Sphericity and Identity Test for High-dimensional Covariance Matrix using Random Matrix Theory

DOI:
10.1007/s10255-021-1004-1

Sphericity and Identity Test for High-dimensional Covariance Matrix using Random Matrix Theory

Shoucheng Yuan[a], Jie Zhou[a*], Jianxin Pan[b] Jieqiong Shen[a]

[a] *College of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China*

[b] *School of Mathematics, University of Manchester, Manchester M13 9PL, UK*

Abstract: This paper addresses the issue of testing sphericity and identity of high-dimensional population covariance matrix when the data dimension exceeds the sample size. The central limit theorem of the first four moments of eigenvalues of sample covariance matrix is derived using random matrix theory for generally distributed populations. Further, some desirable asymptotic properties of the proposed test statistics are provided under the null hypothesis as data dimension and sample size both tend to infinity. Simulations show that the proposed tests have a greater power than existing methods for the spiked covariance model.

Keywords: Sphericity test; Identity test; High-dimensional covariance matrix; Spiked model; Spectral distribution.

1. Introduction

High-dimensional statistical inference problems for covariance matrices are increasingly encountered in many applications such as image processing, stock marketing and genetics. A fundamental problem in such applications is the hypothesis test for covariance matrix when data dimension is much larger than the sample size. With the advancement in computer technology, it is feasible to analyze the high-dimensional data. However, many of the classical multivariate methods may not work properly when the dimension equals or exceeds the sample size. These procedures rely on the classical regime where the sample size tends to infinity while the dimension remains fixed.

Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) p-dimensional vectors with mean zero and covariance matrix Σ_p. We focus on testing two structures for the population covariance matrix:

1) The sphericity test

$$H_{0a} : \Sigma_p = \sigma^2 \mathrm{I}_p \quad \text{vs.} \quad H_{1a} : \Sigma_p \neq \sigma^2 \mathrm{I}_p; \tag{1}$$

2) The identity test

$$H_{0b} : \Sigma_p = \mathrm{I}_p \quad \text{vs.} \quad H_{1b} : \Sigma_p \neq \mathrm{I}_p, \tag{2}$$

*Corresponding author. E-mail: jzhou@scu.edu.cn

where σ is an unknown but finite positive constant, and I_p is the $p \times p$ identity matrix. Traditional tests for covariance matrix based on the likelihood ratio (Anderson (2003)) can not used when the dimensionality diverges at the same rate as the sample size, since the sample covariance matrix (SCM) cannot converge to its population counterpart in high-dimensional regime where both n and p tend to infinity.

In the need of theoretical research and practical application, the study of the above hypothesis tests has attracted statisticians' attentions in high-dimensional regime. Ledoit and Wolf (2002) restudied certain tests originally setting up by John (1971) and Nagao (1973) in a fixed p context and generalized the locally best invariant (LBI) tests to accommodate the situation that both p and n go to infinity with $p/n \to c \in (0, \infty)$. Birke and Dette (2005) later extended these results to the cases $c = 0$ and $c = \infty$ in high-dimensional regime. Meanwhile, Srivastava (2005) refined the LBI tests by applying the unbiased estimators of $\mathrm{tr}\Sigma_p^k/p, k = 1, 2$. Subsequently, some test statistics were investigated based on these unbiased estimators, such as in Srivastava et al. (2011); Srivastava and Yanagihara (2010). Moreover, Fisher et al. (2010); Fisher (2012) studied homogeneous statistics constructed from unbiased estimators of $\mathrm{tr}\Sigma_p^k/p, k = 1, 2, 3, 4$. Although these tests overcame the constraint of the classical regime, they relied heavily on the normality assumption. For non-normal case, Bai et al. (2009) modified the likelihood ratio tests to accommodate the situations where p and n both can be large but $p < n$. Chen et al. (2010) developed a method where the statistics are constituted by some well-selected U-statistics. In Wang and Yao (2013), the likelihood ratio test and LW test proposed in Ledoit and Wolf (2002) for sphericity were modified in the general population using random matrix theory (RMT). Most recently, a robust test for sphericity was addressed in Tian et al. (2015) by taking the maximum of two test statistics proposed in Srivastava (2005) and Fisher et al. (2010). Wang et al. (2018) proposed two statistics for testing the identity of high-dimensional covariance matrix, based on large dimensional random matrix theory. Hu et al. (2019) considered testing the sphericity of elliptical populations by using the linear spectral statistics of high-dimensional sample covariance matrix.

A method based on inequalities was presented to construct test statistics in Srivastava (2005). Fisher et al. (2010) and Fisher (2012) employed successfully this method to design some desired statistics composed of the first four moments of eigenvalues of the SCM and tested the hypotheses (1) and (2) under the assumption of normal distribution. In this paper, we primarily focus on the central limit theorem (CLT) of the first four moments of spectral distribution of the SCM for general populations. Motivated form Srivastava (2005), we construct some new test statistics in order to pursue the higher test powers for the spiked covariances introduced by Johnstone (2001). Whereas, from the technical point of view, the proposed approach differs from Fisher et al. (2010); Fisher (2012); Ledoit and Wolf (2002); Srivastava (2005) and is analogous to the way devised in Bai et al. (2009). The pivotal tool is the CLT for linear spectral statistics of the SCM established in Bai and Silverstein (2004) and later meliorated in Pan and Zhou (2008).

We discuss the testing problems (1) and (2) from the viewpoint of RMT. We propose new test statistics for general populations in high-dimensional regime where both the dimension p and the sample size n tend to infinity with $p/n \to c \in (0, \infty)$. Under this regime, an asymptotic joint distribution of the first four moments of the eigenvalues is deduced from the CLT for linear spectral statistics of the SCM. Then, we derive some desirable asymptotic properties of the proposed test statistics under the null hypotheses utilizing the aforementioned joint distribution. Moreover, some numerical simulations are provided for demonstrating the performance of the proposed tests.

This paper is organized as follows. In Section 2, we review the consistent estimators of $\mathrm{tr}\Sigma_p^k/p$ for $k = 1, 2, 3, 4$ under the general assumptions and derive their asymptotic behaviors with general moment conditions under the null hypotheses. In Section 3, we introduce two new test statistics and deduce their asymptotic normality. Section 4 reports some simulation results that demonstrate the asymptotic behavior of the proposed statistics and the good performance for high-dimensional covariance with spiked structure. Section 5 provides the analysis of actual data, and Section 6 includes concluding remarks. All the technical details including the preliminary results in RMT and the proofs of lemmas and theorems are presented in the appendix.

2. Estimators of $\mathrm{tr}\boldsymbol{\Sigma}^{k}/p$ and their asymptotic properties

Let the SCM be $S_n = \frac{1}{n}\sum_{i=1}^n X_i X_i'$. Suppose that H_p and F_n are respectively spectral distributions of Σ_p and S_n. We define the integer-order moments of H_p and F_n:

$$\alpha_k := \int t^k dH_p(t) = \frac{1}{p}\mathrm{tr}(\Sigma_p^k) \quad \text{and} \quad \hat{\beta}_k := \int t^k dF_n(t) = \frac{1}{p}\mathrm{tr}(S_n^k), \quad k = 1, 2, \ldots$$

Assuming that the observations are normally distributed, the estimators $\hat{\alpha}_k$ of α_k, $k = 1, 2, 3, 4$, were proved to be consistent, unbiased and asymptotically normal as $(n, p) \to \infty$ and adopted in Srivastava (2005); Fisher et al. (2010); Fisher (2012); Tian et al. (2015). These estimators can be expressed as the polynomials of $\hat{\beta}_k$'s:

$$\begin{aligned}
\hat{\alpha}_1 &= \hat{\beta}_1, \\
\hat{\alpha}_2 &= \tau_2\big(\hat{\beta}_2 - c_n\hat{\beta}_1^2\big), \\
\hat{\alpha}_3 &= \tau_3\big(\hat{\beta}_3 - 3c_n\hat{\beta}_2\beta_1 + 2c_n^2\hat{\beta}_1^3\big), \\
\hat{\alpha}_4 &= \tau_4\left(\hat{\beta}_4 - 4c_n\hat{\beta}_3\hat{\beta}_1 - \frac{2n^2+3n-6}{n^2+n+2}c_n\hat{\beta}_2^2 + \frac{10n^2+12n}{n^2+n+2}c_n^2\hat{\beta}_2\hat{\beta}_1^2 - \frac{5n^2+6n}{n^2+n+2}c_n^3\hat{\beta}_1^4\right),
\end{aligned}$$

where

$$\begin{aligned}
c_n &= p/n, \\
\tau_2 &= n^2/((n-1)(n+2)), \\
\tau_3 &= n^4/((n-1)(n-2)(n+2)(n+4)), \\
\tau_4 &= n^5(n^2+n+2)/((n+1)(n+2)(n+4)(n+6)(n-1)(n-2)(n-3)).
\end{aligned}$$

If the underlying distribution is non-normal, we note that the unbiasedness does not hold any more for $\hat{\alpha}_k, k = 2, 3, 4$, but the consistency and asymptotic normality can be retained under some suitable assumptions in Tian et al. (2015).

Assumption 1: The sample size n and the dimension p both tend to infinity with $c_n = p/n \to c \in (0, \infty)$.

Assumption 2: There is a doubly infinite matrix composed of i.i.d. random variables w_{ij} satisfying

$$E(w_{ij}) = 0, \quad E(w_{ij}^2) = 1, \quad E(w_{ij}^4) < \infty, \quad i, j \geq 1.$$

Letting $W_n = (w_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$, the observation vectors can be represented as $X_j = \Sigma_p^{1/2} w_{\cdot j}$, where $w_{\cdot j} = (w_{ij})_{1 \leq i \leq p}$ is the j-th column of the matrix W_n.

Assumption 3: The spectral norm of Σ_p is bounded by a positive constant and the population spectral distribution H_p converges weakly to a non-random distribution H as $p \to \infty$.

It is worth pointing out that we are accustomed to assuming $E(w_{11}^4) = 3 + \Delta$ where Δ is a finite constant which is 0 if w_{ij} is normal in Assumption 2. Under these assumptions, we know that the estimators $\hat{\alpha}_k$ converge almost surely to $\alpha_k, k = 1, 2, 3, 4$ from Tian et al. (2015).

Lemma 2.1: *Under Assumptions 1–3, if $\Sigma_p = I_p$, then*

$$p \begin{pmatrix} \hat{\beta}_1 - 1 \\ \hat{\beta}_2 - (1 + c_n) \\ \hat{\beta}_3 - (1 + 3c_n + c_n^2) \\ \hat{\beta}_4 - (1 + 6c_n + 6c_n^2 + c_n^3) \end{pmatrix} \xrightarrow{D} N_4(\mathbf{m}, \mathbf{V}), \tag{3}$$

where

$$\mathbf{m} = (1 + \Delta)(0, c, 3c + 3c^2, 6c + 17c^2 + 6c^3)',$$

and $\mathbf{V}$ *is a* 4×4 *symmetric matrix with entries*

$$\begin{aligned}
V_{11} &= 2c + c\Delta, \\
V_{12} &= 4c(c+1) + 2c(c+1)\Delta, \\
V_{13} &= 6c(c^2 + 3c + 1) + 3c(c^2 + 3c + 1)\Delta, \\
V_{14} &= 8c(c^3 + 6c^2 + 6c + 1) + 4c(c^3 + 6c^2 + 6c + 1)\Delta, \\
V_{22} &= 4c(2c^2 + 5c + 2) + 4c(c^2 + 2c + 1)\Delta, \\
V_{23} &= 12c(c^3 + 5c^2 + 5c + 1) + 6c(c^3 + 4c^2 + 4c + 1)\Delta, \\
V_{24} &= 8c(2c^4 + 17c^3 + 32c^2 + 17c + 2) + 8c(c^4 + 7c^3 + 12c^2 + 7c + 1)\Delta, \\
V_{33} &= 6c(3c^4 + 24c^3 + 46c^2 + 24c + 3) + 9c(c^4 + 6c^3 + 11c^2 + 6c + 1)\Delta, \\
V_{34} &= 24c(c^5 + 12c^4 + 37c^3 + 37c^2 + 12c + 1) + 12c(c^5 + 9c^4 + 25c^3 + 25c^2 + 9c + 1)\Delta, \\
V_{44} &= 8c(4c^6 + 66c^5 + 300c^4 + 485c^3 + 300c^2 + 66c + 4)
\end{aligned}$$

$$+ 16c(c^6 + 12c^5 + 48c^4 + 74c^3 + 48c^2 + 12c + 1)\Delta.$$

Lemma 2.1 can be viewed as a new CLT of linear spectral statistics $\hat{\beta}_k, k = 1, 2, 3, 4$. It shows that the first four moments of spectral distributions of S_n asymptotically follow normal distribution when the population covariance Σ_p is an identity matrix. If w_{11} is normal, then the formula (3) reduces to a simpler form with $\Delta = 0$. One can observe that the asymptotic distribution crucially depends on the limiting dimension-to-sample ratio c. In addition, the entries in the covariance matrix $\mathbf{V}$ become more and more complicated as the order of moments of sample eigenvalues increases.

Lemma 2.2: *Under Assumptions 1–3, if $\Sigma_p = I_p$, then*

$$n(\hat{\alpha}_1 - 1, \hat{\alpha}_2 - 1, \hat{\alpha}_3 - 1, \hat{\alpha}_4 - 1)' \xrightarrow{D} N(\widetilde{\mathbf{m}}, \widetilde{\mathbf{V}}), \tag{4}$$

where

$$\widetilde{\mathbf{m}} = \Delta(0, 1, 3, c + 6)',$$

and

$$\widetilde{\mathbf{V}} = \frac{1}{c} \begin{pmatrix} 2 + \Delta & 4 + 2\Delta & 6 + 3\Delta & 8 + 4\Delta \\ 4 + 2\Delta & 4(c + 2) + 4\Delta & 12(c + 1) + 6\Delta & 8(3c + 2) + 8\Delta \\ 6 + 3\Delta & 12(c + 1) + 6\Delta & 6(c^2 + 6c + 3) + 9\Delta & 24(c^2 + 3c + 1) + 12\Delta \\ 8 + 4\Delta & 8(3c + 2) + 8\Delta & 24(c^2 + 3c + 1) + 12\Delta & 8(c^3 + 12c^2 + 18c + 4) + 16\Delta \end{pmatrix}.$$

In fact, Lemma 2.2 can also be viewed as a CLT for the estimators $\hat{\alpha}_k$ of $\alpha_k (k = 1, 2, 3, 4)$ when the population covariance Σ_p is identity. One can see from (4) that the unbiasedness of $\hat{\alpha}_k$, $k = 2, 3, 4$, can not retain and their mean vectors and the covariance matrix bring some shifts related to Δ. When neither $\Delta = 0$ nor $\Sigma_p = I_p$, the limiting distribution will become very complicated (see Pan and Zhou (2008)).

3. Test procedures

The testing problems (1) and (2) remain invariant under the scalar transformation $x \to cx (c \neq 0)$, and the orthogonal transformation $x \to Gx$, where G is an orthogonal matrix. Thus, we can assume without loss of generality that $\Sigma_p = \text{diag}(\lambda_1, \ldots, \lambda_p)$ with $\lambda_i > 0$ for $i = 1, \ldots, p$.

3.1. *A test for sphericity*

From Hölder's inequality, we know that

$$\left(\sum_{i=1}^{p}\lambda_i^2\right)^3 \leq \left(\sum_{i=1}^{p}\lambda_i^4\right)\left(\sum_{i=1}^{p}\lambda_i\right)^2$$

with equality holding if and only if $\lambda_1, \ldots, \lambda_p$ are equal. Therefore, the moments of the distribution H_p satisfy

$$\frac{\alpha_4}{\alpha_2^2} \geq \frac{\alpha_2}{\alpha_1^2}, \tag{5}$$

and the equality holds if and only if the hypothesis of sphericity is true (see Srivastava (2005) and Fisher et al. (2010)). Denote

$$\phi := \frac{\alpha_4}{\alpha_2^2} - \frac{\alpha_2}{\alpha_1^2} \geq 0,$$

then $\phi = 0$ if and only if $\Sigma = \sigma^2 I_p$. Hence, the sphericity test (1) is equivalent to the following hypothesis

$$H_{0a} : \phi = 0 \quad \text{vs.} \quad H_{1a} : \phi > 0.$$

We propose the new test statistic

$$\gamma_1 = \frac{\hat{\alpha}_4}{\hat{\alpha}_2^2} - \frac{\hat{\alpha}_2}{\hat{\alpha}_1^2}. \tag{6}$$

It is worth noting that in Srivastava (2005) and Fisher et al. (2010), the test statistics

$$T_s = \frac{\hat{\alpha}_2}{\hat{\alpha}_1^2} - 1 \quad \text{and} \quad T_f = \frac{\hat{\alpha}_4}{\hat{\alpha}_2^2} - 1$$

were designed to test the hypotheses $\alpha_2/\alpha_1^2 - 1 = 0$ and $\alpha_4/\alpha_2^2 - 1 = 0$ respectively. Obviously, the test statistic (6) can be treated as a linear combination of T_s and T_f. From Tian et al. (2015), we know that these two statistics are asymptotically independent while the limiting ratio c is large. Thus, the asymptotic normality of statistic γ_1 still holds as c increases. The following theorem gives the asymptotic distribution of the statistic γ_1 under null hypothesis H_{0a}.

Theorem 3.1: *Under Assumptions 1–3, when the null hypothesis H_{0a} in (1) holds, we have*

$$T_1 = \frac{n\gamma_1 - m_1}{\sqrt{V_1}} \xrightarrow{D} N(0,1),$$

where $m_1 = (c+3)\Delta$, $V_1 = 8c^2 + 96c + 36$. Further, if $E(w_{11}^4) = 3$, then

$$T_1 = \frac{n\gamma_1}{2\sqrt{2c^2 + 24c + 9}} \xrightarrow{D} N(0,1). \tag{7}$$

3.2. *A test for identity*

Here, we consider the problem of testing the covariance matrix to be identity. This is equivalent to the null hypothesis that $\lambda_i = 1$ for all $i = 1, \ldots, p$ against the alternative that $\lambda_i \neq 1$ for at least one $i, i = 1, \ldots, p$. It is clear that

$$\frac{1}{p}\sum_{i=1}^{p}(\lambda_i^2 - \lambda_i)^2 = \frac{1}{p}\big(\mathrm{tr}\Sigma_p^4 - 2\mathrm{tr}\Sigma_p^3 + \mathrm{tr}\Sigma_p^2\big) = \alpha_4 - 2\alpha_3 + \alpha_2 \geq 0$$

with equality holding if and only if $\lambda_i = 1$ for all i. Thus, we propose a new test statistic

$$\gamma_2 = \hat{\alpha}_4 - 2\hat{\alpha}_3 + \hat{\alpha}_2, \tag{8}$$

which is a consistent estimator of $\alpha_4 - 2\alpha_3 + \alpha_2$. The following theorem gives the asymptotic distribution of the statistic γ_2 under the null hypothesis H_{0b}.

Theorem 3.2: *Under Assumptions 1–3, when the null hypothesis H_{0b} in (2) holds, we have*

$$T_2 = \frac{n\gamma_2 - m_2}{\sqrt{V_2}} \xrightarrow{D} N(0,1),$$

where $m_2 = \Delta(c+1), V_2 = 8c^2 + 24c + 4$. Further, if $E(w_{11}^4) = 3$, then

$$T_2 = \frac{n\gamma_2}{2\sqrt{2c^2 + 6c + 1}} \xrightarrow{D} N(0,1). \tag{9}$$

Remark 3.1: *For the case of population mean μ being unknown, the data are represented as*

$$X_j = \mu + \Sigma^{1/2} w_{\cdot j},\ \ j = 1, \ldots, n,$$

then the sample covariance matrix should be taken as

$$S_n^* = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})',$$

where $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$. Because the rank of the matrix $\bar{X}\bar{X}'$ is one, from Theorem A.44 in Bai and Silverstein (2010), substituting S_n for S_n^* will not affect the limiting distributions in Theorems 3.1 and 3.2.

4. Simulation studies

In this section, simulation studies are done to demonstrate the effectiveness of the proposed statistics T_1 and T_2, and to perform comparative studies on the congeneric test statistics in testing the hypotheses (1) and (2).

4.1. *Normality and the attained significance levels*

Consider the following two cases of population distributions:

Case 1: $w_{\cdot j}$ is a p-dimensional normal random vector with mean $\mathbf{0}_p$ and covariance matrix $\Sigma_p = \mathrm{I}_p$;

Case 2: $w_{\cdot j}$ consists of the i.i.d. random variables w_{ij}, $i = 1, \ldots, p$, following the standardized Gamma$(4, 0.5)$ so that $E(w_{ij}) = 0$, $E(w_{ij}^2) = 1$, and $\Sigma_p = \mathrm{I}_p$. It is easy to see $E(w_{11}^4) = 4.5$ and $\Delta = 1.5$.

For the sphericity test (1), we compare the proposed statistics with four statistics given respectively in Srivastava (2005) (referred as T_s), Fisher et al. (2010) (referred as T_f) for Case 1, and Wang and Yao (2013) (referred as T_{CJ}), Tian et al. (2015) (referred as T_m) for Case 2. For the identity test (2), we also take into account four tests given respectively in Srivastava (2005) (still referred as T_s), Fisher (2012) (still referred as T_f) for Case 1, and Wang et al. (2013) (referred as T_{CLW}), Wang et al. (2018) (referred as T_{WH}) for Case 2. Note that the tests in Wang and Yao (2013) and Wang et al. (2013) are the modified versions of those developed in John (1971) and Ledoit and Wolf (2002) respectively for general distributions. We set the significant level as $\alpha = 0.05$, and simulate 10000 independent trials for all tests.

We demonstrate the empirical size of these test statistics in terms of the attained significance level (ASL). To verify the performance of the ASLs, we compare the statistics T_1 and T_2 with the aforementioned test statistics. Tables 1 and 2 respectively present the ASLs of T_1, T_2, T_s, T_f for Case 1 and T_1, T_2, T_{CJ}, T_{CLW} for Case 2 when the sample size n is set as $20, 40, 80, 120$ and the data dimension $p = cn$ with $c = 1, 2, 5, 10$. It can be seen that the ASLs of these tests are close to the nominal significant level α as p and n both increase. At the same time, it is also noted that the ASLs of the proposed statistics T_1 and T_2 perform better in most situations when data come from the gamma population.

Table 1. The ASLs of T_s, T_f and the proposed two test statistics for Case 1

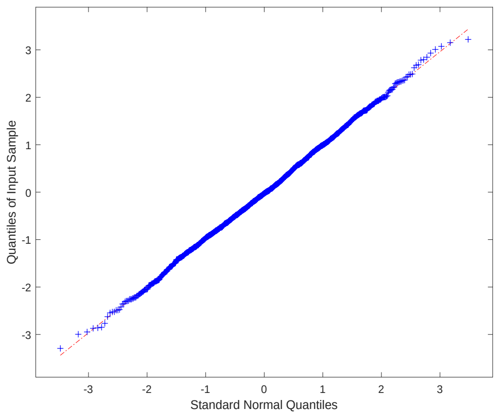
$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
	(a) Sphericity test											
	T_s	T_f	T_1	T_s	T_f	T_1	T_s	T_f	T_1	T_s	T_f	T_1
$n = 20$	0.0529	0.0319	0.0325	0.0507	0.0360	0.0375	0.0514	0.0412	0.0432	0.0544	0.0414	0.0444
$n = 40$	0.0525	0.0486	0.0497	0.0522	0.0496	0.0507	0.0499	0.0506	0.0523	0.0527	0.0494	0.0518
$n = 80$	0.0490	0.0535	0.0521	0.0550	0.0529	0.0534	0.0519	0.0534	0.0530	0.0543	0.0525	0.0536
$n = 120$	0.0521	0.0543	0.0526	0.0479	0.0551	0.0538	0.0487	0.0509	0.0498	0.0499	0.0494	0.0494
	(b) Identity test											
	T_s	T_f	T_2	T_s	T_f	T_2	T_s	T_f	T_2	T_s	T_f	T_2
$n = 20$	0.0572	0.0575	0.0484	0.0520	0.0606	0.0531	0.0565	0.0585	0.0490	0.0501	0.0515	0.0468
$n = 40$	0.0536	0.0610	0.0595	0.0488	0.0602	0.0557	0.0522	0.0588	0.0543	0.0535	0.0581	0.0564
$n = 80$	0.0545	0.0610	0.0580	0.0510	0.0628	0.0589	0.0516	0.0545	0.0534	0.0519	0.0535	0.0512
$n = 120$	0.0493	0.0589	0.0565	0.0517	0.0564	0.0556	0.0535	0.0558	0.0533	0.0514	0.0520	0.0518

We also check the QQ plots of the test statistics T_1 and T_2 under the null hypotheses for sphericity and identity tests. Figures 1 and 2 show the QQ plots of the 10000 observations of the test statistics

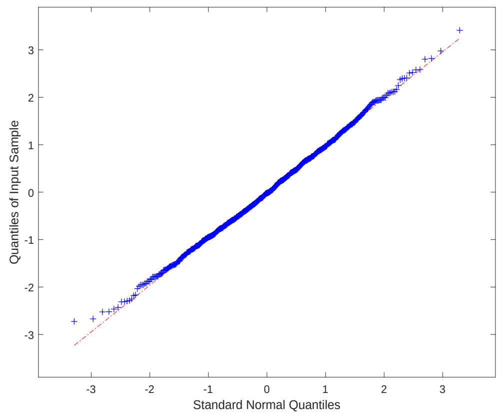
Table 2. The ASLs of T_{CJ}, T_m, T_{CLW}, T_{WH} and the proposed two test statistics for Case 2

$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
	(a) Sphericity test											
	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1
$n = 20$	0.0748	0.0976	0.0461	0.0728	0.1006	0.0488	0.0722	0.1053	0.0422	0.0686	0.1119	0.0383
$n = 40$	0.0683	0.1084	0.0627	0.0623	0.1144	0.0564	0.0641	0.1114	0.0430	0.0573	0.1066	0.0385
$n = 80$	0.0599	0.1062	0.0578	0.0586	0.1024	0.0465	0.0572	0.1080	0.0411	0.0552	0.1064	0.0395
$n = 120$	0.0541	0.0912	0.0539	0.0558	0.0977	0.0461	0.0541	0.1021	0.0356	0.0504	0.0969	0.0326
	(b) Identity test											
	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2
$n = 20$	0.1062	0.1272	0.1035	0.0966	0.1119	0.0890	0.0854	0.0966	0.0713	0.0823	0.0835	0.0553
$n = 40$	0.0808	0.1013	0.0884	0.0764	0.1071	0.0692	0.0685	0.0796	0.0528	0.0670	0.0673	0.0388
$n = 80$	0.0707	0.0979	0.0729	0.0595	0.0855	0.0554	0.0561	0.0679	0.0482	0.0549	0.0607	0.0381
$n = 120$	0.0642	0.0878	0.0624	0.0569	0.0751	0.0516	0.0538	0.0647	0.0410	0.0588	0.0559	0.0357

T_1 and T_2 under the null hypotheses with $n = 200$ and $p = 400$ for Cases 1 and 2 respectively, and indicate that the statistics T_1 and T_2 are normally distributed for sufficiently large n and p under the null hypotheses.



(a) T_1 for Case 1 under H_{0a}



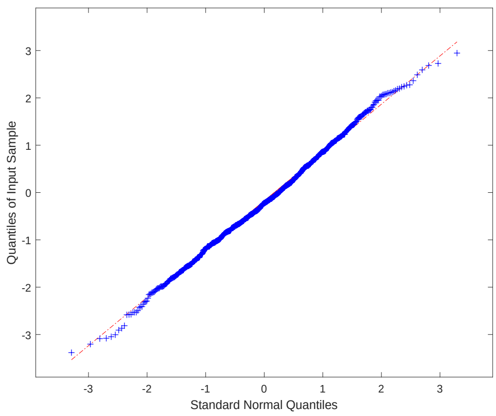
(b) T_2 for Case 1 under H_{0b}

Figure 1. Normal QQ plots for Case 1 from 10000 independent trials.

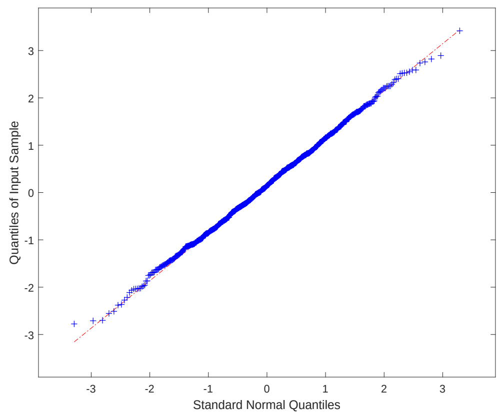
4.2. *The powers of the spiked covariance*

We make simulations on the powers of the proposed tests for normal and gamma populations. To compare the proposed statistics with the statistics given in Srivastava (2005); Fisher et al. (2010); Fisher (2012) for normal population and in Wang et al. (2013); Wang and Yao (2013); Tian et al. (2015); Wang et al. (2018) for gamma population, we discuss some hypothesis testing problems considered in Fisher et al. (2010); Fisher (2012); Srivastava (2006).

Among different structures of population covariance matrices, we are interested in the spiked covariance model which has been applied to signal processing, wireless communication and networking technology (see Nadakuditi and Edelman (2008); Torun et al. (2011); Couillet and Debbah (2011); Bianchi et al. (2011); Couillet and Hachem (2013)). Suppose that the alternative hypothesis



(a) T_1 for Case 2 under H_{0a}



(b) T_2 for Case 2 under H_{0b}

Figure 2. Normal QQ-Plots for Case 2 from 10000 independent trials.

has the following form

$$\Sigma_p = \begin{pmatrix} \Theta_k & \mathbf{0} \\ \mathbf{0} & \mathrm{I}_{p-k} \end{pmatrix} \tag{10}$$

where $k < p$ is a fixed number, Θ_k denotes a $k \times k$ diagonal matrix with all entries not being 1, I_{p-k} is the $(p-k) \times (p-k)$ identity matrix. When the population are normal and gamma distributions, their powers are respectively shown in Tables 3 and 4.

Table 3 shows the powers in the sphericity and identity tests under the normality assumption. The population covariance matrix Σ_p is given by setting $\Theta_1 = 1 + 3\sqrt{c}$ and $\Theta_4 = \mathrm{diag}(0.5, 1.5, 3.5, 5.5)$ in the sphericity test, and $\Theta_2 = \mathrm{I}_2 + \sqrt{c} \cdot \mathrm{diag}(0.2, 3.5)$ and $\Theta_5 = \mathrm{diag}(1.2, 1.5, 1.8, 2.5, 6)$ in the identity test.

Table 4 gives the powers in the sphericity and identity tests for gamma population. In the sphericity test, $\Theta_2 = I_2 + \sqrt{c} \cdot \mathrm{diag}(0.3, 3)$ and $\Theta_4 = \mathrm{diag}(0.5, 2.25, 3.5, 6)$. In the identity test, $\Theta_1 = 1 + 3\sqrt{c}$ and $\Theta_3 = \mathrm{diag}(0.6, 2, 5)$.

Tables 3 and 4 show that the empirical powers of all test statistics increase as the sample size becomes large. Meanwhile, from the simulation results, we can see that the proposed tests are more powerful than the compared tests when k is very small, e.g., $k = 1$ or 2. If k gets bigger, the performance of these tests T_1 and T_2 also becomes better and better as the sample size n increases. At the same time, we note that the power of the proposed statistic T_1, as the linear combination of the statistics T_s and T_f, increases significantly while c becomes large and k is small.

4.3. *Performances for uniform and Student's t distributions*

In our study, we hope that the proposed statistics still work reasonably well in some different distributions, e.g., uniform distribution or Student's t distribution. We mainly look at two scenarios under Assumption 2: (i) each w_{ij} follows uniform distribution over the interval $(-\sqrt{3}, \sqrt{3})$; (ii)

Table 3. Empirical powers from normal data for the sphericity and identity tests

$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
(a) Sphericity test												
	T_s	T_f	T_1	T_s	T_f	T_1	T_s	T_f	T_1	T_s	T_f	T_1
$\Theta_1 = 1 + 3\sqrt{c}$												
$n = 20$	0.6335	0.6754	0.6856	0.6986	0.7437	0.7602	0.7294	0.7861	0.7996	0.7903	0.8545	0.8712
$n = 40$	0.7969	0.8832	0.8885	0.8251	0.9202	0.9290	0.8575	0.9472	0.9690	0.8910	0.9682	0.9796
$n = 80$	0.8930	0.9786	0.9800	0.9161	0.9906	0.9911	0.9366	0.9950	0.9960	0.9395	0.9964	1
$n = 120$	0.9341	0.9946	0.9950	0.9542	0.9975	0.9977	0.9567	0.9995	1	0.9745	0.9990	1
$\Theta_4 = \text{diag}(0.5, 1.5, 3.5, 5.5)$												
$n = 20$	0.9403	0.8944	0.8855	0.8469	0.8017	0.8033	0.5762	0.5555	0.5748	0.3306	0.3140	0.3205
$n = 40$	0.9956	0.9970	0.9971	0.9597	0.9729	0.9735	0.6933	0.7702	0.7757	0.3681	0.4242	0.4256
$n = 80$	1	1	1	0.9936	0.9993	0.9997	0.9154	0.7657	0.9154	0.3689	0.5260	0.5292
$n = 120$	1	1	1	0.9971	1	1	0.7757	0.9576	0.9588	0.3811	0.5744	0.5769
(b) Identity test												
	T_s	T_f	T_2	T_s	T_f	T_2	T_s	T_f	T_2	T_s	T_f	T_2
$\Theta_2 = I_2 + \sqrt{c} \cdot \text{diag}(0.2, 3.5)$												
$n = 20$	0.8118	0.8437	0.8591	0.8390	0.8866	0.8972	0.8625	0.9083	0.9207	0.8800	0.9260	0.9342
$n = 40$	0.9180	0.9624	0.9671	0.9354	0.9774	0.9791	0.9470	0.9870	0.9877	0.9541	0.9889	0.9898
$n = 80$	0.9679	0.9957	0.9966	0.9765	0.9989	0.9991	0.9845	0.9992	0.9993	0.9876	0.9989	0.9990
$n = 120$	0.9885	0.9995	0.9999	0.9912	0.9997	1	0.9936	1	1	0.9990	1	1
$\Theta_5 = \text{diag}(1.2, 1.5, 1.8, 2.5, 6)$												
$n = 20$	0.9718	0.9647	0.9644	0.9018	0.9027	0.8745	0.6476	0.6824	0.6899	0.3886	0.4264	0.4320
$n = 40$	0.9977	0.9982	0.9985	0.9824	0.9840	0.9690	0.7394	0.8370	0.8396	0.3996	0.5463	0.5490
$n = 80$	0.9999	1	1	0.9815	0.9925	0.9993	0.7841	0.9446	0.9474	0.4140	0.6197	0.6370
$n = 120$	1	1	1	0.9982	0.9995	1	0.8030	0.9783	0.9791	0.4150	0.6683	0.6824

Table 4. Empirical powers from gamma data for the sphericity and identity tests

$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
(a) Sphericity test												
	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1
$\Theta_2 = I_2 + \sqrt{c} \cdot \text{diag}(0.3, 3)$												
$n = 20$	0.5060	0.5186	0.4912	0.5897	0.6218	0.5892	0.6757	0.7205	0.7026	0.6977	0.7695	0.7462
$n = 40$	0.7158	0.7864	0.7744	0.7563	0.8462	0.8467	0.8073	0.9116	0.9086	0.8287	0.9269	0.9218
$n = 80$	0.8432	0.9383	0.9414	0.8789	0.9681	0.9702	0.8953	0.9860	0.9875	0.9107	0.9894	0.9912
$n = 120$	0.9056	0.9820	0.9855	0.9238	0.9927	0.9940	0.9413	0.9972	0.9973	0.9433	0.9985	0.9992
$\Theta_4 = \text{diag}(0.5, 2.25, 3.5, 6)$												
$n = 20$	0.9048	0.8946	0.7404	0.8224	0.8149	0.7078	0.5742	0.5884	0.5079	0.3700	0.3887	0.3320
$n = 40$	0.9948	0.9953	0.9892	0.9582	0.9700	0.9566	0.7245	0.7918	0.7513	0.4212	0.5110	0.4754
$n = 80$	1	1	1	0.9932	0.9986	0.9986	0.8124	0.9342	0.9227	0.4666	0.6423	0.6112
$n = 120$	1	1	1	0.9983	1	0.9998	0.8557	0.9689	0.9627	0.4703	0.6883	0.6749
(b) Identity test												
	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2
$\Theta_1 = 1 + 3\sqrt{c}$												
$n = 20$	0.5905	0.6897	0.6010	0.6417	0.7256	0.6800	0.6858	0.7353	0.7476	0.6963	0.7033	0.7882
$n = 40$	0.7500	0.8397	0.8139	0.7819	0.8543	0.8795	0.8130	0.8557	0.9180	0.8210	0.8290	0.9340
$n = 80$	0.8618	0.9352	0.9571	0.8761	0.9373	0.9766	0.8940	0.9328	0.9881	0.9080	0.9021	0.9909
$n = 120$	0.9032	0.9658	0.9902	0.9126	0.9646	0.9960	0.9306	0.9588	0.9968	0.9442	0.9663	1
$\Theta_3 = \text{diag}(0.6, 2, 5)$												
$n = 20$	0.8074	0.8712	0.7944	0.6364	0.7317	0.6398	0.3981	0.4698	0.4191	0.2232	0.2590	0.2299
$n = 40$	0.9338	0.9642	0.9566	0.7846	0.8768	0.8536	0.4581	0.5558	0.5515	0.2291	0.2990	0.2907
$n = 80$	0.9869	0.9960	0.9971	0.8786	0.9487	0.9703	0.4918	0.6137	0.6964	0.2260	0.3000	0.3388
$n = 120$	0.9943	0.9992	1	0.9140	0.9693	0.9914	0.4984	0.6310	0.7558	0.2256	0.2997	0.3529

each w_{ij} is distributed as $\sqrt{\frac{3}{4}}t(8)$, where $t(8)$ means the Student's t distribution with 8 degrees of freedom. The significant level is set as $\alpha = 0.05$. When the true covariance matrix is identity, the ASLs of T_1, T_{CJ}, T_m for sphericity test and T_2, T_{CLW}, T_{WH} for identity test are respectively shown in Tables 5 and 6.

Table 5. The ASLs of test statistics from the uniform distribution, $w_{11} \sim U(-\sqrt{3}, \sqrt{3})$

$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
	(a) Sphericity test											
	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1
$n = 20$	0.0488	0.0642	0.0364	0.0502	0.0734	0.0425	0.0476	0.0775	0.0516	0.0512	0.0865	0.0612
$n = 40$	0.0509	0.0793	0.0508	0.0561	0.0920	0.0604	0.0524	0.0883	0.0671	0.0453	0.0898	0.0683
$n = 80$	0.0496	0.0849	0.0572	0.0481	0.0864	0.0539	0.0493	0.0915	0.0685	0.0520	0.0974	0.0746
$n = 120$	0.0526	0.0886	0.0556	0.0493	0.0854	0.0579	0.0534	0.0875	0.0656	0.0474	0.0902	0.0759
	(b) Identity test											
	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2
$n = 20$	0.0549	0.0469	0.0398	0.0612	0.0538	0.0467	0.0640	0.0551	0.0579	0.0600	0.0535	0.0597
$n = 40$	0.0544	0.0413	0.0463	0.0581	0.0485	0.0568	0.0559	0.0496	0.0645	0.0529	0.0477	0.0709
$n = 80$	0.0542	0.0424	0.0602	0.0518	0.0412	0.0639	0.0505	0.0478	0.0775	0.0513	0.0502	0.0731
$n = 120$	0.0458	0.0360	0.0582	0.0552	0.0394	0.0678	0.0532	0.0414	0.0789	0.0544	0.0504	0.0860

Table 6. The ASLs of test statistics from the Student's t distribution, $w_{11} \sim \sqrt{3/4}\ t(8)$

$p = cn$	$c = 1$			$c = 2$			$c = 5$			$c = 10$		
	(a) Sphericity test											
	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1	T_{CJ}	T_m	T_1
$n = 20$	0.0729	0.0934	0.0479	0.0710	0.1017	0.0537	0.0723	0.1148	0.0475	0.0703	0.1115	0.0417
$n = 40$	0.0727	0.1119	0.0685	0.0678	0.1166	0.0622	0.0654	0.1212	0.0521	0.0573	0.1136	0.0411
$n = 80$	0.0660	0.1142	0.0671	0.0586	0.1158	0.0588	0.0540	0.1070	0.0456	0.0568	0.1024	0.0322
$n = 120$	0.0576	0.0924	0.0558	0.0556	0.1062	0.0520	0.0558	0.1044	0.0412	0.0594	0.1060	0.0332
	(b) Identity test											
	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2	T_{CLW}	T_{WH}	T_2
$n = 20$	0.1050	0.1297	0.1040	0.0999	0.1240	0.1013	0.0881	0.0929	0.0737	0.0834	0.0851	0.0593
$n = 40$	0.0844	0.1276	0.1012	0.0781	0.1100	0.0776	0.0673	0.0788	0.0588	0.0688	0.0684	0.0474
$n = 80$	0.0671	0.1038	0.0829	0.0656	0.0863	0.0630	0.0598	0.0692	0.0429	0.0577	0.0626	0.0376
$n = 120$	0.0586	0.0856	0.0698	0.0568	0.0792	0.0492	0.0594	0.0674	0.0395	0.0600	0.0542	0.0346

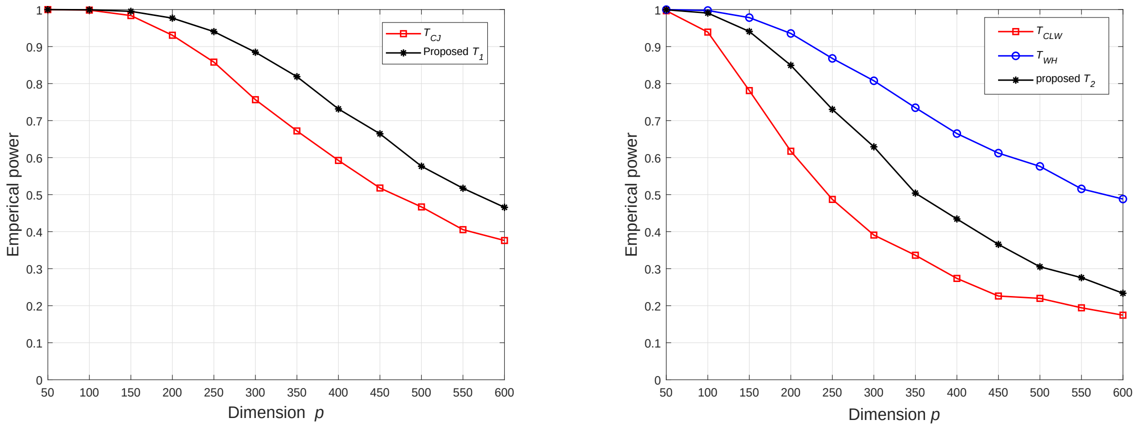
The results in Tables 5 and 6 indicate that the ASLs of the statistics T_1, T_2, T_{CJ}, T_{CLW} and T_{WH} appear to be more conservative around $\alpha = 0.05$ except that T_m is a little bit too big, as the sample size n becomes larger under the null hypotheses, whether the data come from uniform distribution or t distribution. Meanwhile, we also note that the proposed statistics T_1 and T_2 have more accurate ASLs than T_{CJ}, T_{CLW} and T_{WH} in most cases.

Next, we compare the powers of sphericity and identity tests when data come from uniform distribution and t distribution as mentioned above. When $n = 50$, the empirical powers are presented in Figures 3 and 4.

Figure 3 demonstrates that when data are from uniform distribution, the statistic T_1 has better performance than T_{CJ}, but T_2 does not outperform T_{WH}. Meanwhile, if data come from t distribution, we also see from Figure 4 that the powers of the statistics T_1 and T_2 respectively become better than T_{CJ} and T_{CLW}, T_{WH}.

5. Leukemia data analysis

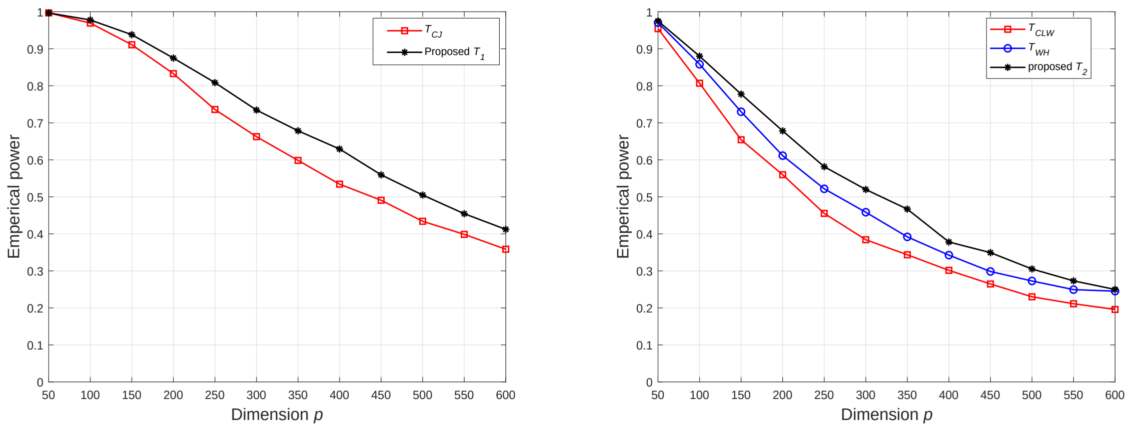
In this section, the statistics are applied to the classical leukemia data in Golub et al. (1999). The data are published at the website: `http://portals.broadinstitute.org/cgi-bin/cancer/publications/pub_paper.cgi?mode=view&paper_id=43`. We still adopt the preprocessing way



(a) $n = 50, \Theta = \mathrm{diag}(0.5, 1.25, 2.5, 5)$

(b) $n = 50, \Theta = \mathrm{diag}(0.6, 2, 5)$

Figure 3. Empirical powers versus dimension p under uniform distribution from 10000 independent trials.



(a) $n = 50, \Theta = \mathrm{diag}(0.5, 1.25, 2.5, 5)$

(b) $n = 50, \Theta = \mathrm{diag}(0.6, 2, 5)$

Figure 4. Empirical powers versus dimension p under t distribution from 10000 independent trials.

attributed to Dudoit et al. (2002) and Dettling and Bühlmann (2003) by filtering, thresholding and a logarithmic transformation. The data are comprised of $p = 3571$ genes and the effective sample size is only $n = 70$. As a cross-sectional example, the data are adopted to verify test procedures for sphericity and identity of the covariance matrices in Fisher et al. (2010), Fisher (2012) and Srivastava (2006). For test of sphericity, the value of statistic T_1 is 213.974, and the corresponding p-value is 0. For test of identity, the value of statistic T_2 is 6967.0553, and the corresponding p-value is 0. Thus, we conclude that the hypothesis of sphericity or identity covariance matrix is rejected, in consistent with the ones by Srivastava (2005), Fisher et al. (2010) and Fisher (2012).

6. Concluding remarks

The issue of testing sphericity or identity of covariance matrix in high-dimensional cases is addressed. Like Srivastava (2005), Fisher et al. (2010) and Fisher (2012), our test statistics are induced by some inequalities. But unlike the T_s and T_f test statistics, we employ the first four moments of the sample eigenvalues under the general distribution assumptions based on the RMT. We derive the asymptotic joint distribution of these moments. Moreover, we propose new statistics constructed by the first four moments of eigenvalues for the sphericity test and identity test, and obtain the desirable asymptotic properties of the proposed statistics T_1 and T_2 under the null hypotheses for the general populations. Simulations indicate that, for the spiked covariance models, whether the underlying distribution is normal or not, the proposed test statistics perform better than the existing test statistics. As an interesting question, we will deduce the asymptotic behavior of the alternative hypothesis in the future.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grants 61374027 and 11871357, and the Sichuan Science and Technology Program under grant 2019YJ0122.

Appendix A. Preliminary results in RMT

Suppose that Assumptions 1–3 hold, from Silverstein (1995), the spectral distribution F_n of the sample covariance matrix S_n converges weakly to a limiting distribution $F^{c,H}$, whose Stieltjes transform $s(z)$ satisfies the following equation:

$$s(z) = \int \frac{1}{t(1-c-czs(z))-z} dH(t), \quad z \in \mathcal{C}^+. \tag{A1}$$

Let F^{c_n,H_p} be a distribution derived from $F^{c,H}$ by substituting c_n and H_p for c and H respectively. Then, the kth moments of F^{c_n,H_p} and $F^{c,H}$ are respectively given by

$$\beta_k = \int t^k dF^{c_n,H_p}(t) \quad \text{and} \quad \widetilde{\beta}_k = \int t^k dF^{c,H}(t), \quad k = 1, 2, \ldots \tag{A2}$$

Obviously, $\beta_k \to \widetilde{\beta}_k$ as $(n,p) \to \infty$.

Pan and Zhou (2008) derives the CLT for the linear spectral statistics. Specifically, denote $G_n(x) = p(F_n(x) - F^{c_n,H_p}(x))$, then for any analytic functions $f_1, \ldots, f_k$ on an open region containing the support of $F^{c,H}$, the random vector

$$\left(\int f_1(x) dG_n(x), \ldots, \int f_k(x) dG_n(x) \right)'$$

converges weakly to a normal vector $(X_{f_1}, \ldots, X_{f_k})'$ under some assumptions on the population covariance matrix Σ.

When $\Sigma_p = I_p$, the distribution $F^{c,H}$ is the well-known Marčenko–Pastur law:

$$g_c(x) = \frac{1}{2\pi cx}\sqrt{((1+\sqrt{c})^2 - x)(x-(1-\sqrt{c})^2)}, \quad (1-\sqrt{c})^2 \le x \le (1+\sqrt{c})^2, \tag{A3}$$

and, from (A1), the Stieltjes transform satisfies

$$s(z) = \frac{1}{1-c-czs(z)-z}.$$

Denoting $\underline{s}(z) = cs(z) - \frac{1-c}{z}$, $\Delta = E(w_{11}^4 - 3)$, and applying the CLT for linear spectral statistics in Pan and Zhou (2008), we obtain that the random vector $(X_{f_1}, \ldots, X_{f_k})'$ follows the normal distribution with the mean function

$$\begin{aligned} E[X_{f_i}] &= -\frac{1}{2\pi i}\oint \frac{c(\underline{s}(z)/(1+\underline{s}(z)))^3 f_i(z)}{(1-c(\underline{s}(z)/(1+\underline{s}(z)))^2)^2}dz - \frac{E(w_{11}^4-3)}{2\pi i}\oint \frac{c(\underline{s}(z)/(1+\underline{s}(z)))^3 f_i(z)}{1-c(\underline{s}(z)/(1+\underline{s}(z)))^2}dz \\ &=: I_1(f_i) + I_2(f_i)\Delta, \quad 1 \le i \le k, \end{aligned} \tag{A4}$$

and the covariance function

$$\begin{aligned} \mathrm{Cov}(X_{f_i}, X_{f_j}) = &-\frac{1}{2\pi^2}\oint\oint \frac{f_i(z_1)f_j(z_2)}{(\underline{s}(z_1)-\underline{s}(z_2))^2}\underline{s}'(z_1)\underline{s}'(z_2)dz_1dz_2 \\ &-\frac{cE(w_{11}^4-3)}{4\pi^2}\oint f_i(z_1)\frac{\partial}{\partial z_1}\Big\{\frac{\underline{s}(z_1)}{1+\underline{s}(z_1)}\Big\}dz_1 \cdot \oint f_j(z_2)\frac{\partial}{\partial z_2}\Big\{\frac{\underline{s}(z_2)}{1+\underline{s}(z_2)}\Big\}dz_2 \\ =&: J_1(f_i,f_j) + J_2(f_i,f_j)\Delta, \quad 1 \le i,j \le k, \end{aligned} \tag{A5}$$

where the integrals are along contours enclosing the support of $F^{c,H}$. However, the concrete applications of the CLT are inconvenient, since those parameters in the CLT are expressed through the integrals on contours that are vague. Therefore, we use another way for calculating conveniently these parameters. Using the idea employed in Zheng (2012), we can convert all these integrals along the unit circle by introducing a change of variable

$$z = 1 + hr\xi + hr^{-1}\bar{\xi} + h^2,$$

where $h = \sqrt{c}$, $r > 1$ but close to 1, and $|\xi| = 1$. In Wang and Yao (2013), the integrals in (A4) and (A5) are written as follows:

$$I_1(f_i) = \lim_{r\downarrow 1}\frac{1}{2\pi i}\oint_{|\xi|=1} f_i(|1+h\xi|^2)\left(\frac{\xi}{\xi^2 - r^{-2}} - \frac{1}{\xi}\right)d\xi, \tag{A6}$$

$$I_2(f_i) = \frac{1}{2\pi i}\oint_{|\xi|=1} f_i(|1+h\xi|^2)\frac{1}{\xi^3}d\xi, \tag{A7}$$

$$J_1(f_i,f_j) = \lim_{r\downarrow 1} -\frac{1}{2\pi^2}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1}\frac{f_i(|1+h\xi_1|^2)f_j(|1+h\xi_2|^2)}{(\xi_1 - r\xi_2)^2}d\xi_1 d\xi_2, \tag{A8}$$

$$J_2(f_i, f_j) \;\; = \;\; -\frac{1}{4\pi^2} \oint_{|\xi_1|=1} \frac{f_i(|1+h\xi_1|^2)}{\xi_1^2} d\xi_1 \oint_{|\xi_2|=1} \frac{f_j(|1+h\xi_2|^2)}{\xi_2^2} d\xi_2. \tag{A9}$$

Appendix B. Proof of Lemma 2.1

Denote $f_k(x) = x^k, k = 1, 2, 3, 4$. Based on the CLT for linear spectral statistics, we note that the random vector

$$\Big(\int f_1(x) dG_n(x), \int f_2(x) dG_n(x), \int f_3(x) dG_n(x), \int f_4(x) dG_n(x) \Big)'$$

converges weakly to a normal vector $(X_{f_1}, X_{f_2}, X_{f_3}, X_{f_4})'$ with the mean

$$\mathbf{m} = (I_1(f_1), I_1(f_2), I_1(f_3), I_1(f_4))' + \Delta (I_2(f_1), I_2(f_2), I_2(f_3), I_2(f_4))', \tag{B1}$$

and the covariance

$$\mathbf{V} = (J_1(f_i, f_j) + J_2(f_i, f_j)\Delta)_{1 \le i,j \le 4}. \tag{B2}$$

From the residue theorem, and (A6) and (A7), we can obtain each term in the entries of the vector $\mathbf{m}$ given by (B1), as follows:

$$\begin{aligned}
&I_1(f_1) = I_2(f_1) = 0, \\
&I_1(f_2) = I_2(f_2) = c, \\
&I_1(f_3) = I_2(f_3) = 3c + 3c^2, \\
&I_1(f_4) = I_2(f_4) = 6c + 17c^2 + 6c^3.
\end{aligned}$$

Similarly, from (A8) and (A9), we get each term in the entries of the matrix $\mathbf{V}$ given by (B2), as follows:

$$\begin{aligned}
&J_1(f_1, f_1) = 2c, \\
&J_1(f_2, f_2) = 4c(2c^2 + 5c + 2), \\
&J_1(f_3, f_3) = 6c(3c^4 + 24c^3 + 46c^2 + 24c + 3), \\
&J_1(f_4, f_4) = 8c(4c^6 + 66c^5 + 300c^4 + 485c^3 + 300c^2 + 66c + 4), \\
&J_1(f_1, f_2) = J_1(f_2, f_1) = 4c(c + 1), \\
&J_1(f_1, f_3) = J_1(f_3, f_1) = 6c(c^2 + 3c + 1), \\
&J_1(f_1, f_4) = J_1(f_4, f_1) = 8c(c^3 + 6c^2 + 6c + 1), \\
&J_1(f_2, f_3) = J_1(f_3, f_2) = 12c(c^5 + 5c^2 + 5c + 1), \\
&J_1(f_2, f_4) = J_1(f_4, f_2) = 8c(2c^4 + 17c^3 + 32c^2 + 17c + 2), \\
&J_1(f_3, f_4) = J_1(f_4, f_3) = 24c(c^5 + 12c^4 + 37c^3 + 37c^2 + 12c + 1),
\end{aligned}$$

$$
\begin{aligned}
&J_2(f_1, f_1) = c,\\
&J_2(f_2, f_2) = 4c(c^2 + 2c + 1),\\
&J_2(f_3, f_3) = 9c(c^4 + 6c^3 + 11c^2 + 6c + 1),\\
&J_2(f_4, f_4) = 16c(c^6 + 12c^5 + 48c^4 + 74c^3 + 48c^2 + 12c + 1),\\
&J_2(f_1, f_2) = J_2(f_2, f_1) = 2c(c + 1),\\
&J_2(f_1, f_3) = J_2(f_3, f_1) = 3c(c^2 + 3c + 1),\\
&J_2(f_1, f_4) = J_2(f_4, f_1) = 4c(c^3 + 6c^2 + 6c + 1),\\
&J_2(f_2, f_3) = J_2(f_3, f_2) = 6c(c^3 + 4c^2 + 4c + 1),\\
&J_2(f_2, f_4) = J_2(f_4, f_2) = 8c(c^4 + 7c^3 + 12c^2 + 7c + 1),\\
&J_2(f_3, f_4) = J_2(f_4, f_3) = 12c(c^5 + 9c^4 + 25c^3 + 25c^2 + 9c + 1).
\end{aligned}
$$

Meanwhile, using the Marčenko–Pastur law given by (A3), we can obtain

$$
\begin{aligned}
\beta_1 &= \int f_1(x) dF^{c_n, H_p}(x) = 1,\\
\beta_2 &= \int f_2(x) dF^{c_n, H_p}(x) = 1 + c_n,\\
\beta_3 &= \int f_3(x) dF^{c_n, H_p}(x) = 1 + 3c_n + c_n^2,\\
\beta_4 &= \int f_4(x) dF^{c_n, H_p}(x) = 1 + 6c_n + 6c_n^2 + c_n^3.
\end{aligned}
$$

Appendix C. Proof of Lemma 2.2

Using Lemma 2.1, we have

$$
n(\hat{\beta}_1 - \beta_1, \hat{\beta}_2 - \beta_2, \hat{\beta}_3 - \beta_3, \hat{\beta}_4 - \beta_4)' \xrightarrow{D} N(\mathbf{m}/c, \mathbf{V}/c^2).
$$

Let $\mathbf{u} = (x, y, z, w)'$. Define a vector function

$$
G_n(\mathbf{u}) = \begin{pmatrix} x \\ \tau_2(y - c_n x^2) \\ \tau_3(z - 3c_n xy + 2c_n^2 x^3) \\ \tau_4\left(w - 4c_n xz - \frac{2n^2 + 3n - 6}{n^2 + n + 2} c_n y^2 + \frac{10n^2 + 12n}{n^2 + n + 2} c_n^2 x^2 y - \frac{5n^2 + 6n}{n^2 + n + 2} c_n^3 x^4\right) \end{pmatrix}.
$$

Obviously, $G_n(\mathbf{u})$ has a continuous partial derivative at $\mathbf{b} = (\beta_1, \beta_2, \beta_3, \beta_4)'$ and the Jacobian matrix is

$$J_n(\mathbf{u}) = \frac{\partial G_n(\mathbf{u})}{\partial \mathbf{u}'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2\tau_2 c_n x & \tau_2 & 0 & 0 \\ \tau_3(6c_n^2 x^2 - 3c_n y) & -3\tau_3 c_n x & \tau_3 & 0 \\ \tau_4(-4c_n z + 2a_2 c_n^2 xy + 3a_3 c_n^3 x^3) & \tau_4(-2a_1 c_n y + a_2 c_n^2 x^2) & -4\tau_4 c_n x & \tau_4 \end{pmatrix},$$

where

$$a_1 = \frac{2n^2 + 3n - 6}{n^2 + n + 2}, \; a_2 = \frac{10n^2 + 12n}{n^2 + n + 2}, \; a_3 = \frac{5n^2 + 6n}{n^2 + n + 2}.$$

Using the Delta method, we can obtain, as $(n, p) \to \infty$,

$$n(\hat{\alpha}_1 - 1, \hat{\alpha}_2 - 1, \hat{\alpha}_3 - 1, \hat{\alpha}_4 - 1)' + n((1,1,1,1)' - G_n(\mathbf{b})) \xrightarrow{D} N\big(J(\mathbf{b})\mathbf{m}/c, J(\mathbf{b})\mathbf{V}J'(\mathbf{b})/c^2\big),$$

Some elementary calculations reveal that

$$n((1,1,1,1)' - G_n(\mathbf{b})) \to (0,0,0,0)',$$

and

$$\frac{J(\mathbf{b})\mathbf{m}}{c} := \widetilde{\mathbf{m}} = \Delta \cdot (0, 1, 3, c+6)',$$

$$\frac{J(\mathbf{b})\mathbf{V}J'(\mathbf{b})}{c^2} := \widetilde{\mathbf{V}} = \frac{1}{c}\begin{pmatrix} 2 & 4 & 6 & 8 \\ 4 & 4(c+2) & 12(c+1) & 8(3c+2) \\ 6 & 12(c+1) & 6(c^2+6c+3) & 24(c^2+3c+1) \\ 8 & 8(3c+2) & 24(c^2+3c+1) & 8(c^3+12c^2+18c+4) \end{pmatrix} + \frac{\Delta}{c}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \end{pmatrix}.$$

Therefore,

$$n(\hat{\alpha}_1 - 1, \hat{\alpha}_2 - 1, \hat{\alpha}_3 - 1, \hat{\alpha}_4 - 1)' \xrightarrow{D} N(\widetilde{\mathbf{m}}, \widetilde{\mathbf{V}}).$$

Appendix D. Proof of Theorem 3.1

Let

$$f_1(\mathbf{t}) = \frac{w}{y^2} - \frac{y}{x^2}, \quad \mathbf{t} = (x, y, w)'.$$

It is clear that $f_1(\mathbf{t})$ has a continuous partial derivative at $\mathbf{t_0} = (1, 1, 1)'$, and

$$J(\mathbf{t}) = \frac{\partial f_1(\mathbf{t})}{\partial \mathbf{t}} = \left(\frac{2y}{x^3}, -\frac{2w}{y^3} - \frac{1}{x^2}, \frac{1}{y^2} \right)'.$$

Note that $f(\mathbf{t_0}) = 0$ and $J(\mathbf{t_0}) = (2, -3, 0, 1)'$. Using the Delta method, we have

$$n\left(\frac{\hat{\alpha}_4}{\hat{\alpha}_2^2} - \frac{\hat{\alpha}_2}{\hat{\alpha}_1^2} - f(\mathbf{t_0}) \right) \xrightarrow{D} N(J(\mathbf{t_0})'\widetilde{\mathbf{m}}, J(\mathbf{t_0})'\widetilde{\mathbf{V}}J(\mathbf{t_0})).$$

By simply calculating, we know

$$J(\mathbf{t_0})'\widetilde{\mathbf{m}} = (c+3)\Delta,$$
$$J(\mathbf{t_0})'\widetilde{\mathbf{V}}J(\mathbf{t_0}) = 8c^2 + 96c + 36.$$

Therefore,

$$n\gamma_1 \xrightarrow{D} N(m_1, V_1),$$

where the mean $m_1 = (c+3)\Delta$ and the covariance $V_1 = 8c^2 + 96c + 36$. Thus,

$$T_1 = \frac{n\gamma_1 - m_1}{\sqrt{V_1}} \xrightarrow{D} N(0, 1).$$

Appendix E. Proof of Theorem 3.2

The proof of this theorem is similar to Theorem 3.1. Let

$$f_2(\mathbf{t}) = w - 2z + y, \quad \mathbf{t} = (x, y, z, w)'.$$

It is clear that $f_2(\mathbf{t})$ has a continuous partial derivative at $\mathbf{t_0} = (1, 1, 1, 1)'$, and

$$J(\mathbf{t_0}) = \frac{\partial f_2(\mathbf{t})}{\partial \mathbf{t}}\Big|_{\mathbf{t}=\mathbf{t_0}} = (0, 1, -2, 1)'.$$

Using the Delta method, we have

$$n\gamma_2 \xrightarrow{D} N(m_2, V_2),$$

with the mean $m_2 = J(\mathbf{t_0})'\widetilde{\mathbf{m}} = (c+1)\Delta$, and the covariance $V_2 = J(\mathbf{t_0})'\widetilde{\mathbf{V}}J(\mathbf{t_0}) = 8c^2 + 24c + 4$. Therefore,

$$T_2 = \frac{n\gamma_2 - m_2}{\sqrt{V_2}} \xrightarrow{D} N(0, 1).$$

References

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons, Hoboken, NJ, third edition.

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

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

Bai, Z. D. and Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices.* Springer, New York, NY, second edition.

Bianchi, P., Debbah, M., Maida, M., and Najim, J. (2011). Performance of statistical tests for single-source detection using random matrix theory. *IEEE Transactions on Information Theory*, 57:2400–2419.

Birke, M. and Dette, H. (2005). A note on testing the covariance matrix for large dimension. *Statistics and Probability Letters*, 74:281–289.

Chen, S. X., Zhang, L. X., and Zhong, P. S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association*, 105:810–819.

Couillet, R. and Debbah, M. (2011). *Random Matrix Methods for Wireless Communications.* Cambridge University Press, New York, NY.

Couillet, R. and Hachem, W. (2013). Fluctuations of spiked random matrix models and failure diagnosis in sensor networks. *IEEE Transactions on Information Theory*, 59:509–525.

Dettling, M. and Bühlmann, P. (2003). Boosting for tumor classification with gene expression data. *Bioinformatics*, 19:1061–1069.

Dudoit, S., Fridlyand, J., and Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association*, 97:77–87.

Fisher, T. J. (2012). On testing for an identity covariance matrix when the dimensionality equals or exceeds the sample size. *Journal of Statistical Planning and Inference*, 142:312–326.

Fisher, T. J., Sun, X., and Gallagher, C. M. (2010). A new test for sphericity of the covariance matrix for high dimensional data. *Journal of Multivariate Analysis*, 101:2554–2570.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286:531–537.

Hu, J., Li, W., Liu, Z., and Zhou, W. (2019). High-dimensional covariance matrices in elliptical distributions with application to spherical test. *The Annals of Statistics*, 47:527–555.

John, S. (1971). Some optimal multivariate tests. *Biometrika*, 58:123–127.

Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *The Annals of Statistics*, 29:295–327.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics*, 30:1081–1102.

Nadakuditi, R. and Edelman, A. (2008). Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Transactions on Signal Processing*, 56:2625–2638.

Nagao, H. (1973). On some test criteria for covariance matrix. *The Annals of Statistics*, 1:700–709.

Pan, G. M. and Zhou, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *The Annals of Applied Probability*, 18:1232–1270.

Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55:331–339.

Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of the Japan Statistical Society*, 35:251–272.

Srivastava, M. S. (2006). Some tests criteria for the covariance matrix with fewer observations than the dimension. *Acta Et Commentationes Universitatis Tartuensis De Mathematica*, 10:77–93.

Srivastava, M. S., Kollo, T., and von Rosen, D. (2011). Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis*, 102:1090–1103.

Srivastava, M. S. and Yanagihara, H. (2010). Testing the equality of several covariance matrices with fewer observations than the dimension. *Journal of Multivariate Analysis*, 101:1319–1329.

Tian, X. T., Lu, Y. T., and Li, W. M. (2015). A robust test for sphericity of high-dimensional covariance matrices. *Journal of Multivariate Analysis*, 141:217–227.

Torun, M., Akansu, A., and Avellaneda, M. (2011). Portfolio risk in multiple frequencies. *IEEE Signal Processing Magazine*, 28:61–71.

Wang, C., Yang, J., Miao, B., and Cao, L. (2013). Identity tests for high dimensional data using RMT. *Journal of Multivariate Analysis*, 118:128–137.

Wang, H., Liu, B., Shi, N.-Z., and Zheng, S. (2018). Testing identity of high-dimensional covariance matrix. *Journal of Statistical Computation and Simulation*, 88:2600–2611.

Wang, Q. W. and Yao, J. F. (2013). On the sphericity test with large-dimensional observations. *Electronic Journal of Statistics*, 7:2164–2192.

Zheng, S. R. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 48:444–476.