delta_{n_1,n_2}) = o_p(1)$. It can be easily shown that $n_l^{-1}\sum_{i=1}^{n_1} R_{li}{}^{-k} = E(R_{li}^{-k})\{1 + O_p(n^{-1/2})\}$, for $l = 1, 2$. Then we know that $\upsilon_{0,n_1,n_2}^{-1}p(\hat{\delta}_{n_1,n_2} - \delta_{n_1,n_2}) = o_p(1)$ is valid for $p = o(n_l^{3/2})$ $(l = 1, 2)$ when $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ have bounded eigenvalues. Details can be found in the Appendix.*

By Theorems 3.1 and 3.2, as $\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2 > 0$ under the alternative, then the acceptance region at the test level $100\alpha\%$ is constructed as follows

$$\{(\mathbf{x}_1, \dots, \mathbf{x}_{n_1}; \mathbf{y}_1, \dots, \mathbf{y}_{n_2}) : \hat{\upsilon}_{0,n_1,n_2}^{-1}(T_{n_1,n_2}^{'} - p\hat{\delta}_{n_1,n_2}) \leq z_\alpha\},$$

where z_α is the upper-α quantile of $N(0, 1)$.

Then we wish to study the power performance for $\{T_{n_1,n_2}^{'} - p\delta_{n_1,n_2}\}$, the asymptotic power function of $\{T_{n_1,n_2}^{'} - p\delta_{n_1,n_2}\}$ can be written as

$$\beta_{n_1,n_2}(\mathbf{S}_1, \mathbf{S}_2, \boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \alpha) = \Phi\{-\mathcal{L}_{n_1,n_2}(\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2)z_\alpha + p\upsilon_{n_1,n_2}^{-1}\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2\},$$

where $\mathcal{L}_{n_1,n_2}(\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2) = \upsilon_{n_1,n_2}^{-1}\hat{\upsilon}_{0,n_1,n_2}$ and $\Phi(\cdot)$ denotes the cumulative probability distribution of $N(0, 1)$. The following theorem gives the consistency of the proposed test. And the proof is given in the Appendix.

Theorem 3.3 *Under Assumption 3.1–3.2, as* $\min\{n_1, n_2\} \to \infty, p \to \infty$*, if* $(n_1 + n_2)p\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ *has larger order than* $\{p^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + p^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)\}$*, there is*

$$\beta_{n_1,n_2}(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{S}_1, \mathbf{S}_2, \alpha) \to 1.$$

Moreover, if both $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are positive definite and have bounded eigenvalues, then $\{p^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + p^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)\} = O(1)$. Thus, Theorem 3.3 ensures that the proposed test is consistent only if $p^{-1}\mathrm{tr}(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2$ is not shrinking faster than n^{-1}, as $p\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ has the same order as $p^{-1}\mathrm{tr}(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2$.

4 | SIMULATION STUDIES

In this section, we perform some simulation studies to evaluate the finite sample properties of our testing method for the hypothesis (1.1) in the high-dimensional case. For comparisons, we also conduct the Bartlett adjusted likelihood ratio test (Flury, 1986), Wald test (Schott, 1999), PLRT (Xu et al., 2014), and LZ test (Liu et al., 2014). With each method, the data are generated from the following two scenarios:

- *Scenario 1 (Multivariate Gaussian distribution)*: Samples $\mathbf{x}_1, ..., \mathbf{x}_{n_1}$ are from $N_p(\mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{y}_1, ..., \mathbf{y}_{n_2}$ are from $N_p(\mathbf{0}, \mathbf{\Sigma}_2)$;
- *Scenario 2 (Multivariate t distribution)*: Samples $\mathbf{x}_1, ..., \mathbf{x}_{n_1}$ are from $t_6(\mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{y}_1, ..., \mathbf{y}_{n_2}$ are from $t_6(\mathbf{0}, \mathbf{\Sigma}_2)$.

It is noted that the above two distributions are elliptically distributed. We choose $\mathbf{\Sigma}_1 = \mathbf{I}_p$ and $\mathbf{\Sigma}_2 = c(\rho^{|i-j|})_{i,j=1}^p$ with $c = 1, 2$, $\rho = 0$ for computing empirical sizes and with different $\rho \neq 0$ for computing empirical powers. The combination of (n_1, n_2) is set to be $(50, 100), (100, 100), (100, 200)$. The nominal significant level is set as $\alpha = 5\%$ and all empirical test sizes and empirical powers are computed based on 5,000 simulations.

Table 2 for Scenario 1 reports the simulation results of Gaussian data for $p = 40, 80$. With large p, both Wald test and Bartlett adjusted likelihood ratio test have large empirical sizes which indicate that these two methods do not work any more for high-dimensional data. Besides, compared to the PLRT and LZ test, our proposed method also has good size performance. To study the robustness of our proposed test on non-normal data, we go on to perform Scenario 2 for multivariate t distribution and the simulation results are displayed in Table 3. All of other tests except our proposed test have unsatisfactorily performance since the empirical sizes are far away from the nominal level 5%. As there is no exsiting method for large $p > \max(n_1, n_2)$, then we choose $p = 400, 800$ to handle our proposed method, we can see that the proposed method has satisfactory size performance for $\rho = 0$ and also has nontrivial powers for a range of c and ρ under both normal distribution and t distribution in Table 4-5.

FIGURE 1 shows the density plot for the 2000 observed values of the proposed method with $\rho = 0.0$ and $c = 2$. In both cases, $p < \min(n_1, n_2)$ and $p > \max(n_1, n_2)$, the normality result appears to be satisfied by the density-plots for large (p, n_1, n_2), which validates our theoretical asymptotic normality results, therefore it is also a good evidence for why the empirical sizes of the proposed method perform very well.

For comparing empirical powers, as the dimension p increases, both Wald test and Bartlett test lead to an increasing test size. Thus we only compared with LZ, PLRT and our proposed method. It can be seen that our proposed method has the greater powers even though the empirical sizes of these three methods are all so close to the nominal level $\alpha = 0.05$ when the sample sizes become greater. Moreover, fixing the sample sizes n_1 and n_2, the powers of PLRT and LZ test decrease but the powers of our proposed test seem to be stable when p becomes greater, these results are illustrated in **FIGURE 2.** All of these simulation results show that our proposed test is valid, robust, especially for heavy-tailed or skewed data.

TABLE 2 Simulation results (in percentage) for Scenario 1 based on 5,000 independent replications.

			Empirical sizes $c = 1$					Empirical sizes $c = 2$				
p	n_1	n_2	Wald	Bartlett	PLRT	LZ	T'_{n_1,n_2}	Wald	Bartlett	PLRT	LZ	T'_{n_1,n_2}
40	50	100	0.66	23.95	5.08	5.98	4.90	6.02	22.92	5.78	6.46	5.36
	100	100	1.18	5.95	5.72	6.68	5.48	5.27	6.25	5.62	6.83	5.20
	100	200	2.43	6.66	5.44	6.34	5.04	6.31	6.36	5.40	5.61	5.59
80	50	100	-	-	-	7.44	4.76	-	-	-	6.92	4.89
	100	100	0.05	61.60	6.66	6.66	4.70	14.93	60.83	6.36	7.28	4.87
	100	200	0.65	51.82	4.70	6.12	5.22	13.35	50.81	5.19	5.79	5.23

TABLE 3 Simulation results (in percentage) for Scenario 2 based on 5,000 independent replications.

			Empirical sizes $c = 1$				Empirical sizes $c = 2$			
p	n_1	n_2	Wald	PLRT	LZ	T'_{n_1,n_2}	Wald	PLRT	LZ	T'_{n_1,n_2}
40	50	100	98.66	98.86	87.44	5.42	99.79	95.08	82.58	5.75
	100	100	99.97	99.74	85.74	4.92	99.97	98.49	80.77	4.82
	100	200	100	100	99.86	5.64	100	99.95	99.43	5.65
80	50	100	-	-	27.10	4.86	-	-	25.70	5.01
	100	100	100	97.30	19.18	4.86	100	93.32	18.87	4.91
	100	200	100	100	100	5.34	100	100	99.97	5.11

TABLE 4 The performance of the proposed test under Scenario 1 with a range of ρ and c.

p	n_1	n_2	$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.4$	
			$c = 1$	2	$c = 1$	2	$c = 1$	2
400	50	100	5.06	4.63	34.89	34.83	100	100
	100	100	5.01	4.67	62.49	62.60	100	100
	100	200	5.11	4.87	83.42	83.56	100	100
800	50	100	4.73	4.89	34.83	35.02	100	100
	100	100	4.71	4.81	63.35	63.57	100	100
	100	200	5.02	5.27	84.21	83.33	100	100

TABLE 5 The performance of the proposed test under Scenario 2 with a range of ρ and c.

p	n_1	n_2	$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.4$	
			$c = 1$	2	$c = 1$	2	$c = 1$	2
400	50	100	4.65	4.80	35.09	36.06	100	100
	100	100	4.79	4.73	62.86	62.47	100	100
	100	200	4.96	4.65	84.15	83.56	100	100
800	50	100	4.97	4.88	36.75	36.37	100	100
	100	100	4.72	4.89	62.58	64.23	100	100
	100	200	5.37	5.05	84.34	84.86	100	100

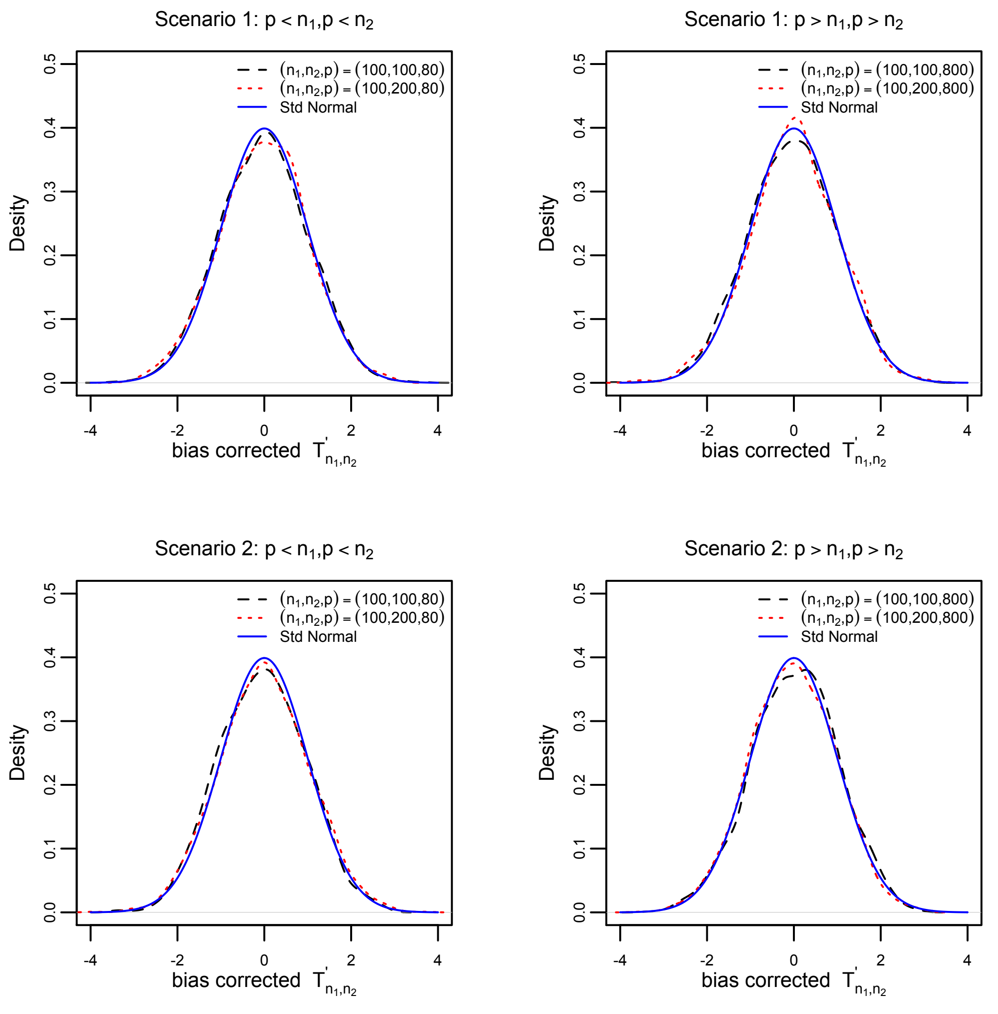


FIGURE 1 The null distributions of standardized and bias corrected $\{T'_{n_1,n_2} - p\delta_{n_1,n_2}\}$ with a range of p, n_1 and n_2 for $c = 2$.

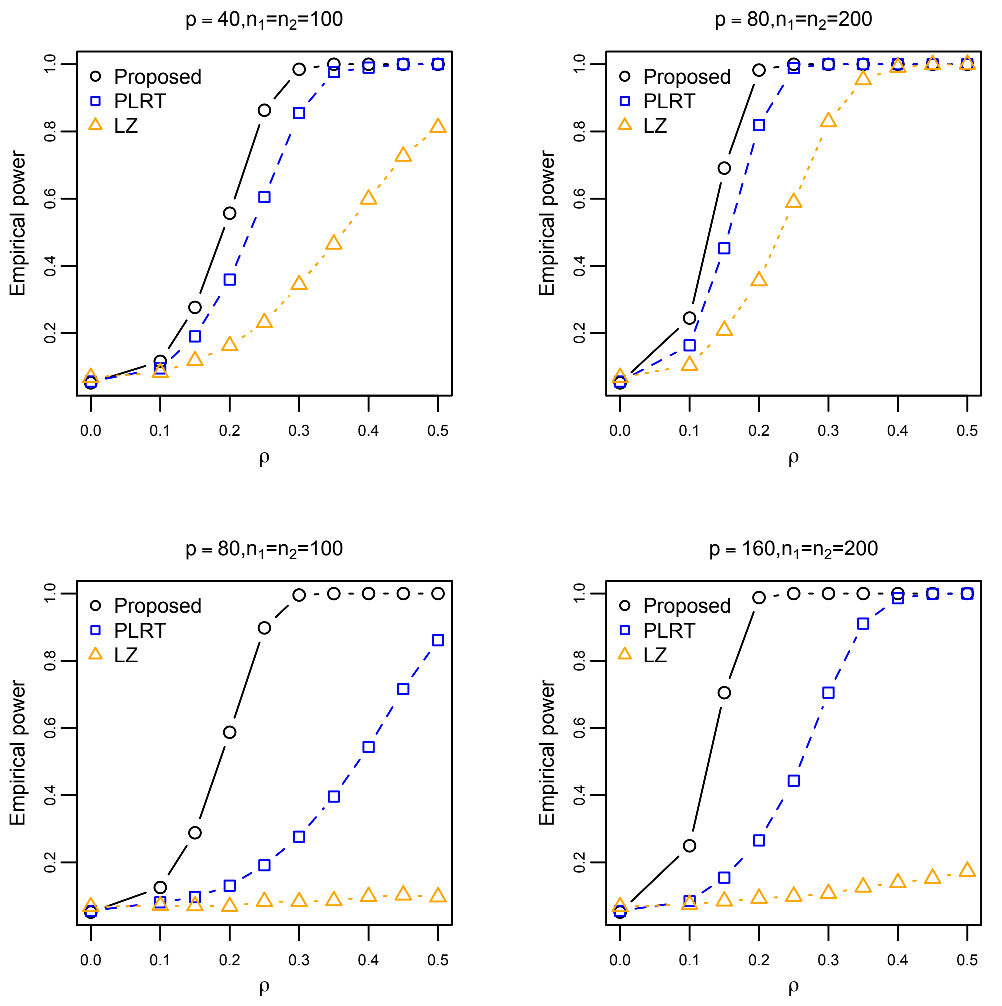


FIGURE 2 Power curves comparison between the PLRT, LZ and the proposed procedure with different p/n_1 , p/n_2 for normal distribution with $c = 2$.

5 | CONCLUSIONS AND DISCUSSIONS

This paper proposes a new nonparametric test based on spatial sign covariance matrix for testing the proportionality of two high-dimensional covariance matrices. Its asymptotic normality is established under both null and alternative hypotheses. The advantage of our proposed test is valid for some heavy-tail distributions, e.g. multivariate t distributions with respect to the dimension much larger than the sample sizes. One limitation of our method is, as did in the bias-correction procedure of Theorem 3.1 with $n = n_1 + n_2$, it requires that p is not growing fast than n^2, that is, $p = O(n^2)$ where $n = n_1 + n_2$. In addition, when δ_{n_1,n_2} in $\{T'_{n_1,n_2} - p\delta_{n_1,n_2}\}$ is replaced by the estimator $\hat{\hat{\delta}}_{n_1,n_2}$, the bias-corrected method is valid for $p = o(n^{3/2})$. In the future, we try to relax this limitation and extend our work to general cases.

6 | APPENDIX: TECHNICAL PROOFS

This section contains the main proofs of Theorem 2.1, Theorem 3.1, Theorem 3.2 and Theorem 3.3.

6.1 | Proof of Theorem 2.1

Suppose that $\psi_h = \psi'_h, h = 1, \dots, p$, that is,

$$\mathrm{E}\left(\frac{\lambda_h Z_h^2}{\sum_{k=1}^p \lambda_k Z_k^2}\right) = \mathrm{E}\left(\frac{\lambda'_h Z_h^2}{\sum_{k=1}^p \lambda'_k Z_k^2}\right), h = 1, ..., p, \tag{6.1}$$

where $Z_1^2, \dots, Z_p^2$ are mutually independent χ^2-variates with degrees of freedom 1.

(1). Thus, if there exists some $h \in \{1, \dots, p\}$ such that $\lambda_h = 0$, then we have $\lambda'_h = 0$.

Let $\rho_h = \lambda'_h/\lambda_h$ for some $\lambda_h \neq 0$ and $\lambda'_h \neq 0$. Denoting $\tilde{Z}_h^2 = \lambda_h Z_h^2$, then we have

$$\mathrm{E}\left(\frac{\tilde{Z}_h^2}{\sum_{k=1}^p \tilde{Z}_k^2}\right) = \mathrm{E}\left(\frac{\rho_h \tilde{Z}_h^2}{\sum_{k=1}^p \rho_k \tilde{Z}_k^2}\right), \quad 1 \leq h \leq p. \tag{6.2}$$

Let $\rho_1, \cdots, \rho_p$ be ordered as $\rho_{(1)} \leq \cdots \leq \rho_{(p)}$.

(2). It is clear that if $\rho_{(1)} = \rho_{(p)}$, then $\lambda_1/\lambda'_1 = ... = \lambda_p/\lambda'_p$.

(3). Suppose $\rho_{(1)} < \rho_{(p)}$ and k_0 satisfying $\rho_{k_0} = \rho_{(p)}$, then we have $\sum_{k=1}^p \rho_k \tilde{Z}_k^2 < \rho_{k_0} \sum_{k=1}^p \tilde{Z}_k^2$. Thus

$$\mathrm{E}\left(\frac{\rho_{k_0} \tilde{Z}_{k_0}^2}{\sum_{k=1}^p \rho_k \tilde{Z}_k^2}\right) > \mathrm{E}\left(\frac{\rho_{k_0} \tilde{Z}_{k_0}^2}{\rho_{k_0} \sum_{k=1}^p \tilde{Z}_k^2}\right) = \mathrm{E}\left(\frac{\tilde{Z}_{k_0}^2}{\sum_{k=1}^p \tilde{Z}_k^2}\right),$$

which contradicts with (6.2). Therefore we have $\rho_{(1)} = \rho_{(p)}$, that is, $\lambda_1/\lambda'_1 = ... = \lambda_p/\lambda'_p$.

6.2 | Proof of Theorem 3.1

Define a sequence of random variables $\{\mathbf{z}_1, ..., \mathbf{z}_{n_1+n_2}\}$ as follows

$$\mathbf{z}_i = \mathbf{u}_i, 1 \leq i \leq n_1, \quad \text{and} \quad \mathbf{z}_{n_1+j} = \mathbf{v}_j, 1 \leq j \leq n_2.$$

Let $\mathrm{E}_k(\cdot)$ denote the conditional expectation conditional on $\{\mathbf{z}_1, \dots, \mathbf{z}_k\}$. Define $D_{n,k} = p^{-1}\{\mathrm{E}_k(T_{n_1,n_2}) - \mathrm{E}_{k-1}(T_{n_1,n_2})\}$. Then it is easy to show that $p^{-1}\{T_{n_1,n_2} - E(T_{n_1,n_2})\} = \sum_{k=1}^{n_1+n_2} D_{n,k}$. Hence the sequence $\{D_{n,1}, \dots, D_{n,n_1+n_2}\}$ constitutes a martingale difference with respect to the σ-fields $\sigma(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)$. To derive martingale central limit theorem, we need the following lemma.

Lemma 6.1 *Under Assumption 3.2 and as* $\min\{n_1, n_2\} \to \infty$, *we have*

$$p^2 \sum_{k=1}^{n_1+n_2} \sigma_{n,k}^2 / \mathrm{Var}(T_{n_1,n_2}) \xrightarrow{p} 1 \text{ and } \sum_{k=1}^{n_1+n_2} \mathrm{E}(D_{nk}^4) = p^{-4} o\{\mathrm{Var}^2(T_{n_1,n_2})\}.$$

where $\sigma_{n,k}^2 = \mathrm{E}_{k-1}(D_{n,k}^2)$.

The above lemma shows the variance of martingale is convergent in probability and the Lindeberg condition can be established. The proof is placed in the Supplementary material.

In order to study the asymptotic normality of T'_{n_1,n_2}, as shown in the proof of Lemma 3.1,

$$\mathrm{E}(T'_{n_1,n_2}) = \mathrm{E}(T_{n_1,n_2}) + p\delta_{n_1,n_2} + o(pn^{-3}) + O(n^{-2}),$$

then we only need to study the asymptotic normality of the leading term in T_{n_1,n_2}.

In the Supplementary material, we also prove that

$$\begin{aligned}
\mathrm{Var}(T_{n_1,n_2}) =& p^2\{\mathrm{Var}(A_{n_1}) + \mathrm{Var}(B_{n_2}) + 4\mathrm{Var}(C_{n_1,n_2}) - 4\mathrm{Cov}(A_{n_1}, C_{n_1,n_2}) - 4\mathrm{Cov}(B_{n_2}, C_{n_1,n_2})\} \\
=& \Bigg\{ \frac{4}{n_1(n_1-1)} \frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_1^2)}{(p+2)^2} + \frac{8}{n_1} \frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1^4) - \mathrm{tr}^2(\boldsymbol{\Lambda}_1^2)}{p^2(p+2)} + \frac{4}{n_2(n_2-1)} \frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_2^2)}{(p+2)^2} \\
&+ \frac{8}{n_2} \frac{p\mathrm{tr}(\boldsymbol{\Lambda}_2^4) - \mathrm{tr}^2(\boldsymbol{\Lambda}_2^2)}{p^2(p+2)} - \frac{16}{n_1} \frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1^3\boldsymbol{\Lambda}_2) - \mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)\mathrm{tr}(\boldsymbol{\Lambda}_1^2)}{p^2(p+2)} \\
&+ \frac{8}{n_1 n_2} \frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)}{(p+2)^2} + (\frac{8}{n_1} + \frac{8}{n_2}) \frac{p\mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)^2 - \mathrm{tr}^2(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)}{p^2(p+2)} \\
&- \frac{16}{n_2} \frac{p\mathrm{tr}(\boldsymbol{\Lambda}_2^3\boldsymbol{\Lambda}_1) - \mathrm{tr}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)\mathrm{tr}(\boldsymbol{\Lambda}_2^2)}{p^2(p+2)} \Bigg\} (1 + o(1)).
\end{aligned}$$

Following Lemma 6.1 and applying the martingale central limit theorem (Hall and Hyde, 1980), we conclude that

$$\frac{T_{n_1,n_2} - \mathrm{E}(T_{n_1,n_2})}{\mathrm{Var}(T_{n_1,n_2})} \xrightarrow{d} N(0,1).$$

6.3 | Proof of Theorem 3.2

Recall that $\mathrm{E}(pA'_{n_1}) = p\mathrm{tr}(\mathbf{S}_1^2) + p\delta_{n_1} + o(n_1^{-1})$, as $\tilde{A}_{n_1} = A'_{n_1} - \delta_{n_1}$, then $\mathrm{E}(p\tilde{A}_{n_1}) = p\mathrm{tr}(\mathbf{S}_1^2) + o(n_1^{-1})$. Notice that

$$\mathrm{Var}(p^2\tilde{A}_{n_1}/\mathrm{tr}(\boldsymbol{\Lambda}_1^2)) = O\left[\frac{p^4}{\mathrm{tr}^2\boldsymbol{\Lambda}_1^2}\left\{\frac{2}{n_1(n_1-1)}\left(\frac{3\mathrm{tr}^2(\boldsymbol{\Lambda}_1^2) + 6\mathrm{tr}(\boldsymbol{\Lambda}_1^4)}{p^2(p+2)^2} - \frac{\mathrm{tr}^2(\boldsymbol{\Lambda}_1^2)}{p^4}\right)\right\}\right].$$

As $\mathrm{tr}(\boldsymbol{\Lambda}_1^4)/\mathrm{tr}^2(\boldsymbol{\Lambda}_1^2) \to 0$, hence $p^2\tilde{A}_{n_1}/\mathrm{tr}(\boldsymbol{\Lambda}_1^2) \xrightarrow{P} 1$. Moreover, under H_0, there are $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}$, $p^2\tilde{A}_{n_1}/\mathrm{tr}(\boldsymbol{\Lambda}^2) \xrightarrow{P} 1$. As

$$\hat{R}_{1i} - R_{1i} = \|\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1\| - \|\mathbf{x}_i\| \le \|\hat{\boldsymbol{\mu}}_1\| = O_p(c_1^{-1}n_1^{-1/2}),$$

we have

$$\hat{R}_{1i}^{-1} - R_{1i}^{-1} = \mathrm{E}(R_{1i}^{-1})(O_p(n_1^{-1/2})), \;\; n_1^{-1}\sum_{i=1}^{n_1}\hat{R}_{1i}^{-1} = n_1^{-1}\sum_{i=1}^{n_1} R_{1i}{}^{-1}\{1+O_p(n^{-1/2})\}.$$

Moreover, we have $n_1^{-1}\sum_{i=1}^{n_1} R_{1i}{}^{-k} = \mathrm{E}(R_{1i}^{-k})(1+O_p(n_1^{-1/2}))$ by the assumption $\mathrm{E}^2(R_{1i}^{-k})/\{\mathrm{E}(R_{1i}^{-k})\}^2 = O(1)$. Thus we have

$$n_1\sum_{i=1}^{n_1}\hat{R}_{1i}^{-2}/(\sum_{i=1}^{n_1}\hat{R}_{1i}^{-1})^2 = \mathrm{E}(R_{1i}^{-2})/\{\mathrm{E}(R_{1i}^{-1})\}^2\{1+O_p(n_1^{-1/2})\},$$

and

$$n_1^2\sum_{i=1}^{n_1}\hat{R}_{1i}^{-3}/(\sum_{i=1}^{n_1}\hat{R}_{1i}^{-1})^3 = \mathrm{E}(R_{1i}^{-3})/\{\mathrm{E}(R_{1i}^{-1})\}^3\{1+o_p(1)\}.$$

Therefore, $\hat{\delta}_{n_1}$ converges to δ_{n_1} in probability and $p\hat{\delta}_{n_1} = p\delta_{n_1}\{1+o_p(1)\}$ with $p = O(n_1^2)$. Then $p^2\tilde{A}_{n_1}^{'}/\mathrm{tr}(\mathbf{\Lambda}^2)\xrightarrow{P} 1$, where $\tilde{A}_{n_1}^{'} = A_{n_1}^{'} - \hat{\delta}_{n_1}$. Similarly, we have $p^2\tilde{B}_{n_2}/\mathrm{tr}(\mathbf{\Lambda}^2)\xrightarrow{P} 1$. Hence, we proved that $\hat{\upsilon}_{0,n_1,n_2}/\upsilon_{0,n_1,n_2}\xrightarrow{P} 1$.

Moreover, when $\hat{\upsilon}_{0,n_1,n_2}^{-1}p(\hat{\delta}_{n_1,n_2} - \delta_{n_1,n_2}) = o_p(1)$, we have $\hat{\upsilon}_{0,n_1,n_2}^{-1}(T_{n_1,n_2}^{'} - p\hat{\delta}_{n_1,n_2})\xrightarrow{D} N(0,1)$, note that $\upsilon_{0,n_1,n_2}^{-1}p(\hat{\delta}_{n_1,n_2} - \delta_{n_1,n_2}) = O(np\delta_{n_1,n_2}n^{-1/2})$. Thus, the test is valid for using $\hat{\delta}_{n_1,n_2}$ instead of δ_{n_1,n_2} in $\{T_{n_1,n_2}^{'} - p\delta_{n_1,n_2}\}$ with $p = o(n^{3/2})$.

6.4 | Proof of Theorem 3.3

First, we will show that $\mathcal{L}_{n_1,n_2}(\mathbf{\Lambda}_1,\mathbf{\Lambda}_2)$ is bounded. In fact, $\hat{\upsilon}_{0,n_1,n_2}$ is a ratio-consistent estimator of $2p^{-1}n_2^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + 2p^{-1}n_1^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)$. Hence we only need to show that $\upsilon_{n_1,n_2}^{-1}[2p^{-1}n_2^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + 2p^{-1}n_1^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)]$ is bounded. Notice that $\upsilon_{n_1,n_2}^2 \geq 4n_1^{-2}p^{-2}\mathrm{tr}^2(\mathbf{\Lambda}_1^2) + 4n_2^{-2}p^{-2}\mathrm{tr}^2(\mathbf{\Lambda}_2^2)$. Then we have

$$\mathcal{L}_{n_1,n_2}(\mathbf{\Lambda}_1,\mathbf{\Lambda}_2) \leq \frac{n_2^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + n_1^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)}{\sqrt{n_1^{-2}\mathrm{tr}^2(\mathbf{\Lambda}_1^2) + n_2^{-2}\mathrm{tr}^2(\mathbf{\Lambda}_2^2)}}.$$

Let $u_0 = \mathrm{tr}(\mathbf{\Lambda}_1^2)/\mathrm{tr}(\mathbf{\Lambda}_2^2)$. The right hand side in the above inequality can be written as a function of u_0, say, $h(u_0)$. It is easy to obtain that $h(u_0)$ is maximized at $u_0 = k_n/(1-k_n)$, where $k_n = n_1/(n_1+n_2)$. Then we have

$$\beta_{n_1,n_2}(\mathbf{\Lambda}_1,\mathbf{\Lambda}_2,\mathbf{S}_1,\mathbf{S}_2,\alpha) \geq \Phi\left(-z_\alpha k_n^{-1}(1-k_n)^{-1} + p\upsilon_{n_1,n_2}^{-1}\mathrm{tr}(\mathbf{S}_1-\mathbf{S}_2)^2\right),$$

where the right hand side indicates the lower bound of power function and also implies that the test is powerful when $p\mathrm{tr}(\mathbf{S}_1-\mathbf{S}_2)^2$ is at the same or a larger order of υ_{n_1,n_2}.

Under Assumption 3.2, as $k_n \to \eta_1$, to prove the consistency of the proposed test, it is only enough to prove $p\upsilon_{n_1,n_2}^{-1}\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ tending to infinity when both n and p tend to infinity. By the Cauchy-Schwarz inequality, we note that the variance of $\{T_{n_1,n_2}^{'} - p\delta_{n_1,n_2}\}$ can be dominated by

$$4p^{-2}\{n_1^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + n_2^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)\}^2 + 8p^{-2}\{n_1^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + n_2^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)\}\mathrm{tr}(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)^2.$$

Thus, the test statistic $\{T_{n_1,n_2}^{'} - p\delta_{n_1,n_2}\}$ has nontrivial power $p\mathrm{tr}(\mathbf{S}_1-\mathbf{S}_2)^2$ at least order of $n_1^{-1}p^{-1}\mathrm{tr}(\mathbf{\Lambda}_1^2) + n_2^{-1}p^{-1}\mathrm{tr}(\mathbf{\Lambda}_2^2)$.

The Supplementary Material Supplement to "Testing Equality of Two High dimensional Spatial Sign Covariance Matrices". It consists of the proofs of Lemma 3.1 and Lemma 6.1, and the derivation of $\mathrm{Var}(T_{n_1,n_2})$.

ACKNOWLEDGEMENTS

The research is in part supported by NSFC 11671258, 11522105 and 11690012. The second author's research is supported by Department of Education of Liaoning Province (NO. LN2017ZD001). We would also like to appreciate the reviewers for their constructive comments, which help to substantially improve this manuscript.

Eriksen, P.S. (1987). Proportionality of covariance matrices. *Ann. Statist.* **15,** 732-748.

Federer, W.T. (1951). Testing proportionality of covariance matrices. *Ann. Math. Statist.* **22**, 102-106.

Flury, B.K. (1986). Proportionality of *k* covariance matrices. *Statist. Prob. Lett.* **4**, 29-33.

Flury, B.K. & Riedwyl, H. (1988). Multivariate Statistics: A Practical Approach Chapman and Hall, London.

Fang, K. T., Kotz, S. & Ng, K. W. (1990). Symmetric Multivariate and Related Distributions. Chapman and Hall, London.

Jensen, S.T. & Madsen, J. (2004). Estimation of proportional covariances in the presence of certain linear restrictions. *Ann. Statist.* **32**, 219-232.

Kim, D.Y. (1971). Statistical inference for constants of proportionality between covariance matrices. Technical Report, **59**, Stanford University.

Li, J., & Chen, S. X. (2012). Two sample tests for high dimensional covariance matrices. *Ann. Statist.* **40** 908-940.

Liu, B., Xu, L., Zheng, S. & Tian, G. (2014). A new test for the proportionality of two large-dimensional covariance matrices. *Journal of Multivariate Analysis.* **131** 293-308.

Magyar, A. & Tyler, D. (2014). The asymptotic inadmissibility of the spatial sign covariance matrix for elliptically symmetric distributions. *Biometrika.* **3**, 673-688.

Mottonen J. & Oja, H. (1995). Multivariate spatial sign and rank methods. *J. Nonparam. Statist.* **5**, 201-213.

Hall, P.G. & Hyde, C.C. (1980). Martingale central limit theory and its applications. Academic Press, New York.

Purdom, E. & Holmes, S.P. (2005). Error distribution for gene expression Data, *Statistical Applications in Genetics and Molecular Biology*. **4**, Article 16.

Rao, C.R. (1983). Likelihood ratio tests for relationships between two covariance matrices. In: Karlin, S., Amemiya, T., Goodman, L.A. (Eds.), Studies in Econometrics, Time Series and Multivariate Statistics. Academic Press, New York, pp. 529-543.

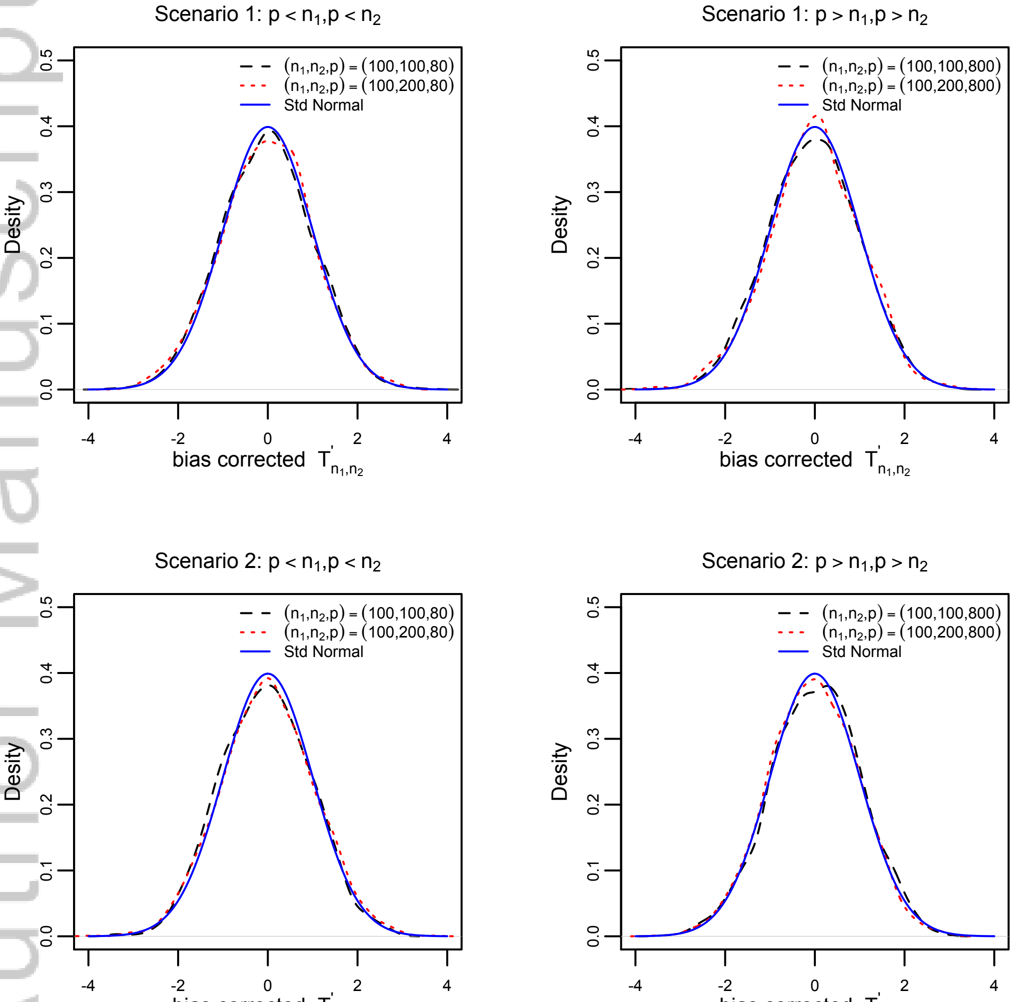
Schott, J.R. (1991). Some tests for common principal component subspaces in several groups. *Biometrika* **78**, 771-777.

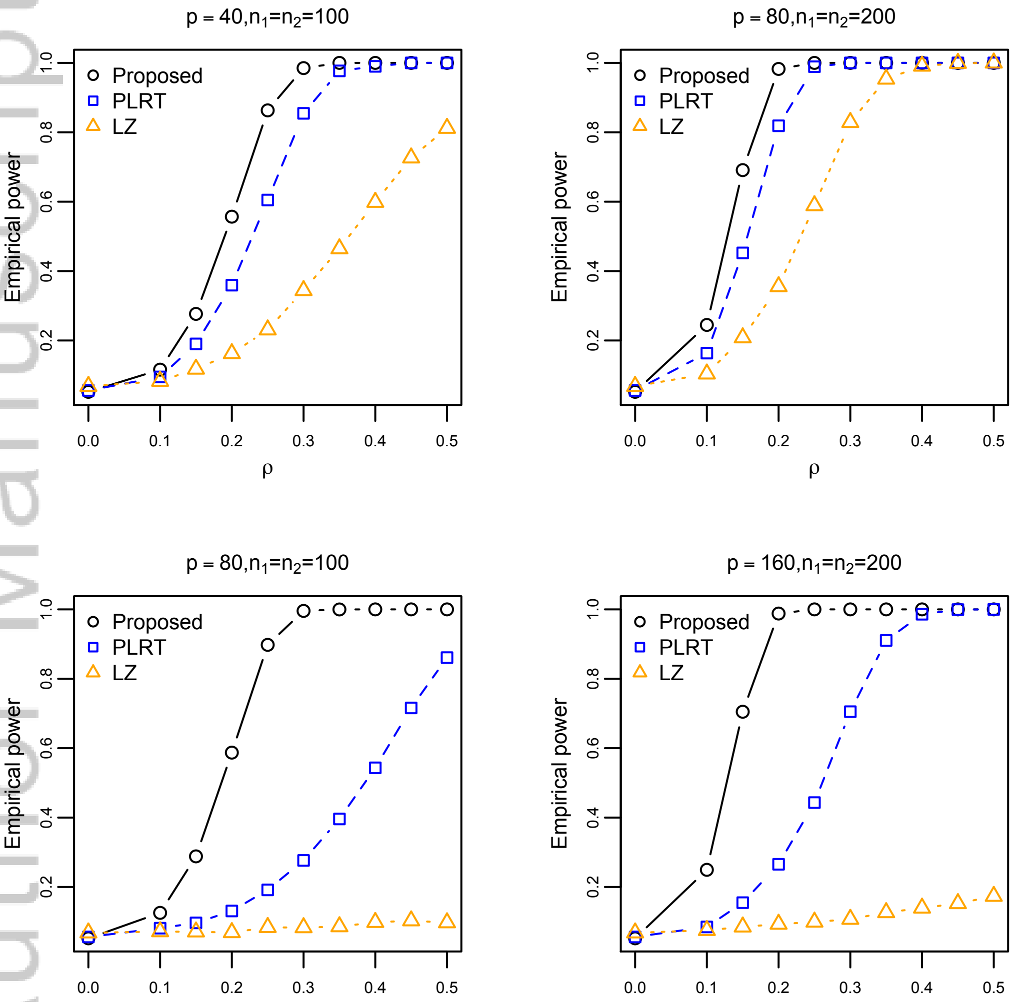
Schott, J.R. (1999). A test for proportional covariance matrices . *Computational Statistics & Data Analysis* **32**, 135-146.

Wang, L., Peng, B. & Li, R. (2015). A high-dimensional nonparametric multivariate test for mean vector. *J. Amer. Statist. Assoc.*. **110**, 1658-1669.

Xu, L., Liu, B., Zheng, S. & Bao, S. (2014). Testing proportionality of two large-dimensional covariance matrices. *Computational Statistics & Data Analysis* **78**,43-55.

Zou, C. L., Peng, L. H., Feng, L. & Wang, Z. J. (2014). Multivariate sign-based high-dimensional tests for sphericity. *Biometrika* **101**, 229-236.


Scenario 1: p < n1,p < n2
Scenario 1: p > n1,p > n2
Scenario 2: p < n1,p < n2
Scenario 2: p > n1,p > n2
(n1,n2,p) = (100,100,80)
(n1,n2,p) = (100,200,80)
(n1,n2,p) = (100,100,800)
(n1,n2,p) = (100,200,800)
Std Normal
Desity
bias corrected T'n1,n2
0.0
0.1
0.2
0.3
0.4
0.5
-4
-2
0
2
4



p = 40,n$_1$=n$_2$=100
p = 80,n$_1$=n$_2$=200
p = 80,n$_1$=n$_2$=100
p = 160,n$_1$=n$_2$=200
Proposed
PLRT
LZ
Empirical power
ρ
0.0
0.1
0.2
0.3
0.4
0.5
0.2
0.4
0.6
0.8
1.0