Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(8): 5200–5216.
DOI: 10.3934/era.2026231
Received: 30 March 2026
Revised: 28 May 2026
Accepted: 09 June 2026
Published: 17 June 2026

Research article

Hypothesis testing for sphericity covariance structure without a normality condition

Jieqiong Shen*

School of Computer and Data Engineering, NingboTech University, Ningbo 315100, China

* **Correspondence:** Email: sj_cde@nbt.edu.cn.

Abstract: This paper investigated the sphericity test for high-dimensional covariance matrices, which is a fundamental problem in probability and statistics courses. While several existing powerful tests in textbooks and literature are available, they often rely on either the normality assumption or an unknown fourth moment. This work aimed to relax these restrictions, serving as a complement and extension of traditional course content on this subject. Specifically, we constructed a consistent estimator related to the fourth moment and proposed a new test procedure. The proposed test statistic depends on the second, third, and fourth arithmetic means of the eigenvalues of the sample covariance matrix, but does not require the normality assumption. Numerical simulations illustrated that the proposed test maintains satisfactory empirical size, and achieves power comparable to, or in some cases superior to, the existing tests.

Keywords: hypothesis test; sphericity; covariance matrix; non-normality; high-dimensional data

1. Introduction

In probability and statistics courses, the covariance matrix is a vital distribution characteristic, and its statistical inference is of central theoretical and practical importance in, e.g., educational sciences, gene analysis, and radar detection [1–6]. In such inferences, the covariance matrix of the population is sometimes assumed to have a spherical structure (e.g., [2]). Validating this assumption via a sphericity test is therefore crucial for students and researchers, as it serves as a critical step to ensure the robustness of the subsequent analyses.

Accordingly, this paper is devoted to testing whether the covariance matrix possesses a spherical structure, which is of particular interest for in-depth course study and application. Let $\boldsymbol{X}_i = (x_{i,1}, \ldots, x_{i,p})^T$, $i = 1, \ldots, N$, be an independent and identically distributed sample from a p-dimensional population with the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, where N denotes the sample size. Our interest is to test the following hypothesis:

$$H_0 : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p \quad \text{vs.} \;\; H_1 : \boldsymbol{\Sigma} \neq \sigma^2 \boldsymbol{I}_p, \tag{1.1}$$

where $\boldsymbol{I}_p$ is the $p \times p$ identity matrix and σ^2 is an unknown but finite positive constant. Several traditional methods as in textbooks and literature are available for testing hypothesis (1.1) in the case that the data dimension p is negligible relative to the sample size N. These include the classical likelihood ratio test (e.g., [7]), and the locally most powerful invariant test [8, 9].

However, the rapid advancement of technology has led to multivariate analysis frequently encountering high-dimensional datasets, a challenge that students and researchers often face in real-world applications. For instance, microarray DNA analysis involves thousands of gene expressions but only a small sample of individuals; in the field of education, the data collected through online learning platforms can encompass a multitude of behavioral and interaction variables despite a small number of students. In such common scenarios, where the data dimension p exceeds or equals the sample size N, the sample covariance matrix becomes singular, rendering the classical likelihood ratio test inapplicable. Faced with this challenge, researchers have proposed several testing methods for the sphericity hypothesis (1.1) in the high-dimensional regime. When p/N tends to $c \in (0, \infty)$, Ledoit and Wolf [10] revisited the test proposed in [8, 9] and demonstrated the robustness of the test statistic. This finding was subsequently extended by Birke and Dette in [11] to accommodate $c \in [0, \infty]$. Meanwhile, Srivastava [12] proposed a test statistic based on the unbiased estimators of $\mathrm{tr}(\boldsymbol{\Sigma})$ and $\mathrm{tr}(\boldsymbol{\Sigma}^2)$, constructed by the first and second arithmetic means of the eigenvalues of the sample covariance matrix. Building upon this idea, Fisher et al. [13] later developed another test statistic by employing the unbiased estimators for $\mathrm{tr}(\boldsymbol{\Sigma}^2)$ and $\mathrm{tr}(\boldsymbol{\Sigma}^4)$ that involve the second and fourth arithmetic means of the sample eigenvalues.

It is worth noting that although the tests mentioned above successfully overcome the limitation of the classical regime, their performance can be adversely affected in non-normal settings due to their reliance on the normality assumption. However, such an assumption can be relaxed in some sense. For instance, Srivastava et al. [14] proposed an improved version of the test statistic from [12] for general populations by utilizing a more accurate and reliable estimator of $\mathrm{tr}(\boldsymbol{\Sigma}^2)$, Chen et al. [15] proposed another test statistic by virtue of some well-selected U-statistics, and Wang and Yao [16] modified the likelihood ratio test proposed in [10] using random matrix theory. With a finite fourth moment, Tian et al. [17] constructed a test statistic by taking the maximum of two test statistics proposed in [12, 13]. Yuan et al. [18] developed a test statistic based on random matrix theory that involves the first four arithmetic means of the sample eigenvalues.

Numerical simulations indicate that the tests proposed in [13, 17, 18] achieve higher power than those in [10, 12, 16], especially when the covariance matrix deviates from the sphericity structure only in a few elements. This advantage of the former may be attributed to their leveraging of higher arithmetic means of the sample eigenvalues. However, the practical application of these powerful tests is limited by their reliance on either the normality assumption [13] or a finite fourth moment [17, 18]. We remark that such a finite fourth moment is unknown due to the absence of the underlying distribution, a fact that is also a difficulty often encountered by students in coursework and practical training. To address this gap, we aim to develop a new sphericity test that relaxes the normality assumption by constructing a consistent estimator related to the unknown fourth moment. This work will deepen and extend the traditional pedagogical content on hypothesis testing for covariance matrices in probability and statistics courses.

The remainder of this paper is organized as follows. Section 2 provides the preliminaries necessary for subsequent developments. In Section 3, a new test procedure for hypothesis (1.1) is proposed by establishing a consistent estimator related to the unknown fourth moment. Section 4 reports numerical simulations designed to evaluate the finite-sample behavior of the proposed test. Section 5 concludes the paper with a discussion.

2. Notations and preliminaries

This section provides the necessary preliminaries for the subsequent analysis. Let $\boldsymbol{S}$ be the sample covariance matrix, defined as

$$\boldsymbol{S} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T, \tag{2.1}$$

where $\bar{\boldsymbol{X}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{X}_i$. Throughout the following, we abbreviate

$$a_j = \frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^j), \qquad j = 1, 2, 3, 4. \tag{2.2}$$

The estimators $\hat{a}_j$ for a_j are then stated as follows:

$$\hat{a}_1 = \frac{1}{p}\mathrm{tr}(\boldsymbol{S}), \tag{2.3}$$

$$\hat{a}_2 = \frac{n^2}{p(n-1)(n+2)}\left(\mathrm{tr}(\boldsymbol{S}^2) - \frac{1}{n}(\mathrm{tr}(\boldsymbol{S}))^2\right), \tag{2.4}$$

$$\hat{a}_3 = \frac{\tau_1}{p}\left(\mathrm{tr}(\boldsymbol{S}^3) - \frac{3}{n}\mathrm{tr}(\boldsymbol{S}^2)\mathrm{tr}(\boldsymbol{S}) + \frac{2}{n^2}(\mathrm{tr}(\boldsymbol{S}))^3\right), \tag{2.5}$$

$$\begin{aligned}\hat{a}_4 = \frac{\tau_2}{p}\Big(&\mathrm{tr}(\boldsymbol{S}^4) - \frac{4}{n}\mathrm{tr}(\boldsymbol{S}^3)\mathrm{tr}(\boldsymbol{S}) - \frac{2n^2+3n-6}{n^3+n^2+2n}(\mathrm{tr}(\boldsymbol{S}^2))^2 \\ &+ \frac{10n+12}{n^3+n^2+2n}\mathrm{tr}(\boldsymbol{S}^2)(\mathrm{tr}(\boldsymbol{S}))^2 - \frac{5n+6}{n^4+n^3+2n^2}(\mathrm{tr}(\boldsymbol{S}))^4\Big),\end{aligned} \tag{2.6}$$

where $n = N - 1$, and

$$\begin{aligned}\tau_1 &= \frac{n^4}{(n-1)(n-2)(n+2)(n+4)}, \\ \tau_2 &= \frac{n^7+n^6+2n^5}{(n+1)(n+2)(n+4)(n+6)(n-1)(n-2)(n-3)}.\end{aligned}$$

It is worth pointing out that the estimators $\hat{a}_1$ and $\hat{a}_2$ are given in [12], whereas $\hat{a}_3$ and $\hat{a}_4$ are provided in [19] and [13], respectively. Moreover, these estimators $\hat{a}_j$ ($j = 1, 2, 3, 4$) are consistent under some mild assumptions. To ensure these properties, we work under the assumptions laid out in [17], which can be stated as follows.

Assumption 1. *Both N and p tend to infinity with $p/N \to c \in (0, \infty)$.*

Assumption 2. *The random vectors $\boldsymbol{X}_j$ are represented as $\boldsymbol{X}_j = \boldsymbol{\Sigma}^{1/2}\boldsymbol{W}_{\cdot j} + \mu$, where $\boldsymbol{W}_{\cdot j}$ denotes the j-th column of $\boldsymbol{W}$. Here $\boldsymbol{W} := (w_{ij})_{1\le i\le p, 1\le j\le N}$ and $\{w_{ij}\}$ are independent and identically distributed, satisfying*

$$\mathbb{E}(w_{11}) = 0, \ \ \mathbb{E}(w_{11}^2) = 1, \ \ \mathbb{E}(w_{11}^4) < \infty.$$

Assumption 3. *Define the spectral distribution of the population covariance matrix $\boldsymbol{\Sigma}$ as*

$$H_p(x) = \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}_{\{\lambda_i \le x\}},$$

where λ_i, $i = 1, \ldots, p$, are the eigenvalues of $\boldsymbol{\Sigma}$. We assume that H_p converges weakly to a non-random distribution H as $p \to \infty$, and that the spectral norm of $\boldsymbol{\Sigma}$ is uniformly bounded in p, i.e., there exists a constant $C > 0$ independent of p such that $\|\boldsymbol{\Sigma}\|_2 \le C$.

Under Assumptions 1–3, one can verify that $\hat{a}_1$ is unbiased and consistent, while $\hat{a}_2$, $\hat{a}_3$, and $\hat{a}_4$ are consistent but biased estimators of a_2, a_3, and a_4, respectively. Recognizing the bias of $\hat{a}_2$, Srivastava et al. [14] proposed an unbiased estimator for a_2, given by

$$\hat{a}_{2s} = \frac{1}{pf}\Big((N-2)(N-1)^3\mathrm{tr}(\boldsymbol{S}^2) - N(N-1)\mathrm{tr}(\boldsymbol{D}^2) + (N-1)^2(\mathrm{tr}(\boldsymbol{S}))^2\Big), \tag{2.7}$$

where $f = N(N-1)(N-2)(N-3)$ and

$$D = \mathrm{diag}\Big((\boldsymbol{X}_1 - \bar{\boldsymbol{X}})^\top(\boldsymbol{X}_1 - \bar{\boldsymbol{X}}), \ldots, (\boldsymbol{X}_N - \bar{\boldsymbol{X}})^\top(\boldsymbol{X}_N - \bar{\boldsymbol{X}})\Big).$$

For further details, the reader is referred to [12, 13, 17, 19].

3. Test procedure

Note that the hypothesis problem (1.1) remains invariant under the scalar transformation $x \to ax$ and the orthogonal transformation $x \to \boldsymbol{G}x$, where $a \ne 0$ and $\boldsymbol{G}$ is an orthogonal matrix. Therefore, we can assume without loss of generality that $\boldsymbol{\Sigma} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ with $\lambda_i > 0$ for $i = 1, \ldots, p$. By employing Hölder's inequality, we have

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

It is evident that the equality holds if and only if $\lambda_1, \ldots, \lambda_p$ are equal. The measure δ between the null and alternative hypotheses is therefore characterized by

$$\delta = \frac{a_2 a_4}{a_3^2},$$

where a_j is given by (2.2), $j = 2, 3, 4$. Note that $\delta = 1$ if H_0 is true, while $\delta > 1$ if H_1 holds. The corresponding test statistic, based on the estimator of δ, is then constructed by

$$T = \frac{\hat{a}_2\hat{a}_4}{\hat{a}_3^2}, \tag{3.1}$$

where $\hat{a}_2$, $\hat{a}_3$, and $\hat{a}_4$ are given by (2.4), (2.5), and (2.6), respectively. The following theorem establishes the asymptotic normality of the test statistic T.

Theorem 3.1. *Under Assumptions* 1–3, *if* $H_0 : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p$ *holds, then we have*

$$N(T-1) \xrightarrow{D} \mathcal{N}(\Delta, 8c^2 + 24c + 4),$$

where $\Delta = \mathbb{E}(w_{11}^4) - 3$.

Proof. Letting $f(\boldsymbol{t}) = yw/z^2$ with $\boldsymbol{t} = (x, y, z, w)^T$, we have

$$\frac{\partial f(\boldsymbol{t})}{\partial x} = 0, \quad \frac{\partial f(\boldsymbol{t})}{\partial y} = \frac{w}{z^2}, \quad \frac{\partial f(\boldsymbol{t})}{\partial z} = \frac{-2yw}{z^3}, \quad \frac{\partial f(\boldsymbol{t})}{\partial w} = \frac{y}{z^2}.$$

Obviously, $f(\boldsymbol{t})$ has a continuous partial derivative at $\boldsymbol{t}_0 = (a_1, a_2, a_3, a_4)^T$ with the Jacobian vector

$$\boldsymbol{J}(\boldsymbol{t_0}) = \frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}_0} = \left(0, \frac{a_4}{a_3^2}, \frac{-2a_2a_4}{a_3^3}, \frac{a_2}{a_3^2}\right)^T.$$

By employing Lemma 2 in [17], under $H_0 : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p$, it holds that

$$N\left(\hat{a}_1 - \sigma^2,\ \hat{a}_2 - \sigma^4,\ \hat{a}_3 - \sigma^6,\ \hat{a}_4 - \sigma^8\right)^T \xrightarrow{D} \mathcal{N}_4\left(\boldsymbol{\nu}, \boldsymbol{B}\right),$$

where $\boldsymbol{\nu} = \Delta\left(0, \sigma^4, 3\sigma^6, 6\sigma^8\right)^T$ and

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}.$$

The matrix $\boldsymbol{B}$ is symmetric, in which

$$\begin{aligned}
b_{11} &= (2 + \Delta)\sigma^4/c, \\
b_{12} &= (4 + 2\Delta)\sigma^6/c, \\
b_{13} &= (6 + 3\Delta)\sigma^8/c, \\
b_{14} &= (8 + 4\Delta)\sigma^{10}/c, \\
b_{22} &= (8 + 4c + 4\Delta)\sigma^8/c, \\
b_{23} &= (12 + 12c + 6\Delta)\sigma^{10}/c, \\
b_{24} &= (16 + 24c + 8\Delta)\sigma^{12}/c, \\
b_{33} &= (18 + 36c + 6c^2 + 9\Delta)\sigma^{12}/c, \\
b_{34} &= (24 + 72c + 24c^2 + 12\Delta)\sigma^{14}/c, \\
b_{44} &= (32 + 144c + 96c^2 + 8c^3 + 16\Delta)\sigma^{16}/c.
\end{aligned}$$

On the other hand, under $H_0 : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p$, we obtain that

$$\boldsymbol{J}(\boldsymbol{t_0})|_{H_0} = \left(0, \frac{1}{\sigma^4}, \frac{-2}{\sigma^6}, \frac{1}{\sigma^8}\right)^T, \quad \frac{a_2a_4}{a_3^2} = 1.$$

By employing the delta method, it follows that

$$N\left(\frac{\hat{a}_2\hat{a}_4}{\hat{a}_3^2}-1\right)\xrightarrow{D}\mathcal{N}(\tau,\boldsymbol{D}),$$

where

$$\begin{aligned}\tau &= \left(\boldsymbol{J}(\boldsymbol{t_0})|_{H_0}\right)^T\boldsymbol{\nu}\\ &= \left(0,\frac{1}{\sigma^4},\frac{-2}{\sigma^6},\frac{1}{\sigma^8}\right)\Delta\left(0,\sigma^4,3\sigma^6,6\sigma^8\right)^T\\ &= \Delta\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{D} &= \left(\boldsymbol{J}(\boldsymbol{t_0})|_{H_0}\right)^T\boldsymbol{B}\left(\boldsymbol{J}(\boldsymbol{t_0})|_{H_0}\right)\\ &= \left(0,\frac{1}{\sigma^4},\frac{-2}{\sigma^6},\frac{1}{\sigma^8}\right)\boldsymbol{B}\left(0,\frac{1}{\sigma^4},\frac{-2}{\sigma^6},\frac{1}{\sigma^8}\right)^T\\ &= 8c^2+24c+4.\end{aligned}$$

Therefore, this completes the proof of Theorem 3.1. □

A key observation is that, similar to the tests in [17,18], the asymptotic normality of our test statistic T involves the quantity $\Delta=\mathbb{E}(w_{11}^4)-3$. Note that Δ is inherently distribution-dependent, equaling zero in the case of a standard normal distribution. This dependence limits the test's applicability due to the uncertainty regarding the underlying distribution. To resolve this issue, we will construct a consistent estimator for Δ, and thereby obtain a modified version of the proposed test. As a first step, we derive the expression for Δ. Observe that

$$\begin{aligned}\mathbb{E}(\boldsymbol{W}_{.j}^T\boldsymbol{\Sigma}\boldsymbol{W}_{.j})^2 &= \mathbb{E}\left(\sum_{i=1}^p w_{ij}^2\sigma_{ii}+\sum_{i\neq k}^p w_{ij}w_{kj}\sigma_{ik}\right)^2\\ &= \sum_{i=1}^p\sigma_{ii}^2\mathbb{E}\left(w_{ij}^4\right)+\sum_{i\neq k}^p\sigma_{ii}\sigma_{kk}\mathbb{E}\left(w_{ij}^2w_{kj}^2\right)+\sum_{i\neq k}^p\left(\sigma_{ik}^2+\sigma_{ik}\sigma_{ki}\right)\mathbb{E}\left(w_{ij}^2w_{kj}^2\right)\\ &= (\Delta+3)\sum_{i=1}^p\sigma_{ii}^2+\sum_{i,k=1}^p\sigma_{ii}\sigma_{kk}-\sum_{i=1}^p\sigma_{ii}^2+2\left(\sum_{i,k=1}^p\sigma_{ik}^2-\sum_{i=1}^p\sigma_{ii}^2\right)\\ &= (\Delta+3)\,\mathrm{tr}(\boldsymbol{\Sigma}\odot\boldsymbol{\Sigma})+(\mathrm{tr}(\boldsymbol{\Sigma}))^2-\mathrm{tr}(\boldsymbol{\Sigma}\odot\boldsymbol{\Sigma})+2\mathrm{tr}(\boldsymbol{\Sigma}^2)-2\mathrm{tr}(\boldsymbol{\Sigma}\odot\boldsymbol{\Sigma})\\ &= \Delta\mathrm{tr}(\boldsymbol{\Sigma}\odot\boldsymbol{\Sigma})+(\mathrm{tr}(\boldsymbol{\Sigma}))^2+2\mathrm{tr}(\boldsymbol{\Sigma}^2),\end{aligned}$$

which yields the desired expression for Δ, denoted by

$$\Delta=\frac{\kappa}{\mathrm{tr}\left(\boldsymbol{\Sigma}\odot\boldsymbol{\Sigma}\right)},$$

where

$$\kappa:=\mathbb{E}(\boldsymbol{W}_{.j}^T\boldsymbol{\Sigma}\boldsymbol{W}_{.j})^2-(\mathrm{tr}(\boldsymbol{\Sigma}))^2-2\mathrm{tr}(\boldsymbol{\Sigma}^2).$$

We now seek to develop the estimator for $\mathrm{tr}(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})$. In fact, its estimator exists in [20] and is denoted by

$$\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)} = \sum_{l=1}^{p} \left(\frac{1}{N(N-1)} \sum_{a,b}^{*} x_{a,l}^2 x_{b,l}^2 - \frac{2}{N(N-1)(N-2)} \sum_{a,b,c}^{*} x_{a,l} x_{b,l} x_{c,l}^2 + \frac{1}{N(N-1)(N-2)(N-3)} \sum_{a,b,c,d}^{*} x_{a,l} x_{b,l} x_{c,l} x_{d,l} \right),$$

where $\sum^*$ denotes summation over mutually different indices. We note that the estimator $\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)}$ is prohibitively expensive to compute for large N and p, as it involves $O\left(N^4 p\right)$ multiplications. This computational burden motivates us to find a more efficient unbiased estimator for $\mathrm{tr}(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})$. Observe that $\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)}$ is invariant with respect to the location transformation that maps $x_{a,l}$ to $x_{a,l} + m$ for an arbitrary constant m. By this invariance, it can be equivalently expressed as

$$\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)} = \sum_{l=1}^{p} \left(\frac{1}{N(N-1)} \sum_{a,b}^{*} \tilde{x}_{a,l}^2 \tilde{x}_{b,l}^2 - \frac{2}{N(N-1)(N-2)} \sum_{a,b,c}^{*} \tilde{x}_{a,l} \tilde{x}_{b,l} \tilde{x}_{c,l}^2 + \frac{1}{N(N-1)(N-2)(N-3)} \sum_{a,b,c,d}^{*} \tilde{x}_{a,l} \tilde{x}_{b,l} \tilde{x}_{c,l} \tilde{x}_{d,l} \right), \tag{3.2}$$

where $\tilde{x}_{a,l} = x_{a,l} - \bar{x}_l$ with $\bar{x}_l = \frac{1}{N} \sum_{a=1}^{N} x_{a,l}$. Denoting the (l_1, l_2)-th entry of the sample covariance matrix $\boldsymbol{S}$ by

$$s_{l_1,l_2} = \frac{1}{N-1} \sum_{a=1}^{N} \tilde{x}_{a,l_1} \tilde{x}_{a,l_2}, \tag{3.3}$$

a straightforward calculation shows that

$$\sum_{a,b}^{*} \tilde{x}_{a,l}^2 \tilde{x}_{b,l}^2 = \left(\sum_{a=1}^{N} \tilde{x}_{a,l}^2 \right)^2 - \sum_{a=1}^{N} \tilde{x}_{a,l}^4 = (N-1)^2 s_{l,l}^2 - v_{l,l}, \tag{3.4}$$

$$\sum_{a,b,c}^{*} \tilde{x}_{a,l} \tilde{x}_{b,l} \tilde{x}_{c,l}^2 = 2 \sum_{a=1}^{N} \tilde{x}_{a,l}^4 - \left(\sum_{a=1}^{N} \tilde{x}_{a,l}^2 \right)^2 = 2 v_{l,l} - (N-1)^2 s_{l,l}^2, \tag{3.5}$$

$$\sum_{a,b,c,d}^{*} \tilde{x}_{a,l} \tilde{x}_{b,l} \tilde{x}_{c,l} \tilde{x}_{d,l} = -6 \sum_{a=1}^{N} \tilde{x}_{a,l}^4 + 3 \left(\sum_{a=1}^{N} \tilde{x}_{a,l}^2 \right)^2 = -6 v_{l,l} + 3(N-1)^2 s_{l,l}^2, \tag{3.6}$$

where $v_{l,l} = \sum_{a=1}^{N} \tilde{x}_{a,l}^4$. By substituting (3.4), (3.5), and (3.6) into (3.2), an equivalent form of $\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)}$ is given by

$$\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})} = \sum_{l=1}^{p} \left(\frac{(N-1)(N^2 - 3N + 3)}{N(N-2)(N-3)} s_{l,l}^2 - \frac{1}{(N-2)(N-3)} v_{l,l} \right). \tag{3.7}$$

We emphasize that although $\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}$ is equivalent to $\mathrm{tr}\,\widehat{(\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma})}_{(1)}$, the former requires only $O(Np)$ multiplications, and thereby is computationally more efficient than the estimator given in [20]. The proposed estimator is of independent methodological interest; moreover, its construction strategy can serve as a pedagogical reference for constructing related estimators in probability and statistics courses. Keeping the estimator (3.7) in mind, an estimator of Δ follows.

Theorem 3.2. *Under Assumptions* 1–3*, a ratio consistent estimator of* Δ *can be constructed as*

$$\hat{\Delta} = \frac{\hat{\kappa}}{\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})}, \tag{3.8}$$

where $\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})$ *is given by* (3.7) *and*

$$\hat{\kappa} = \frac{-1}{(N-2)(N-3)}\left(2(N-1)^2\mathrm{tr}(\boldsymbol{S}^2) + (N-1)^2\,(\mathrm{tr}(\boldsymbol{S}))^2 - N(N+1)Q\right),$$

with

$$Q = \frac{1}{N-1}\sum_{i=1}^{N}\left(\left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)^T\left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)\right)^2.$$

Proof. Indeed, it was proved by Himeno and Yamada in [21] that $\hat{\kappa}$ is a ratio-consistent estimator of κ, and the ratio-consistency of $\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})_{(1)}$ is established by Xu in [20], i.e.,

$$\frac{\hat{\kappa}}{\kappa} \xrightarrow{P} 1, \qquad \frac{\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})_{(1)}}{\mathrm{tr}\,(\mathbf{\Sigma} \odot \mathbf{\Sigma})} \xrightarrow{P} 1.$$

By recalling that $\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})_{(1)} = \mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})$, we have

$$\frac{\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})}{\mathrm{tr}\,(\mathbf{\Sigma} \odot \mathbf{\Sigma})} \xrightarrow{P} 1.$$

By means of the definitions of Δ and $\hat{\Delta}$, it follows that

$$\frac{\hat{\Delta}}{\Delta} = \frac{\hat{\kappa}}{\kappa} \times \frac{\mathrm{tr}\,(\mathbf{\Sigma} \odot \mathbf{\Sigma})}{\mathrm{tr}\,(\widehat{\mathbf{\Sigma} \odot \mathbf{\Sigma}})} \xrightarrow{P} 1.$$

This completes the proof of Theorem 3.2. □

Remark 3.1. *Note that based on consistent yet biased estimators for* κ *and* $\mathrm{tr}(\Sigma \odot \Sigma)$*, Lopes et al. [22] constructed a consistent estimator for* Δ *under the zero-mean condition, which was later extended to the nonzero-mean case in [23]. In contrast to these existing approaches, in our framework, the estimators for* κ *and* $\mathrm{tr}(\Sigma \odot \Sigma)$ *are both unbiased and consistent.*

By invoking Theorem 3.1 and Theorem 3.2, a new test statistic for the hypothesis problem (1.1) can be constructed as

$$T_{new} = \frac{N(T-1) - \hat{\Delta}}{\sqrt{8c^2 + 24c + 4}}, \tag{3.9}$$

where T and $\hat{\Delta}$ are given by (3.1) and (3.8), respectively. Moreover, under $H_0 : \mathbf{\Sigma} = \sigma^2\boldsymbol{I}_p$, it holds that

$$T_{new} \xrightarrow{D} \mathcal{N}(0,1).$$

Hence, we can test the hypothesis problem (1.1) at significance level α by

$$\text{rejecting } H_0 \Leftrightarrow T_{new} > Z_\alpha,$$

where Z_α is the upper α quantile of the standard normal distribution.

4. Numerical simulations

In this section, we assess the performance of the proposed test via Monte Carlo simulations. The simulation study serves two main purposes: (i) to evaluate the performance of the consistent estimator $\hat{\Delta}$ defined in (3.8); (ii) to examine the empirical size and power of the proposed test statistic T_{new} given by (3.9) and to compare it with several existing tests. All results are based on 1000 replications.

4.1. Evaluate the consistence of $\hat{\Delta}$

To investigate the performance of the estimator $\hat{\Delta}$, we consider the following simulation setup. Without loss of generality, we set $\boldsymbol{\mu} = \mathbf{0}$ and consider two covariance structures: $\boldsymbol{\Sigma} = 2\boldsymbol{I}_p$ (spherical) and $\boldsymbol{\Sigma} = 0.8\boldsymbol{I}_p + 0.2\boldsymbol{J}_p$ (non-spherical), where $\boldsymbol{J}_p$ is the $p \times p$ matrix of all ones. We generate w_{ij} from the following six distinct distributions:

Case 1. *w_{ij} follows the standard normal distribution, i.e., $w_{ij} \sim \mathcal{N}(0,1)$, in which case $\Delta = 0$.*

Case 2. *w_{ij} follows the standardized Gamma$(5,2)$, i.e., $w_{ij} \sim (5/4)^{-1/2}(\mathcal{G}(5,2) - 5/2)$, in which case $\Delta = 1.2$.*

Case 3. *w_{ij} follows the standardized Gamma$(4,2)$, i.e., $w_{ij} \sim \mathcal{G}(4,2) - 2$, in which case $\Delta = 1.5$.*

Case 4. *w_{ij} follows the standardized Uniform$(-\sqrt{3}, \sqrt{3})$, i.e., $w_{ij} \sim \mathcal{U}(-\sqrt{3}, \sqrt{3})$, in which case $\Delta = -1.2$.*

Case 5. *w_{ij} follows the standardized Student's t(7), i.e., $w_{ij} \sim (7/5)^{-1/2}t(7)$, in which case $\Delta = 2$.*

Case 6. *w_{ij} follows the standardized contaminated normal distribution, i.e.,*

$$w_{ij} \sim (1 - \epsilon + \epsilon b^2)^{-1/2}\left((1-\epsilon)\mathcal{N}(0,1) + \epsilon\mathcal{N}(0,b^2)\right)$$

with $\epsilon = 0.05$ and $b = 3.3$, in which case $\Delta \approx 6.2$.

Figure 1 displays the values of the estimator $\hat{\Delta}$ under Cases 1–6 for varying sample size N, while Figure 2 presents the corresponding results for increasing dimension p. From the figures, we find that $\hat{\Delta}$ remains close to the true value Δ across all six cases. Furthermore, the approximation improves as either N or p increases, providing strong empirical support for the consistency of $\hat{\Delta}$ established in Theorem 3.2.

We further compare the proposed estimator $\hat{\Delta}$ with two existing estimators: $\hat{\Delta}_L$ from [22] and $\hat{\Delta}_D$ from [23]. Figure 3 exhibits the values of these estimators under the normal distribution for the non-spherical covariance structure, while Figure 4 illustrates the corresponding results under the non-normal distribution for the spherical structure. As shown in the figures, all three estimators converge to the true value Δ as N increases. However, for small N, $\hat{\Delta}$ outperforms both $\hat{\Delta}_L$ and $\hat{\Delta}_D$, highlighting its advantage in smaller sample sizes. The improved performance of $\hat{\Delta}$ may be attributed to the use of unbiased and consistent estimators for both the numerator and denominator of Δ, in contrast to competing methods that rely on biased estimators for these components. Although the ratio of two unbiased estimators is not necessarily unbiased, our strategy appears to mitigate some of the bias that could accumulate from separate estimation, thereby enhancing finite-sample performance.

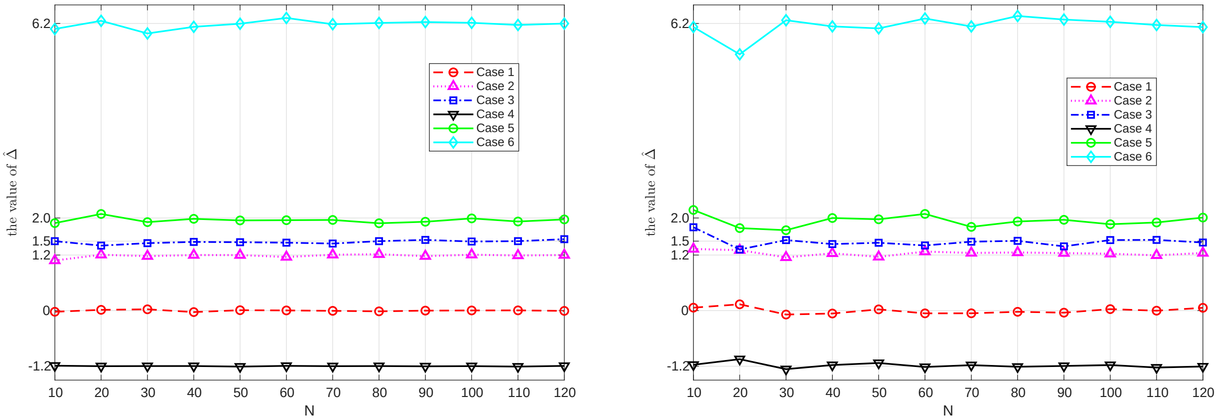

Figure 1. The estimator $\hat{\Delta}$ under Cases 1–6 for $p = 60$ with the true value $\Delta = 0, 1.2, 1.5, -1.2, 2, 6.2$, respectively. The left panel corresponds to $\mathbf{\Sigma} = 2\boldsymbol{I}_p$ and the right panel to $\mathbf{\Sigma} = 0.8\boldsymbol{I}_p + 0.2\boldsymbol{J}_p$.

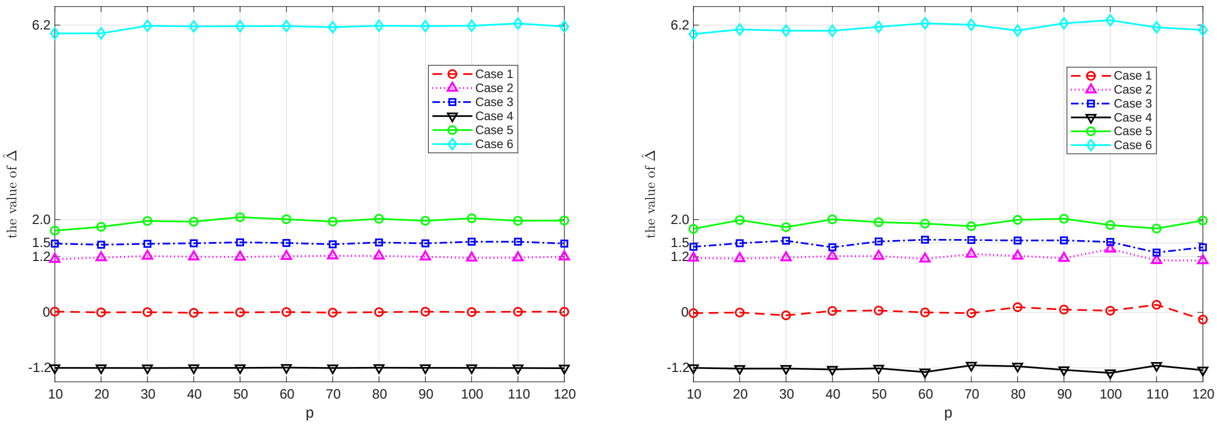

Figure 2. The estimator $\hat{\Delta}$ under Cases 1–6 for $N = 100$ with the true value $\Delta = 0, 1.2, 1.5, -1.2, 2, 6.2$, respectively. The left panel corresponds to $\mathbf{\Sigma} = 2\boldsymbol{I}_p$ and the right panel to $\mathbf{\Sigma} = 0.8\boldsymbol{I}_p + 0.2\boldsymbol{J}_p$.

4.2. The proposed test vs. the existing tests

In this subsection, we evaluate the empirical size and power of the proposed test T_{new} and compare it with two existing tests, referred to as T_F in [13] and T_S in [14]. For the sake of reference, the explicit

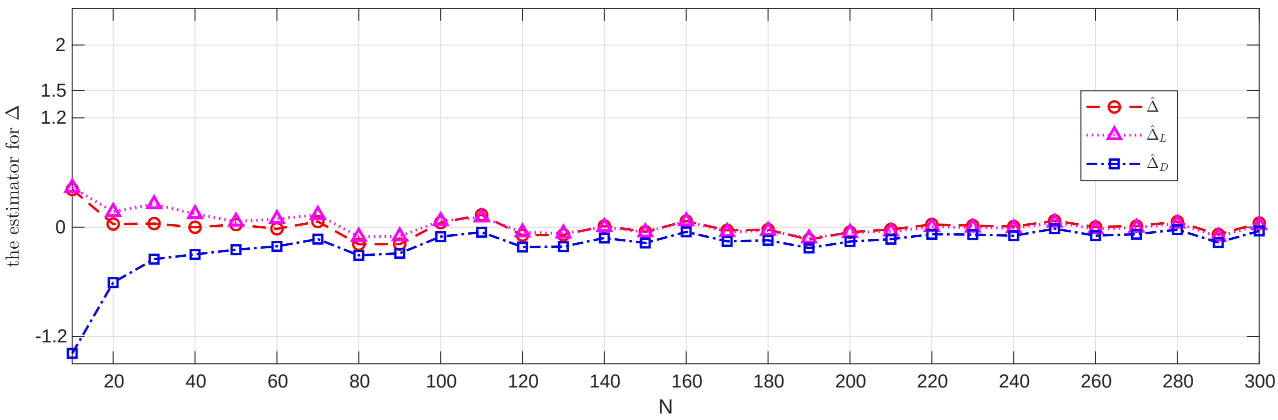

Figure 3. The estimators $\hat{\Delta}$, $\hat{\Delta}_L$, and $\hat{\Delta}_D$ under Case 1 for $\mathbf{\Sigma} = 0.8\boldsymbol{I}_p + 0.2\boldsymbol{J}_p$ with $p = 100$. The true value $\Delta = 0$.

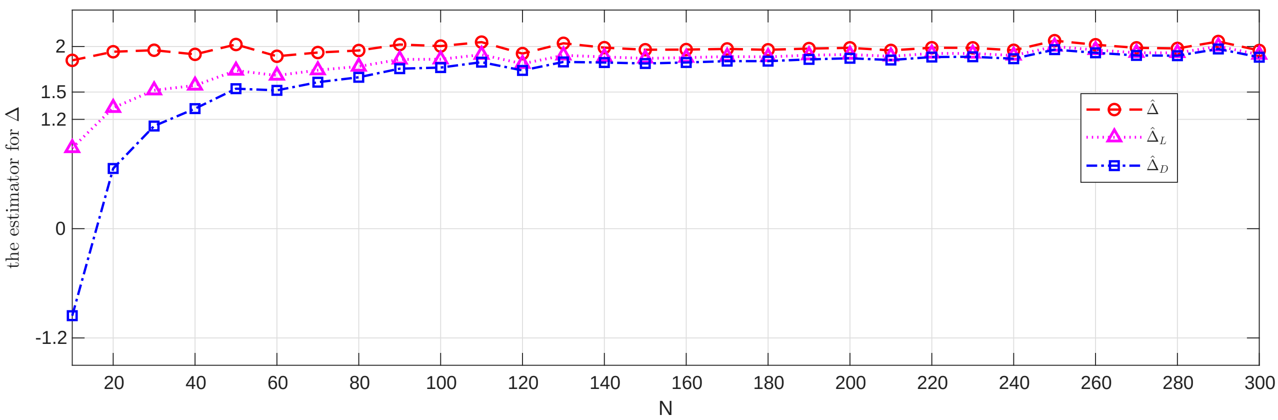

Figure 4. The estimators $\hat{\Delta}$, $\hat{\Delta}_L$, and $\hat{\Delta}_D$ under Case 5 for $\mathbf{\Sigma} = 2\boldsymbol{I}_p$ with $p = 60$. The true value $\Delta = 2$.

forms of T_F and T_S are presented below:

$$T_F = \frac{N-1}{\sqrt{8c^2 + 96c + 64}}\left(\frac{\hat{a}_4}{\hat{a}_2^2} - 1\right), \qquad T_S = \frac{N-1}{2}\left(\frac{\hat{a}_{2s}}{\hat{a}_1^2} - 1\right),$$

where $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_{2s}$, and $\hat{a}_4$ are given in (2.3), (2.4), (2.7), and (2.6), respectively.

To examine the empirical size, we consider the null hypothesis $H_0 : \mathbf{\Sigma} = \boldsymbol{I}_p$, while the following three alternative hypotheses are used to evaluate the empirical power:

$$\text{(I)} \qquad H_1 : \mathbf{\Sigma} = \mathrm{diag}(3, \underbrace{1, \cdots, 1}_{p-1}),$$

$$\text{(II)} \qquad H_1 : \mathbf{\Sigma} = \mathrm{diag}(1.2, 2.6, 3.8, \underbrace{1, \cdots, 1}_{p-3}),$$

$$\text{(III)} \qquad H_1 : \mathbf{\Sigma} = (1-\rho)\boldsymbol{I}_p + \rho\boldsymbol{J}_p.$$

Here, we set the nominal significance level $\alpha = 0.05$ and $\rho = 0.02$.

Table 1 presents the empirical sizes of the tests T_F, T_S, and T_{new} under the null hypothesis for Cases 1, 2, and 5. As shown in the table, under the normal distribution, all three tests control the empirical sizes well around the nominal level of 0.05. Under the gamma distribution and Student's t distribution, however, T_F exhibits empirical sizes roughly between 0.1 and 0.2, indicating a clear deviation from the nominal level. In contrast, T_S and T_{new} maintain empirical sizes well-controlled around 0.05. This is unsurprising, as T_F is derived under the normality assumption, whereas the other two tests do not rely on this assumption and are therefore more robust to the underlying distribution.

Table 1. Empirical sizes of the tests T_F, T_S, and T_{new} under H_0.

p	N	Normal			Gamma(5, 2)			Student's $t(7)$		
		T_F	T_S	T_{new}	T_F	T_S	T_{new}	T_F	T_S	T_{new}
20	60	0.065	0.069	0.038	0.122	0.071	0.036	0.186	0.060	0.042
	80	0.054	0.061	0.039	0.133	0.071	0.040	0.182	0.057	0.039
	100	0.046	0.063	0.051	0.128	0.057	0.046	0.199	0.062	0.061
	120	0.047	0.053	0.041	0.124	0.056	0.051	0.216	0.072	0.069
	160	0.056	0.069	0.045	0.121	0.049	0.042	0.209	0.063	0.059
80	60	0.054	0.060	0.036	0.102	0.066	0.036	0.149	0.047	0.038
	80	0.058	0.052	0.041	0.114	0.063	0.042	0.185	0.044	0.056
	100	0.051	0.054	0.048	0.127	0.041	0.051	0.198	0.053	0.058
	120	0.051	0.056	0.042	0.118	0.054	0.052	0.169	0.061	0.055
	160	0.052	0.048	0.046	0.107	0.052	0.041	0.187	0.056	0.061
100	60	0.052	0.058	0.036	0.108	0.057	0.040	0.157	0.059	0.045
	80	0.042	0.066	0.046	0.112	0.048	0.038	0.150	0.061	0.044
	100	0.046	0.058	0.041	0.113	0.060	0.045	0.193	0.058	0.054
	120	0.052	0.046	0.052	0.108	0.052	0.042	0.178	0.056	0.051
	160	0.052	0.042	0.045	0.112	0.056	0.051	0.197	0.057	0.058
120	60	0.062	0.056	0.038	0.097	0.050	0.039	0.175	0.056	0.042
	80	0.045	0.063	0.046	0.098	0.046	0.036	0.188	0.056	0.051
	100	0.046	0.065	0.051	0.086	0.052	0.042	0.175	0.061	0.055
	120	0.055	0.062	0.052	0.111	0.054	0.046	0.193	0.051	0.061
	160	0.056	0.045	0.058	0.115	0.052	0.051	0.184	0.057	0.048
160	60	0.051	0.045	0.041	0.098	0.052	0.038	0.146	0.066	0.042
	80	0.056	0.064	0.045	0.113	0.061	0.041	0.147	0.058	0.045
	100	0.042	0.054	0.043	0.096	0.058	0.042	0.180	0.057	0.066
	120	0.056	0.058	0.059	0.095	0.051	0.052	0.159	0.049	0.051
	160	0.055	0.056	0.046	0.105	0.047	0.056	0.179	0.065	0.058

Tables 2, 3, and 4 reveal the empirical powers of T_F, T_S, and T_{new} under the three alternative hypotheses for Cases 1, 2, and 5, respectively. Overall, the proposed test T_{new} performs comparably to the existing tests: it is slightly less powerful than T_F but outperforms T_S across these alternatives. We note that the power of the test depends not only on p and N but also on the specific alternative hypothesis. Specifically, under the alternatives (I) and (II), the largest eigenvalue is fixed while the noise dimension increases with p, which dilutes the signal and reduces power. In contrast, under the compound symmetry alternative (III), the largest eigenvalue grows linearly with p, so power increases with p; whereas when p is small, the deviation from sphericity is minimal, resulting in low power.

As a complement, we further compare the proposed test T_{new} with T_F by varying the degrees of freedom ν of Student's t distribution, with the results shown in Table 5. Specifically, under normality ($\nu = \infty$) or mild non-normality (e.g., $\nu = 20, 30$), both T_F and T_{new} exhibit reasonable size control, while T_F has higher power, making it competitive with the proposed test T_{new}. However, under pronounced non-normality (e.g., $\nu = 6, 7$), T_F suffers from serious size distortion, whereas T_{new} maintains proper size control; in such regimes, the apparent high power of T_F is misleading, and T_{new} is clearly preferred.

Table 2. Empirical powers of the tests T_F, T_S, and T_{new} under Case 1.

p	N	Empirical power (I)			Empirical power (II)			Empirical power (III)		
		T_F	T_S	T_{new}	T_F	T_S	T_{new}	T_F	T_S	T_{new}
20	60	0.944	0.891	0.919	0.999	0.981	0.988	0.083	0.072	0.068
	80	0.988	0.970	0.987	1.000	1.000	1.000	0.118	0.102	0.092
	100	0.997	0.987	0.998	1.000	1.000	1.000	0.133	0.108	0.116
	120	1.000	0.999	1.000	1.000	1.000	1.000	0.152	0.123	0.136
	160	1.000	1.000	1.000	1.000	1.000	1.000	0.210	0.174	0.198
80	60	0.584	0.387	0.458	0.912	0.828	0.830	0.381	0.244	0.288
	80	0.766	0.558	0.699	0.985	0.948	0.972	0.540	0.367	0.469
	100	0.872	0.669	0.856	0.999	0.993	0.998	0.687	0.443	0.641
	120	0.955	0.782	0.946	1.000	0.999	1.000	0.798	0.577	0.783
	160	0.986	0.917	0.985	1.000	1.000	1.000	0.969	0.719	0.945
100	60	0.458	0.302	0.334	0.880	0.735	0.751	0.488	0.320	0.365
	80	0.642	0.424	0.574	0.968	0.909	0.936	0.686	0.457	0.611
	100	0.783	0.517	0.766	0.992	0.958	0.987	0.833	0.563	0.806
	120	0.904	0.654	0.879	1.000	0.993	1.000	0.924	0.676	0.912
	160	0.985	0.831	0.983	1.000	1.000	1.000	0.982	0.844	0.981
120	60	0.352	0.253	0.259	0.797	0.654	0.655	0.598	0.428	0.478
	80	0.549	0.323	0.486	0.946	0.842	0.872	0.786	0.557	0.726
	100	0.724	0.451	0.669	0.991	0.942	0.982	0.913	0.696	0.887
	120	0.815	0.566	0.799	0.998	0.976	0.996	0.960	0.787	0.958
	160	0.954	0.742	0.951	1.000	0.998	1.000	0.997	0.918	0.991
160	60	0.284	0.158	0.163	0.654	0.452	0.462	0.763	0.533	0.656
	80	0.412	0.246	0.316	0.878	0.706	0.791	0.910	0.711	0.876
	100	0.569	0.358	0.508	0.963	0.836	0.938	0.982	0.865	0.976
	120	0.695	0.406	0.666	0.993	0.912	0.981	0.992	0.919	0.989
	160	0.879	0.581	0.869	0.998	0.988	0.998	1.000	0.976	1.000

We remark that size and power results for non-normal Case 3, Case 4, and Case 6 can be obtained similarly, which are consistent with those reported for Case 2 and Case 5; they are therefore omitted for brevity. Overall, the proposed test demonstrates robust performance across various distributions in terms of both size and power. It maintains well-controlled empirical size while achieving power comparable to or better than the existing tests. It is therefore recommended accordingly.

Table 3. Empirical powers of the tests T_F, T_S, and T_{new} under Case 2.

p	N	Empirical power (I)			Empirical power (II)			Empirical power (III)		
		T_F	T_S	T_{new}	T_F	T_S	T_{new}	T_F	T_S	T_{new}
20	60	0.907	0.815	0.822	0.995	0.962	0.967	0.173	0.076	0.062
	80	0.986	0.927	0.941	0.998	0.996	0.998	0.195	0.102	0.082
	100	0.995	0.968	0.984	1.000	1.000	1.000	0.246	0.107	0.112
	120	0.999	0.987	0.996	1.000	1.000	1.000	0.254	0.128	0.131
	160	1.000	0.998	1.000	1.000	1.000	1.000	0.356	0.182	0.199
80	60	0.591	0.371	0.395	0.902	0.762	0.764	0.458	0.261	0.264
	80	0.769	0.514	0.639	0.986	0.926	0.938	0.624	0.374	0.421
	100	0.882	0.648	0.806	0.993	0.976	0.989	0.751	0.441	0.634
	120	0.933	0.742	0.888	1.000	0.995	1.000	0.829	0.556	0.741
	160	0.988	0.859	0.982	1.000	1.000	1.000	0.961	0.764	0.932
100	60	0.519	0.302	0.306	0.859	0.688	0.698	0.560	0.336	0.342
	80	0.678	0.428	0.546	0.976	0.879	0.899	0.726	0.445	0.567
	100	0.809	0.542	0.708	0.988	0.951	0.978	0.870	0.580	0.771
	120	0.895	0.632	0.813	0.998	0.981	0.996	0.923	0.668	0.879
	160	0.971	0.806	0.946	1.000	0.996	1.000	0.982	0.839	0.975
120	60	0.439	0.228	0.239	0.802	0.629	0.636	0.624	0.395	0.416
	80	0.622	0.343	0.458	0.928	0.788	0.829	0.839	0.556	0.701
	100	0.742	0.456	0.621	0.987	0.912	0.956	0.922	0.655	0.866
	120	0.816	0.556	0.726	0.997	0.952	0.988	0.975	0.747	0.950
	160	0.938	0.719	0.919	1.000	0.991	0.999	1.000	0.928	0.999
160	60	0.348	0.151	0.156	0.703	0.451	0.456	0.807	0.565	0.644
	80	0.472	0.269	0.308	0.856	0.683	0.716	0.932	0.734	0.866
	100	0.604	0.319	0.456	0.957	0.816	0.889	0.982	0.835	0.963
	120	0.736	0.425	0.616	0.981	0.902	0.958	0.995	0.902	0.995
	160	0.881	0.542	0.822	0.999	0.966	0.999	1.000	0.982	1.000

5. Conclusions

We have proposed a new sphericity test for high-dimensional covariance matrices, which does not require a normality condition, by constructing a consistent estimator related to the fourth moment. From a pedagogical perspective, this work also responds to a common learning difficulty for students, as high-dimensional and non-normal data are commonly encountered in course study and practical training, representing a challenge in real-world applications. In this way, our work serves as a complement and extension to standard textbook content and existing literature, offering a deeper exploration of hypothesis testing for covariance matrices in probability and statistics courses.

A key technical condition underlying our theoretical framework is that the data structure satisfies Assumption 2, in which the components $\{w_{ij}\}$ are independent and identically distributed. This assumption is widely adopted in the high-dimensional sphericity testing literature (e.g., [14, 15, 17]), which enables the use of random matrix theory and moment methods to derive the asymptotic distribution of test statistics when both p and N tend to infinity. Without such a condition, theoretical derivations may be intractable in high-dimensional settings. As a preliminary empirical assessment of robustness against this assumption, we conducted some numerical simulations under mild violations (e.g., some components are normal while others are non-normal). The results seem to indicate that the proposed test T_{new} maintains reasonable size control, at least in the settings considered. We therefore

Table 4. Empirical powers of the tests T_F, T_S, and T_{new} under Case 5.

p	N	Empirical power (I)			Empirical power (II)			Empirical power (III)		
		T_F	T_S	T_{new}	T_F	T_S	T_{new}	T_F	T_S	T_{new}
20	60	0.915	0.761	0.766	0.995	0.961	0.966	0.217	0.078	0.065
	80	0.973	0.920	0.930	1.000	0.991	0.996	0.287	0.105	0.096
	100	0.997	0.969	0.986	1.000	0.999	0.999	0.311	0.124	0.127
	120	0.999	0.987	0.997	1.000	1.000	1.000	0.362	0.135	0.139
	160	1.000	0.999	0.999	1.000	1.000	1.000	0.417	0.187	0.205
80	60	0.652	0.399	0.421	0.916	0.716	0.728	0.488	0.246	0.261
	80	0.779	0.498	0.596	0.978	0.908	0.913	0.676	0.340	0.442
	100	0.868	0.587	0.760	0.995	0.970	0.978	0.813	0.434	0.606
	120	0.939	0.720	0.871	0.999	0.989	0.995	0.881	0.545	0.762
	160	0.984	0.862	0.961	1.000	0.999	1.000	0.969	0.746	0.921
100	60	0.552	0.286	0.292	0.852	0.692	0.699	0.606	0.327	0.344
	80	0.724	0.431	0.516	0.963	0.835	0.867	0.791	0.451	0.556
	100	0.851	0.538	0.714	0.991	0.938	0.974	0.898	0.588	0.776
	120	0.894	0.625	0.791	0.998	0.976	0.989	0.931	0.674	0.873
	160	0.974	0.788	0.934	1.000	0.998	0.999	0.995	0.879	0.981
120	60	0.476	0.224	0.244	0.820	0.610	0.628	0.687	0.411	0.438
	80	0.663	0.376	0.431	0.935	0.789	0.798	0.859	0.533	0.688
	100	0.775	0.419	0.585	0.977	0.885	0.935	0.935	0.671	0.862
	120	0.844	0.527	0.724	0.998	0.956	0.986	0.984	0.779	0.950
	160	0.956	0.704	0.905	1.000	0.991	0.999	1.000	0.926	0.998
160	60	0.382	0.161	0.167	0.729	0.461	0.468	0.804	0.516	0.565
	80	0.541	0.259	0.313	0.858	0.640	0.683	0.946	0.712	0.852
	100	0.682	0.327	0.466	0.960	0.807	0.880	0.979	0.832	0.952
	120	0.762	0.396	0.564	0.984	0.882	0.951	0.999	0.918	0.987
	160	0.871	0.566	0.784	0.999	0.972	0.998	1.000	0.976	0.999

Table 5. Empirical sizes and powers under Student's t distribution with varying degrees of freedom for $p = 100$ and $N = 100$.

ν	Δ	Empirical size		Power (I)		Power (II)		Power (III)	
		T_F	T_{new}	T_F	T_{new}	T_F	T_{new}	T_F	T_{new}
∞	0.00	0.058	0.046	0.802	0.756	0.996	0.992	0.818	0.790
30	0.23	0.058	0.041	0.801	0.738	0.991	0.984	0.816	0.765
20	0.38	0.057	0.051	0.795	0.743	0.993	0.985	0.854	0.785
15	0.55	0.066	0.046	0.809	0.729	0.994	0.988	0.839	0.781
10	1.00	0.102	0.042	0.804	0.712	0.992	0.979	0.846	0.752
8	1.50	0.151	0.056	0.806	0.678	0.994	0.968	0.855	0.761
7	2.00	0.198	0.055	0.824	0.658	0.992	0.962	0.891	0.762
6	3.00	0.262	0.068	0.842	0.642	0.992	0.952	0.916	0.754

conjecture that Assumption 2 can be relaxed in some sense. Nevertheless, a complete theoretical analysis under the relaxed assumption is challenging and lies beyond the scope of this paper. This is therefore left for future exploration.

Use of AI tools declaration

No Artificial Intelligence (AI) tools are used in the creation of this article.

Acknowledgements

The author is very grateful to the reviewers for their detailed comments and valuable suggestions, which greatly improved the manuscript.

Conflict of interest

There are no conflicts of interest.

References

1. T. Pavlenko, A. Björkström, A. Tillander, Covariance structure approximation via gLasso in high-dimensional supervised classification, *J. Appl. Stat.*, **39** (2012), 1643–1666. https://doi.org/10.1080/02664763.2012.663346
2. L. Huang, Y. Xiao, K. Liu, H. So, J. Zhang, Bayesian Information criterion for source enumeration in large-scale adaptive antenna array, *IEEE Trans. Veh. Technol.*, **65** (2016), 3018–3032. https://doi.org/10.1109/TVT.2015.2436060
3. M. Tang, Y. Rong, A. Maio, C. Chen, J. Zhou, Adaptive radar detection in gaussian disturbance with structured covariance matrix via invariance theory, *IEEE Trans. Signal Proces.*, **67** (2019), 5671–5685. https://doi.org/10.1109/TSP.2019.2941119
4. Y. Rong, A. Aubry, A. Maio, M. Tang, Adaptive radar detection in low-rank heterogeneous clutter via invariance theory, *IEEE Trans. Signal Proces.*, **69** (2021), 1492–1506. https://doi.org/10.1109/TSP.2021.3058447
5. C. Chen, M. Tang, Y. Rong, Detection of a rare multichannel gaussian signal via higher criticism, *IEEE Signal Process. Lett.*, **29** (2022), 2063–2067. https://doi.org/10.1109/LSP.2022.3212631
6. X. Wang, B. Zhang, Target selection in shrinkage estimation of covariance matrix: A structural similarity approach, *Stat. Probab. Lett.*, **208** (2024), 110048. https://doi.org/10.1016/j.spl.2023.110048
7. T. Anderson, *An Introduction to Multivariate Statistical Analysis, 3rd Edition*, WILEY, 2003.
8. S. John, Some optimal multivariate tests, *Biometrika*, **58** (1971), 123–127. https://doi.org/10.1093/BIOMET/58.1.123
9. S. John, The distribution of a statistic used for testing sphericity of normal distributions, *Biometrika*, **59** (1972), 169–173. https://doi.org/10.1093/biomet/59.1.169
10. O. Ledoit, M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, *Ann. Stat.*, **4** (2002), 1081–1102. https://doi.org/10.1214/aos/1031652392
11. M. Birke, H. Dette, A note on testing the covariance matrix for large dimension, *Stat. Probab. Lett.*, **74** (2005), 281–289. https://doi.org/10.17877/DE290R-15067
12. M. Srivastava, Some tests concerning the covariance matrix in high dimensional data, *J. Jpn. Stat. Soc.*, **35** (2005), 251–272. https://doi.org/10.14490/jjss.35.251

13. T. Fisher, X. Sun, C. Gallagher, A new test for sphericity of the covariance matrix for high dimensional data, *J. Multivar. Anal.*, **101** (2010), 2554–2570. https://doi.org/10.1016/J.JMVA.2010.07.004
14. M. Srivastava, H. Yanagihara, T. Kubokawa, Tests for covariance matrices in high dimension with less sample size, *J. Multivar. Anal.*, **130** (2014), 289–309. https://doi.org/10.1016/j.jmva.2014.06.003
15. S. Chen, L. Zhang, P. Zhong, Tests for high-dimensional covariance matrices, *J. Am. Stat. Assoc.*, **105** (2010), 810–819. https://doi.org/10.1198/jasa.2010.tm09560
16. Q. Wang, J. Yao, On the sphericity test with large-dimensional observations, *Electron. J. Stat.*, **7** (2013), 2164–2192. https://doi.org/10.1214/13-EJS842
17. X. Tian, Y. Lu, W. Li, A robust test for sphericity of high-dimensional covariance matrices, *J. Multivar. Anal.*, **141** (2015), 217–227. https://doi.org/10.1016/j.jmva.2015.07.010
18. S. Yuan, J. Zhou, J. Pan, J. Shen, Sphericity and identity test for high-dimensional covariance matrix using random matrix theory, *Acta Math. Appl. Sin. E.*, **37** (2021), 214–231. https://doi.org/10.1007/s10255-021-1004-1
19. T. Fisher, On testing for an identity covariance matrix when the dimensionality equals or exceeds the sample size, *J. Stat. Plann. Inference*, **142** (2012), 312–326. https://doi.org/10.1016/j.jspi.2011.07.019
20. K. Xu, Testing diagonality of high-dimensional covariance matrix under non-normality, *J. Stat. Comput. Simul.*, **87** (2017), 3208–3224. https://doi.org/10.1080/00949655.2017.1322630
21. T. Himeno, T. Yamada, Estimations for some functions of covariance matrix in high dimension under non-normality and its applications, *J. Multivar. Anal.*, **130** (2014), 27–44. https://doi.org/10.1016/j.jmva.2014.04.014
22. M. Lopes, A. Blandino, A. Aue, Bootstrapping spectral statistics in high dimensions, *Biometrika*, **106** (2019), 781–801. https://doi.org/10.1093/biomet/asz031
23. N. Dörnemann, H. Dette, Sequential eigenvalue statistics for change-point detection in covariance matrices, preprint, arXiv:2409.15588. https://doi.org/10.48550/arXiv.2409.15588

AIMS Press